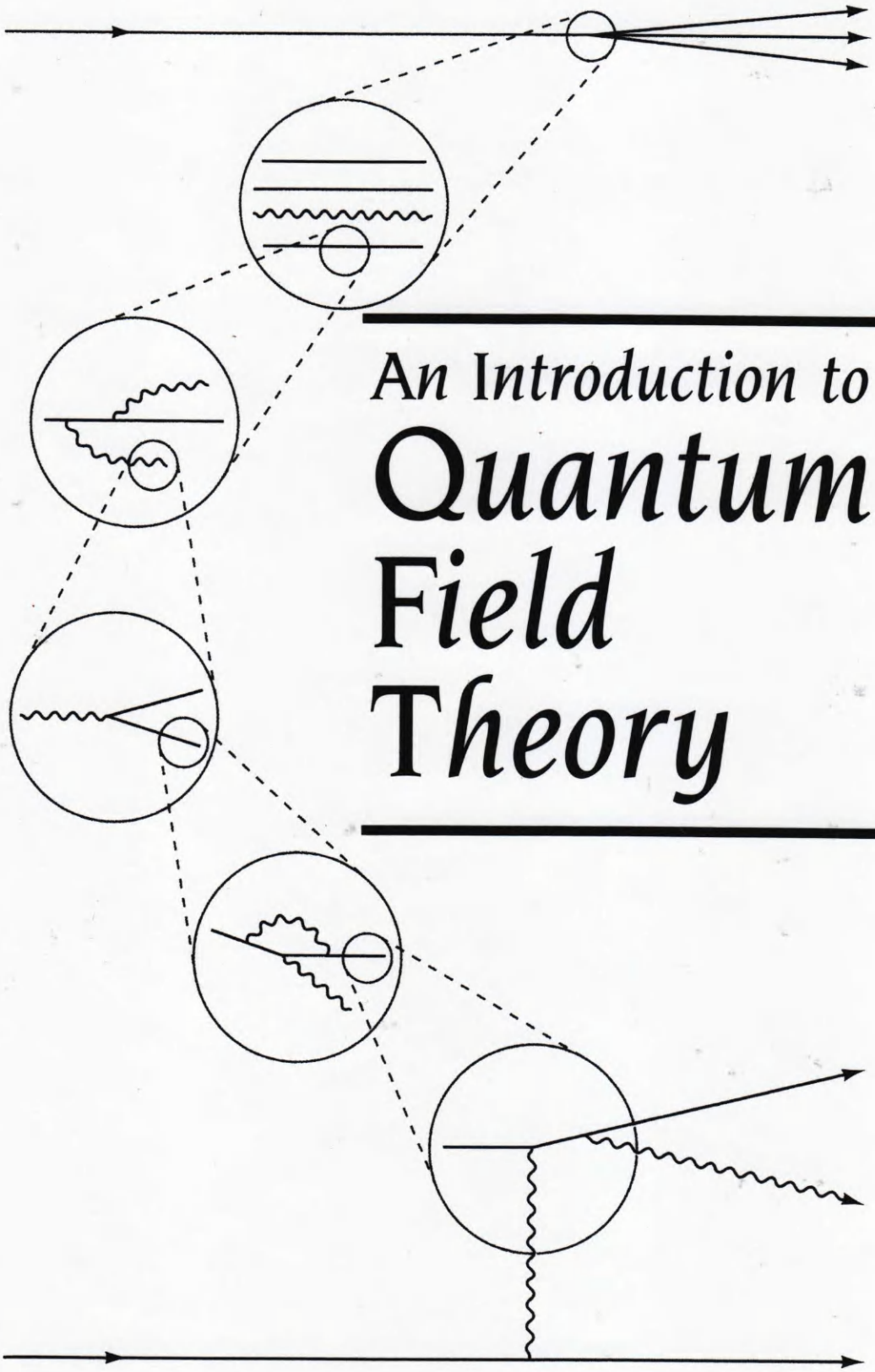


An Introduction to Quantum Field Theory

Michael E. Peskin ♦ Daniel V. Schroeder

ABP

An Introduction to Quantum Field Theory

An Introduction to Quantum Field Theory

Michael E. Peskin
Stanford Linear Accelerator Center

Daniel V. Schroeder
Weber State University

THE ADVANCED BOOK PROGRAM

A Member of the Perseus Books Group

Figures 5.2 and 14.2 have been adapted, with permission, from figures in *Physical Review Letters.*
Table 13.1 has been adapted, with permission, from a set of tables in *Physical Review B.*
Table 17.1 and Figures 5.3, 5.7, 7.9, 17.13, 17.23, and 18.8 have been adapted, with permission, from tables/figures in *Physical Review D.*
Figure 17.5 has been adapted, with permission, from a table in *Zeitschrift für Physik C.*
Table 20.1 and Figures 17.22 and 20.5 have been reprinted, with permission, from *Physical Review D.*

In all cases, the specific source is cited in the caption of the figure or table.

Library of Congress Cataloging-in-Publication Data
Peskin, Michael Edward, 1951–
Introduction to quantum field theory / Michael E. Peskin, Daniel V. Schroeder.
p. cm.
Includes bibliographical references and index.
ISBN-13 978-0-201-50397-5 ISBN-10 0-201-50397-2
1. Quantum field theory. 2. Feynman diagrams. 3. Renormalization (Physics) 4. Gauge fields (Physics) I. Schroeder, Daniel V. II. Title.
QC174.45.P465 1995
530. 1 43–dc20 89-27170
CIP

Westview Press is a member of the Perseus Books Group
Find us on the World Wide Web at http://www.westviewpress.com

Cover art by Michael E. Peskin and Daniel V. Schroeder
Cover design by Lynne Reed
Typeset by Daniel V. Schroeder in 10 pt. Computer Modern using TEXtures for the Macintosh.

EBA 21 22 08 07 06

Contents

Part II: Renormalization

Part III: Non-Abelian Gauge Theories

Epilogue

*These sections may be omitted in a one-year course emphasizing the less formal aspects of elementary particle physics.

Preface

Quantum field theory is a set of ideas and tools that combines three of the major themes of modern physics: the quantum theory, the field concept, and the principle of relativity. Today, most working physicists need to know some quantum field theory, and many others are curious about it. The theory underlies modern elementary particle physics, and supplies essential tools to nuclear physics, atomic physics, condensed matter physics, and astrophysics. In addition, quantum field theory has led to new bridges between physics and mathematics.

One might think that a subject of such power and widespread application would be complex and difficult. In fact, the central concepts and techniques of quantum field theory are quite simple and intuitive. This is especially true of the many pictorial tools (Feynman diagrams, renormalization group flows, and spaces of symmetry transformations) that are routinely used by quantum field theorists. Admittedly, these tools take time to learn, and tying the subject together with rigorous proofs can become extremely technical. Nevertheless, we feel that the basic concepts and tools of quantum field theory can be made accessible to all physicists, not just an elite group of experts.

A number of earlier books have succeeded in making parts of this subject accessible to students. The best known of these is the two-volume text written in the 1960s by our Stanford colleagues Bjorken and Drell. In our opinion, their text contains an ideal mixture of abstract formalism, intuitive explanations, and practical calculations, all presented with great care and clarity. Since the 1960s, however, the subject of quantum field theory has developed enormously, both in its conceptual framework (the renormalization group, new types of symmetries) and in its areas of application (critical exponents in condensed matter systems, the standard model of elementary particle physics). It is long overdue that a textbook of quantum field theory should appear that provides a complete survey of the subject, including these newer developments, yet still offers the same accessibility and depth of treatment as Bjorken and Drell. We have written this book with that goal in mind.

An Outline of the Book

This textbook is composed of three major sections. The first is mainly concerned with the quantum theory of electromagnetism, which provided the first example of a quantum field theory with direct experimental applications. The third part of the book is mainly concerned with the particular quantum field theories that appear in the standard model of particle interactions. The second part of the book is a bridge between these two subjects; it is intended to introduce some of the very deep concepts of quantum field theory in a context that is as straightforward as possible.

Part I begins with the study of fields with linear equations of motion, that is, fields without interactions. Here we explore the combined implications of quantum mechanics and special relativity, and we learn how particles arise as the quantized excitations of fields. We then introduce interactions among these particles and develop a systematic method of accounting for their effects. After this introduction, we carry out explicit computations in the quantum theory of electromagnetism. These illustrate both the special features of the behavior of electrons and photons and some general aspects of the behavior of interacting quantum fields.

In several of the calculations in Part I, naive methods lead to infinite results. The appearance of infinities is a well-known feature of quantum field theory. At times, it has been offered as evidence for the inconsistency of quantum field theory (though a similar argument could be made against the classical electrodynamics of point particles). For a long time, it was thought sufficient to organize calculations in such a way that no infinities appear in quantities that can be compared directly to experiment. However, one of the major insights of the more recent developments is that these formal infinities actually contain important information that can be used to predict the qualitative behavior of a system. In Part II of the book, we develop this theory of infinities systematically. The development makes use of an analogy between quantum-mechanical and thermal fluctuations, which thus becomes a bridge between quantum field theory and statistical mechanics. At the end of Part II we discuss applications of quantum field theory to the theory of phase transitions in condensed matter systems.

Part III deals with the generalizations of quantum electrodynamics that have led to successful models of the forces between elementary particles. To derive these generalizations, we first analyze and generalize the fundamental symmetry of electrodynamics, then work out the consequences of quantizing a theory with this generalized symmetry. This analysis leads to intricate and quite nontrivial applications of the concepts introduced earlier. We conclude Part III with a presentation of the standard model of particle physics and a discussion of some of its experimental tests.

The Epilogue to the book discusses qualitatively the frontier areas of research in quantum field theory and gives references that can guide a student to the next level of study.

Where a choice of viewpoints is possible, we have generally chosen to explain ideas in language appropriate to the applications to elementary particle physics. This choice reflects our background and research interests. It also reflects our strongly held opinion, even in this age of intellectual relativism, that there is something special about unraveling the behavior of Nature at the deepest possible level. We are proud to take as our subject the structure of the fundamental interactions, and we hope to convey to the reader the grandeur and continuing vitality of this pursuit.

How to Use This Book

This book is an *introduction* to quantum field theory. By this we mean, first and foremost, that we assume no prior knowledge of the subject on the part of the reader. The level of this book should be appropriate for students taking their first course in quantum field theory, typically during the second year of graduate school at universities in the United States. We assume that the student has completed graduate-level courses in classical mechanics, classical electrodynamics, and quantum mechanics. In Part II we also assume some knowledge of statistical mechanics. It is not necessary to have mastered every topic covered in these courses, however. Crucially important prerequisites include the Lagrangian and Hamiltonian formulations of dynamics, the relativistic formulation of electromagnetism using tensor notation, the quantization of the harmonic oscillator using ladder operators, and the theory of scattering in nonrelativistic quantum mechanics. Mathematical prerequisites include an understanding of the rotation group as applied to the quantum mechanics of spin, and some facility with contour integration in the complex plane.

Despite being an "introduction", this book is rather lengthy. To some extent, this is due to the large number of explicit calculations and worked examples in the text. We must admit, however, that the total number of topics covered is also quite large. Even students specializing in elementary particle theory will find that their first research projects require only a part of this material, together with additional, specialized topics that must be gleaned from the research literature. Still, we feel that students who want to become experts in elementary particle theory, and to fully understand its unified view of the fundamental interactions, should master every topic in this book. Students whose main interest is in other fields of physics, or in particle experimentation, may opt for a much shorter "introduction", omitting several chapters.

The senior author of this book once did succeed in "covering" 90% of its content in a one-year lecture course at Stanford University. But this was a mistake; at such a pace, there is not enough time for students of average preparation to absorb the material. Our saner colleagues have found it more reasonable to cover about one Part of the book per *semester*. Thus, in planning a one-year course in quantum field theory, they have chosen either to reserve

Part III for study at a more advanced level or to select about half of the material from Parts II and III, leaving the rest for students to read on their own.

We have designed the book so that it can be followed from cover to cover, introducing all of the major ideas in the field systematically. Alternatively, one can follow an accelerated track that emphasizes the less formal applications to elementary particle physics and is sufficient to prepare students for experimental or phenomenological research in that field. Sections that can be omitted from this accelerated track are marked with an asterisk in the Table of Contents; none of the unmarked sections depend on this more advanced material. Among the unmarked sections, the order could also be varied somewhat: Chapter 10 does not depend on Chapters 8 and 9; Section 11.1 is not needed until just before Chapter 20; and Chapters 20 and 21 are independent of Chapter 17.

Those who wish to study some, but not all, of the more advanced sections should note the following table of dependencies:

Before reading . . .	one should read all of . . .
Chapter 13	Chapters 11, 12
Section 16.6	Chapter 11
Chapter 18	Sections 12.4, 12.5, 16.5
Chapter 19	Sections 9.6, 15.3
Section 19.5	Section 16.6
Section 20.3	Sections 19.1–19.4
Section 21.3	Chapter 11

Within each chapter, the sections marked with an asterisk should be read sequentially, except that Sections 16.5 and 16.6 do not depend on 16.4.

A student whose main interest is in statistical mechanics would want to read the book sequentially, confronting the deep formal issues of Part II but ignoring most of Part III, which is mainly of significance to high-energy phenomena. (However, the material in Chapters 15 and 19, and in Section 20.1, does have beautiful applications in condensed matter physics.)

We emphasize to all students the importance of working actively with the material while studying. It probably is not possible to understand any section of this book without carefully working out the intermediate steps of every derivation. In addition, the problems at the end of each chapter illustrate the general ideas and often apply them in nontrivial, realistic contexts. However, the most illustrative exercises in quantum field theory are too long for ordinary homework problems, being closer to the scale of small research projects. We have provided one of these lengthy problems, broken up into segments with hints and guidance, at the end of each of the three Parts of the book. The volume of time and paper that these problems require will be well invested.

At the beginning of each Part we have included a brief "Invitation" chapter, which previews some of the upcoming ideas and applications. Since these

chapters are somewhat easier than the rest of the book, we urge all students to read them.

What This Book is Not

Although we hope that this book will provide a thorough grounding in quantum field theory, it is in no sense a complete education. A dedicated student of physics will want to supplement our treatment in many areas. We summarize the most important of these here.

First of all, this is a book about theoretical methods, not a review of observed phenomena. We do not review the crucial experiments that led to the standard model of elementary particle physics or discuss in detail the more recent experiments that have confirmed its predictions. Similarly, in the chapters that deal with applications to statistical mechanics, we do not discuss the beautiful and varied experiments on phase transitions that led to the confirmation of field theory models. We strongly encourage the student to read, in parallel with this text, a modern presentation of the experimental development of each of these fields.

Although we present the elementary aspects of quantum field theory in full detail, we state some of the more advanced results without proof. For example, it is known rigorously, to all orders in the standard expansion of quantum electrodynamics, that formal infinities can be removed from all experimental predictions. This result, known as *renormalizability*, has important consequences, which we explore in Part II. We do not present the general proof of renormalizability. However, we do demonstrate renormalizability explicitly in illustrative, low-order computations, we discuss intuitively the issues that arise in the complete proof, and we give references to a more complete demonstration. More generally, we have tried to motivate the most important results (usually through explicit examples) while omitting lengthy, purely technical derivations.

Any introductory survey must classify some topics as beyond its scope. Our philosophy has been to include what can be learned about quantum field theory by considering weakly interacting particles and fields, using series expansions in the strength of the interaction. It is amazing how much insight one can obtain in this way. However, this definition of our subject leaves out the theory of bound states, and also phenomena associated with nontrivial solutions to nonlinear field equations. We give a more complete listing of such advanced topics in the Epilogue.

Finally, we have not attempted in this book to give an accurate record of the history of quantum field theory. Students of physics do need to understand the history of physics, for a number of reasons. The most important is to acquire a precise understanding of the experimental basis of the subject. A second important reason is to gain an idea of how science progresses as a human endeavor, how ideas develop as small steps taken by individuals to

become the major achievements of the community as a whole.*

In this book we have not addressed either of these needs. Rather, we have included only the kind of mythological history whose purpose is to motivate new ideas and assign names to them. A principle of physics usually has a name that has been assigned according to the community's consensus on who deserves credit for its development. Usually the real credit is only partial, and the true historical development is quite complex. But the clear assignment of names is essential if physicists are to communicate with one another.

Here is one example. In Section 17.5 we discuss a set of equations governing the structure of the proton, which are generally known as the Altarelli-Parisi equations. Our derivation uses a method due to Gribov and Lipatov (GL). The original results of GL were rederived in a more abstract language by Christ, Hasslacher, and Mueller (CHM). After the discovery of the correct fundamental theory of the strong interactions (QCD), Georgi, Politzer, Gross, and Wilczek (GPGW) used the technique of CHM to derive formal equations for the variation of the proton structure. Parisi gave the first of a number of independent derivations that converted these equations into a useful form. The combination of his work with that of GPGW gives the derivation of the equations that we present in Section 18.5. Dokhshitzer later obtained these equations more simply by direct application of the method of GL. Sometime later, but independently, Altarelli and Parisi obtained these equations again by the same route. These last authors also popularized the technique, explaining it very clearly, encouraging experimentalists to use the equations in interpreting their data, and prodding theorists to compute the systematic higher-order corrections to this picture. In Section 17.5 we have presented the shortest path to the end of this convoluted historical road and hung the name 'Altarelli-Parisi' on the final result.

There is a fourth reason for students to read the history of physics: Often the original breakthrough papers, though lacking a textbook's advantages of hindsight, are filled with marvelous personal insights. We strongly encourage students to go back to the original literature whenever possible and see what the creators of the field had in mind. We have tried to aid such students with references provided in footnotes. Though occasionally we refer to papers merely to give credit, most of the references are included because we feel the reader should not miss the special points of view that the authors put forward.

---

*The history of the development of quantum field theory and particle physics has recently been reviewed and debated in a series of conference volumes: *The Birth of Particle Physics*, L. M. Brown and L. Hoddeson, eds. (Cambridge University Press, 1983); *Pions to Quarks*, L. M. Brown, M. Dresden, and L. Hoddeson, eds. (Cambridge University Press, 1989); and *The Rise of the Standard Model*, L. M. Brown, M. Dresden, L. Hoddeson, and M. Riordan, eds. (Cambridge University Press, 1995). The early history of quantum electrodynamics is recounted in a fascinating book by Schweber (1994).

Acknowledgments

The writing of this book would not have been possible without the help and encouragement of our many teachers, colleagues, and friends. It is our great pleasure to thank these people.

Michael has been privileged to learn field theory from three of the subject's contemporary masters—Sidney Coleman, Steven Weinberg, and Kenneth Wilson. He is also indebted to many other teachers, including Sidney Drell, Michael Fisher, Kurt Gottfried, John Kogut, and Howard Georgi, and to numerous co-workers, in particular, Orlando Alvarez, John Preskill, and Edward Witten. His association with the laboratories of high-energy physics at Cornell and Stanford, and discussions with such people as Gary Feldman, Martin Perl, and Morris Swartz, have also shaped his viewpoint. To these people, and to many other people who have taught him points of physics over the years, Michael expresses his gratitude.

Dan has learned field theory from Savas Dimopoulos, Leonard Susskind, Richard Blankenbecler, and many other instructors and colleagues, to whom he extends thanks. For his broader education in physics he is indebted to many teachers, but especially to Thomas Moore. In addition, he is grateful to all the teachers and friends who have criticized his writing over the years.

For their help in the construction and writing of this book, many people are due more specific thanks. This book originated as graduate courses taught at Cornell and at Stanford, and we thank the students in these courses for their patience, criticism, and suggestions. During the years in which this book was written, we were supported by the Stanford Linear Accelerator Center and the Department of Energy, and Dan also by the Pew Charitable Trusts and the Departments of Physics at Pomona College, Grinnell College, and Weber State University. We are especially grateful to Sidney Drell for his support and encouragement. The completion of this book was greatly assisted by a sabbatical Michael spent at the University of California, Santa Cruz, and he thanks Michael Nauenberg and the physics faculty and students there for their gracious hospitality.

We have circulated preliminary drafts of this book widely and have been continually encouraged by the response to these drafts and the many suggestions and corrections they have elicited. Among the people who have helped us in this way are Curtis Callan, Edward Farhi, Jonathan Feng, Donald Finnell, Paul Frampton, Jeffrey Goldstone, Brian Hatfield, Barry Holstein, Gregory Kilcup, Donald Knuth, Barak Kol, Phillip Nelson, Matthias Neubert, Yossi Nir, Arvind Rajaraman, Pierre Ramond, Katsumi Tanaka, Xerxes Tata, Arkady Vainshtein, Rudolf von Bunau, Shimon Yankielowicz, and Eran Yehudai. For help with more specific technical problems or queries, we are grateful to Michael Barnett, Marty Breidenbach, Don Groom, Toichiro Kinoshita, Paul Langacker, Wu-Ki Tung, Peter Zerwas, and Jean Zinn-Justin. We thank our Stanford colleagues James Bjorken, Stanley Brodsky, Lance Dixon, Bryan Lynn, Helen Quinn, Leonard Susskind, and Marvin Weinstein,

on whom we have tested many of the explanations given here. We are especially grateful to Michael Dine, Martin Einhorn, Howard Haber, and Renata Kallosh, and to their long-suffering students, who have used drafts of this book as textbooks over a period of several years and have brought us back important and illuminating criticism.

We pause to remember two friends of this book who did not see its completion—Donald Yennie, whose long career made him one of the heroes of Quantum Electrodynamics, and Brian Warr, whose brilliant promise in theoretical physics was cut short by AIDS.

We thank our editors at Addison-Wesley—Allan Wylde, Barbara Holland, Jack Repcheck, Heather Mimnaugh, and Jeffrey Robbins—for their encouragement of this project, and for not losing faith that it would be completed. We are grateful to Mona Zeftel and Lynne Reed for their advice on the formatting of this book, and to Jeffrey Towson and Cristine Jennings for their skillful preparation of the illustrations. The final typeset pages were produced at the WSU Printing Services Department, with much help from Allan Davis. Of course, the production of this book would have been impossible without the many individuals who have kept our computers and network links up and running.

We thank the many friends whose support we have relied on during the long course of this project. We are particularly grateful to our parents, Dorothy and Vernon Schroeder and Gerald and Pearl Peskin. Michael also thanks Valerie, Thomas, and Laura for their love and understanding.

Finally, we thank you, the reader, for your time and effort spent studying this book. Though we have tried to cleanse this text of conceptual and typographical errors, we apologize in advance for those that have slipped through. We will be glad to hear your comments and suggestions for further improvements in the presentation of quantum field theory.

Michael E. Peskin
Daniel V. Schroeder

Notations and Conventions

Units

We will work in "God-given" units, where

$$\hbar = c = 1.$$

In this system,

$$[\text{length}] = [\text{time}] = [\text{energy}]^{-1} = [\text{mass}]^{-1}.$$

The mass (m) of a particle is therefore equal to its rest energy (mc^2), and also to its inverse Compton wavelength ($mc/\hbar$). For example,

$$m_{\text{electron}} = 9.109 \times 10^{-28}\,\text{g} = 0.511\,\text{MeV} = \left(3.862 \times 10^{-11}\,\text{cm}\right)^{-1}.$$

A selection of other useful numbers and conversion factors is given in the Appendix.

Relativity and Tensors

Our conventions for relativity follow Jackson (1975), Bjorken and Drell (1964, 1965), and nearly all recent field theory texts. We use the metric tensor

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

with Greek indices running over 0, 1, 2, 3 or t, x, y, z. Roman indices—i, j, etc.—denote only the three spatial components. Repeated indices are summed in all cases. Four-vectors, like ordinary numbers, are denoted by light italic type; three-vectors are denoted by boldface type; unit three-vectors are denoted by a light italic label with a hat over it. For example,

$$x^\mu = (x^0, \mathbf{x}), \qquad x_\mu = g_{\mu\nu}x^\nu = (x^0, -\mathbf{x});$$

$$p \cdot x = g_{\mu\nu}p^\mu x^\nu = p^0 x^0 - \mathbf{p} \cdot \mathbf{x}.$$

A massive particle has

$$p^2 = p^\mu p_\mu = E^2 - |\mathbf{p}|^2 = m^2.$$

Note that the displacement vector x^μ is "naturally raised", while the derivative operator,

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \Big(\frac{\partial}{\partial x^0}, \boldsymbol{\nabla}\Big),$$

is "naturally lowered".

We define the totally antisymmetric tensor $\epsilon^{\mu\nu\rho\sigma}$ so that

$$\epsilon^{0123} = +1.$$

Be careful, since this implies $\epsilon_{0123} = -1$ and $\epsilon^{1230} = -1$. (This convention agrees with Jackson but not with Bjorken and Drell.)

Quantum Mechanics

We will often work with the Schrödinger wavefunctions of single quantum-mechanical particles. We represent the energy and momentum operators acting on such wavefunctions following the usual conventions:

$$E = i\frac{\partial}{\partial x^0}, \qquad \mathbf{p} = -i\boldsymbol{\nabla}.$$

These equations can be combined into

$$p^\mu = i\partial^\mu;$$

raising the index on ∂^μ conveniently accounts for the minus sign. The plane wave $e^{-ik\cdot x}$ has momentum k^μ, since

$$i\partial^\mu\big(e^{-ik\cdot x}\big) = k^\mu\, e^{-ik\cdot x}.$$

The notation 'h.c.' denotes the Hermitian conjugate.

Discussions of spin in quantum mechanics make use of the Pauli sigma matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Products of these matrices satisfy the identity

$$\sigma^i\sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k.$$

It is convenient to define the linear combinations $\sigma^\pm = \frac{1}{2}(\sigma^1 \pm i\sigma^2)$; then

$$\sigma^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \sigma^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Fourier Transforms and Distributions

We will often make use of the Heaviside step function $\theta(x)$ and the Dirac delta function $\delta(x)$, defined as follows:

$$\theta(x) = \begin{cases} 0 & x < 0, \\ 1 & x > 0; \end{cases} \qquad \delta(x) = \frac{d}{dx}\theta(x).$$

The delta function in n dimensions, denoted $\delta^{(n)}(\mathbf{x})$, is zero everywhere except at $\mathbf{x} = 0$ and satisfies

$$\int d^n x\, \delta^{(n)}(\mathbf{x}) = 1.$$

In Fourier transforms the factors of 2π will always appear with the momentum integral. For example, in four dimensions,

$$f(x) = \int \frac{d^4k}{(2\pi)^4}\, e^{-ik\cdot x}\, \tilde{f}(k);$$

$$\tilde{f}(k) = \int d^4x\, e^{ik\cdot x}\, f(x).$$

(In three-dimensional transforms the signs in the exponents will be + and −, respectively.) The tilde on $\tilde{f}(k)$ will sometimes be omitted when there is no potential for confusion. The other important factors of 2π to remember appear in the identity

$$\int d^4x\, e^{ik\cdot x} = (2\pi)^4 \delta^{(4)}(k).$$

Electrodynamics

We use the Heaviside-Lorentz conventions, in which the factors of 4π appear in Coulomb's law and the fine-structure constant rather than in Maxwell's equations. Thus the Coulomb potential of a point charge Q is

$$\Phi = \frac{Q}{4\pi r},$$

and the fine-structure constant is

$$\alpha = \frac{e^2}{4\pi} = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137}.$$

The symbol e stands for the charge of the electron, a negative quantity (although the sign rarely matters). We generally work with the relativistic form of Maxwell's equations:

$$\epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} = 0, \qquad \partial_\mu F^{\mu\nu} = ej^\nu,$$

where

$$A^\mu = (\Phi, \mathbf{A}), \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

and we have extracted the e from the 4-vector current density j^μ.

Dirac Equation

Some of our conventions differ from those of Bjorken and Drell (1964, 1965) and other texts: We use a chiral basis for Dirac matrices, and relativistic normalization for Dirac spinors. These conventions are introduced in Sections 3.2 and 3.3, and are summarized in the Appendix.

Editor's Foreword

The problem of communicating in a coherent fashion recent developments in the most exciting and active fields of physics continues to be with us. The enormous growth in the number of physicists has tended to make the familiar channels of communication considerably less effective. It has become increasingly difficult for experts in a given field to keep up with the current literature; the novice can only be confused. What is needed is both a consistent account of a field and the presentation of a definite "point of view" concerning it. Formal monographs cannot meet such a need in a rapidly developing field, while the review article seems to have fallen into disfavor. Indeed, it would seem that the people most actively engaged in developing a given field are the people least likely to write at length about it.

Frontiers in Physics was conceived in 1961 in an effort to improve the situation in several ways. Leading physicists frequently give a series of lectures, a graduate seminar, or a graduate course in their special fields of interest. Such lectures serve to summarize the present status of a rapidly developing field and may well constitute the only coherent account available at the time. Often, notes on lectures exist (prepared by the lecturer, by graduate students, or by postdoctoral fellows) and are distributed in photocopied form on a limited basis. One of the principal purposes of the *Frontiers in Physics* series is to make such notes available to a wider audience of physicists.

As *Frontiers in Physics* has evolved, a second category of book, the informal text/monograph, an intermediate step between lecture notes and formal texts or monographs, has played an increasingly important role in the series. In an informal text or monograph an author has reworked his/her lecture notes to the point at which the manuscript represents a coherent summation of a newly developed field, complete with references and problems, suitable for either classroom teaching or individual study.

During the past two decades significant advances have been made in both the conceptual framework of quantum field theory and its application to condensed matter physics and elementary particle physics. Given the fact that the study of quantum field theory has become an essential part of the education of graduate students in physics, a textbook which makes these recent developments accessible to the novice, while not neglecting the basic concepts, is highly desirable. Michael Peskin and Daniel Schroeder have written just such a book, describing in lucid fashion quantum electrodynamics, renormalization, and non-Abelian gauge theories while offering the reader a taste of what is to come. It is therefore quite appropriate to include this very polished text/monograph in the *Frontiers in Physics* series, and it gives me pleasure to welcome them to the ranks of its authors.

Aspen, Colorado
August 1995

David Pines

Part I

Feynman Diagrams and Quantum Electrodynamics

Chapter 1

Invitation: Pair Production in e^+e^- Annihilation

The main purpose of Part I of this book is to develop the basic calculational method of quantum field theory, the formalism of Feynman diagrams. We will then apply this formalism to computations in Quantum Electrodynamics, the quantum theory of electrons and photons.

Quantum Electrodynamics (QED) is perhaps the best fundamental physical theory we have. The theory is formulated as a set of simple equations (Maxwell's equations and the Dirac equation) whose form is essentially determined by relativistic invariance. The quantum-mechanical solutions of these equations give detailed predictions of electromagnetic phenomena from macroscopic distances down to regions several hundred times smaller than the proton.

Feynman diagrams provide for this elegant theory an equally elegant procedure for calculation: Imagine a process that can be carried out by electrons and photons, draw a diagram, and then use the diagram to write the mathematical form of the quantum-mechanical amplitude for that process to occur.

In this first part of the book we will develop both the theory of QED and the method of Feynman diagrams from the basic principles of quantum mechanics and relativity. Eventually, we will arrive at a point where we can calculate observable quantities that are of great interest in the study of elementary particles. But to reach our goal of deriving this simple calculational method, we must first, unfortunately, make a serious detour into formalism. The three chapters that follow this one are almost completely formal, and the reader might wonder, in the course of this development, where we are going. We would like to partially answer that question in advance by discussing the physics of an especially simple QED process—one sufficiently simple that many of its features follow directly from physical intuition. Of course, this intuitive, bottom-up approach will contain many gaps. In Chapter 5 we will return to this process with the full power of the Feynman diagram formalism. Working from the top down, we will then see all of these difficulties swept away.

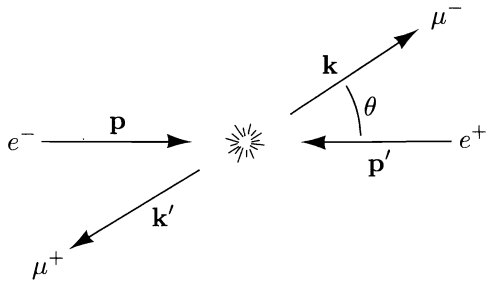


Figure 1.1. The annihilation reaction $e^+e^- \to \mu^+\mu^-$, shown in the center-of-mass frame.

The Simplest Situation

Since most particle physics experiments involve scattering, the most commonly calculated quantities in quantum field theory are scattering cross sections. We will now calculate the cross section for the simplest of all QED processes: the annihilation of an electron with its antiparticle, a positron, to form a pair of heavier leptons (such as muons). The existence of antiparticles is actually a prediction of quantum field theory, as we will discuss in Chapters 2 and 3. For the moment, though, we take their existence as given.

An experiment to measure this annihilation probability would proceed by firing a beam of electrons at a beam of positrons. The measurable quantity is the cross section for the reaction $e^+e^- \to \mu^+\mu^-$ as a function of the center-of-mass energy and the relative angle θ between the incoming electrons and the outgoing muons. The process is illustrated in Fig. 1.1. For simplicity, we work in the center-of-mass (CM) frame where the momenta satisfy $\mathbf{p}' = -\mathbf{p}$ and $\mathbf{k}' = -\mathbf{k}$. We also assume that the beam energy E is much greater than either the electron or the muon mass, so that $|\mathbf{p}| = |\mathbf{p}'| = |\mathbf{k}| = |\mathbf{k}'| = E \equiv E_{\rm cm}/2$. (We use boldface type to denote 3-vectors and ordinary italic type to denote 4-vectors.)

Since both the electron and the muon have spin 1/2, we must specify their spin orientations. It is useful to take the axis that defines the spin quantization of each particle to be in the direction of its motion; each particle can then have its spin polarized parallel or antiparallel to this axis. In practice, electron and positron beams are often unpolarized, and muon detectors are normally blind to the muon polarization. Hence we should average the cross section over electron and positron spin orientations, and sum the cross section over muon spin orientations.

For any given set of spin orientations, it is conventional to write the differential cross section for our process, with the μ^- produced into a solid angle $d\Omega$, as

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 E_{\rm cm}^2} \cdot |\mathcal{M}|^2 \,. \tag{1.1}$$

The factor $E_{\rm cm}^{-2}$ provides the correct dimensions for a cross section, since in our units (energy)$^{-2}$ $\sim$ (length)^2. The quantity $\mathcal{M}$ is therefore dimensionless; it is the quantum-mechanical amplitude for the process to occur (analogous to the scattering amplitude f in nonrelativistic quantum mechanics), and we must now address the question of how to compute it from fundamental theory. The other factors in the expression are purely a matter of convention. Equation (1.1) is actually a special case, valid for CM scattering when the final state contains two massless particles, of a more general formula (whose form cannot be deduced from dimensional analysis) which we will derive in Section 4.5.

Now comes some bad news and some good news.

The bad news is that even for this simplest of QED processes, the exact expression for $\mathcal{M}$ is not known. Actually this fact should come as no surprise, since even in nonrelativistic quantum mechanics, scattering problems can rarely be solved exactly. The best we can do is obtain a formal expression for $\mathcal{M}$ as a perturbation series in the strength of the electromagnetic interaction, and evaluate the first few terms in this series.

The good news is that Feynman has invented a beautiful way to organize and visualize the perturbation series: the method of *Feynman diagrams*. Roughly speaking, the diagrams display the flow of electrons and photons during the scattering process. For our particular calculation, the lowest-order term in the perturbation series can be represented by a single diagram, shown in Fig. 1.2. The diagram is made up of three types of components: external lines (representing the four incoming and outgoing particles), internal lines (representing "virtual" particles, in this case one virtual photon), and vertices. It is conventional to use straight lines for fermions and wavy lines for photons. The arrows on the straight lines denote the direction of negative charge flow, not momentum. We assign a 4-momentum vector to each external line, as shown. In this diagram, the momentum q of the one internal line is determined by momentum conservation at either of the vertices: $q = p + p' = k + k'$. We must also associate a spin state (either "up" or "down") with each external fermion.

According to the *Feynman rules*, each diagram can be translated directly into a contribution to $\mathcal{M}$. The rules assign a short algebraic factor to each element of a diagram, and the product of these factors gives the value of the corresponding term in the perturbation series. Getting the resulting expression for $\mathcal{M}$ into a form that is usable, however, can still be nontrivial. We will develop much useful technology for doing such calculations in subsequent chapters. But we do not have that technology yet, so to get an answer to our particular problem we will use some heuristic arguments instead of the actual Feynman rules.

Recall that in quantum-mechanical perturbation theory, a transition amplitude can be computed, to first order, as an expression of the form

$$\langle \text{final state} | H_I | \text{initial state} \rangle , \tag{1.2}$$

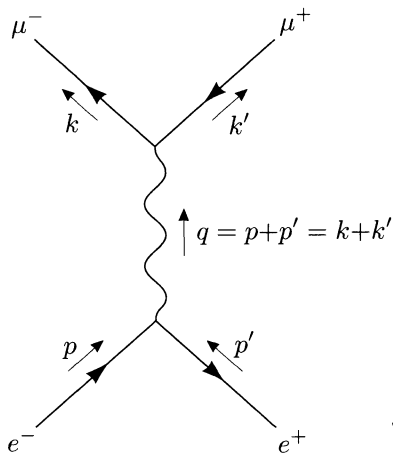


Figure 1.2. Feynman diagram for the lowest-order term in the $e^+e^- \to \mu^+\mu^-$ cross section. At this order the only possible intermediate state is a photon (γ).

where H_I is the "interaction" part of the Hamiltonian. In our case the initial state is $|e^+e^-\rangle$ and the final state is $\langle\mu^+\mu^-|$. But our interaction Hamiltonian couples electrons to muons only through the electromagnetic field (that is, photons), not directly. So the first-order result (1.2) vanishes, and we must go to the second-order expression

$$\mathcal{M} \sim \langle\mu^+\mu^-|\, H_I\, |\gamma\rangle^\mu \,\langle\gamma|\, H_I\, |e^+e^-\rangle_\mu \,. \tag{1.3}$$

This is a heuristic way of writing the contribution to $\mathcal{M}$ from the diagram in Fig. 1.2. The external electron lines correspond to the factor $|e^+e^-\rangle$; the external muon lines correspond to $\langle\mu^+\mu^-|$. The vertices correspond to H_I, and the internal photon line corresponds to the operator $|\gamma\rangle\langle\gamma|$. We have added vector indices (μ) because the photon is a vector particle with four components. There are four possible intermediate states, one for each component, and according to the rules of perturbation theory we must sum over intermediate states. Note that since the sum in (1.3) takes the form of a 4-vector dot product, the amplitude $\mathcal{M}$ will be a Lorentz-invariant scalar as long as each half of (1.3) is a 4-vector.

Let us try to guess the form of the vector $\langle\gamma|\, H_I\, |e^+e^-\rangle_\mu$. Since H_I couples electrons to photons with a strength e (the electron charge), the matrix element should be proportional to e. Now consider one particular set of initial and final spin orientations, shown in Fig. 1.3. The electron and muon have spins parallel to their directions of motion; they are "right-handed". The antiparticles, similarly, are "left-handed". The electron and positron spins add up to one unit of angular momentum in the $+z$ direction. Since H_I should conserve angular momentum, the photon to which these particles couple must have the correct polarization vector to give it this same angular momentum:

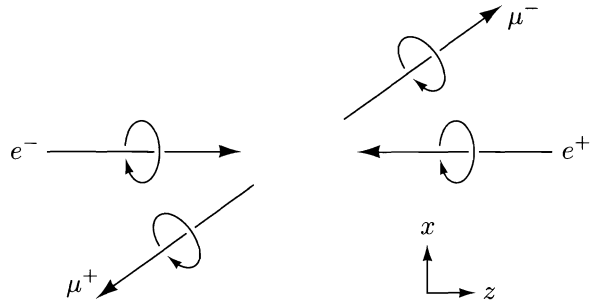


Figure 1.3. One possible set of spin orientations. The electron and the negative muon are right-handed, while the positron and the positive muon are left-handed.

$\epsilon^\mu = (0, 1, i, 0)$. Thus we have

$$\langle \gamma | H_I | e^+e^- \rangle^\mu \propto e\,(0, 1, i, 0). \tag{1.4}$$

The muon matrix element should, similarly, have a polarization corresponding to one unit of angular momentum along the direction of the μ^- momentum $\mathbf{k}$. To obtain the correct vector, rotate (1.4) through an angle θ in the xz-plane:

$$\langle \gamma | H_I | \mu^+\mu^- \rangle^\mu \propto e\,(0, \cos\theta, i, -\sin\theta). \tag{1.5}$$

To compute the amplitude $\mathcal{M}$, we complex-conjugate this vector and dot it into (1.4). Thus we find, for this set of spin orientations,

$$\mathcal{M}(RL \to RL) = -e^2\,(1 + \cos\theta)\ . \tag{1.6}$$

Of course we cannot determine the overall factor by this method, but in (1.6) it happens to be correct, thanks to the conventions adopted in (1.1). Note that the amplitude vanishes for $\theta = 180°$, just as one would expect: A state whose angular momentum is in the $+z$ direction has no overlap with a state whose angular momentum is in the $-z$ direction.

Next consider the case in which the electron and positron are both right-handed. Now their total spin angular momentum is zero, and the argument is more subtle. We might expect to obtain a longitudinally polarized photon with a Clebsch-Gordan coefficient of $1/\sqrt{2}$, just as when we add angular momenta in three dimensions, $|{\uparrow\downarrow}\rangle = (1/\sqrt{2})\big(|j = 1, m = 0\rangle + |j = 0, m = 0\rangle\big)$. But we are really adding angular momenta in the four-dimensional Lorentz group, so we must take into account not only spin (the transformation properties of states under rotations), but also the transformation properties of states under boosts. It turns out, as we shall discuss in Chapter 3, that the Clebsch-Gordan coefficient that couples a 4-vector to the state $|e_R^- e_R^+\rangle$ of massless fermions is zero. (For the record, the state is a superposition of scalar and antisymmetric tensor pieces.) Thus the amplitude $\mathcal{M}(RR \to RL)$ is zero, as are the eleven

other amplitudes in which either the initial or final state has zero total angular momentum.

The remaining nonzero amplitudes can be found in the same way that we found the first one. They are

$$\begin{aligned} \mathcal{M}(RL \to LR) &= -e^2\,(1-\cos\theta), \\ \mathcal{M}(LR \to RL) &= -e^2\,(1-\cos\theta), \\ \mathcal{M}(LR \to LR) &= -e^2\,(1+\cos\theta). \end{aligned} \tag{1.7}$$

Inserting these expressions into (1.1), averaging over the four initial-state spin orientations, and summing over the four final-state spin orientations, we find

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_{\rm cm}^2}\big(1+\cos^2\theta\big), \tag{1.8}$$

where $\alpha = e^2/4\pi \simeq 1/137$. Integrating over the angular variables θ and ϕ gives the total cross section,

$$\sigma_{\rm total} = \frac{4\pi\alpha^2}{3E_{\rm cm}^2}. \tag{1.9}$$

Results (1.8) and (1.9) agree with experiments to about 10%; almost all of the discrepancy is accounted for by the next term in the perturbation series, corresponding to the diagrams shown in Fig. 1.4. The qualitative features of these expressions—the angular dependence and the sharp decrease with energy—are obvious in the actual data. (The properties of these results are discussed in detail in Section 5.1.)

Embellishments and Questions

We obtained the angular distribution predicted by Quantum Electrodynamics for the reaction $e^+e^- \to \mu^+\mu^-$ by applying angular momentum arguments, with little appeal to the underlying formalism. However, we used the simplifying features of the high-energy limit and the center-of-mass frame in a very strong way. The analysis we have presented will break down when we relax any of our simplifying assumptions. So how does one perform general QED calculations? To answer that question we must return to the Feynman rules.

As mentioned above, the Feynman rules tell us to draw the diagram(s) for the process we are considering, and to associate a short algebraic factor with each piece of each diagram. Figure 1.5 shows the diagram for our reaction, with the various assignments indicated.

For the internal photon line we write $-ig_{\mu\nu}/q^2$, where $g_{\mu\nu}$ is the usual Minkowski metric tensor and q is the 4-momentum of the virtual photon. This factor corresponds to the operator $|\gamma\rangle\langle\gamma|$ in our heuristic expression (1.3).

For each vertex we write $-ie\gamma^\mu$, corresponding to H_I in (1.3). The objects γ^μ are a set of four 4×4 constant matrices. They do the "addition of angular

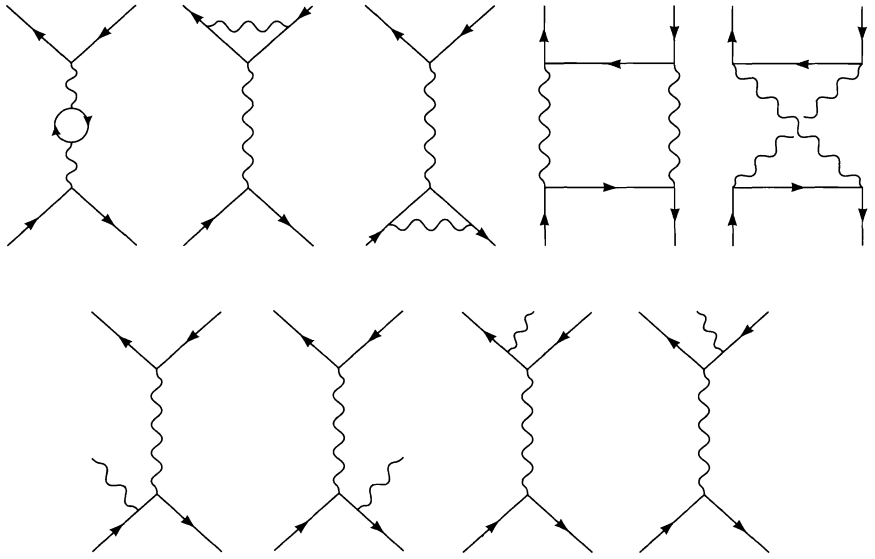


Figure 1.4. Feynman diagrams that contribute to the α^3 term in the $e^+e^- \to \mu^+\mu^-$ cross section.

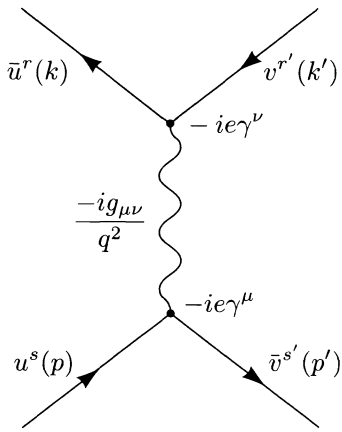


Figure 1.5. Diagram of Fig. 1.2, with expressions corresponding to each vertex, internal line, and external line.

momentum" for us, coupling a state of two spin-1/2 particles to a vector particle.

The external lines carry expressions for four-component column-spinors u, v, or row-spinors $\bar{u}$, $\bar{v}$. These are essentially the momentum-space wavefunctions of the initial and final particles, and correspond to $|e^+e^-\rangle$ and $\langle\mu^+\mu^-|$ in (1.3). The indices s, s', r, and r' denote the spin state, either up or down.

We can now write down an expression for $\mathcal{M}$, reading everything straight off the diagram:

$$\begin{aligned}\mathcal{M} &= \bar{v}^{s'}(p')(-ie\gamma^\mu)u^s(p)\left(\frac{-ig_{\mu\nu}}{q^2}\right)\bar{u}^r(k)(-ie\gamma^\nu)v^{r'}(k') \\ &= \frac{ie^2}{q^2}\big(\bar{v}^{s'}(p')\gamma^\mu u^s(p)\big)\big(\bar{u}^r(k)\gamma_\mu v^{r'}(k')\big).\end{aligned} \tag{1.10}$$

It is instructive to compare this in detail with Eq. (1.3).

To derive the cross section (1.8) from (1.10), we could return to the angular momentum arguments used above, supplemented with some concrete knowledge about γ matrices and Dirac spinors. We will do the calculation in this manner in Section 5.2. There are, however, a number of useful tricks that can be employed to manipulate expressions like (1.10), especially when one wants to compute only the *unpolarized* cross section. Using this "Feynman trace technology" (so-called because one must evaluate traces of products of γ-matrices), it isn't even necessary to have explicit expressions for the γ-matrices and Dirac spinors. The calculation becomes almost completely mindless, and the answer (1.8) is obtained after less than a page of algebra. But since the Feynman rules and trace technology are so powerful, we can also relax some of our simplifying assumptions. To conclude this section, let us discuss several ways in which our calculation could have been more difficult.

The easiest restriction to relax is that the muons be massless. If the beam energy is not much greater than the mass of the muon, all of our predictions should depend on the ratio $m_\mu/E_{\rm cm}$. (Since the electron is 200 times lighter than the muon, it can be considered massless whenever the beam energy is large enough to create muons.) Using Feynman trace technology, it is extremely easy to restore the muon mass to our calculation. The amount of algebra is increased by about fifty percent, and the relation (1.1) between the amplitude and the cross section must be modified slightly, but the answer is worth the effort. We do this calculation in detail in Section 5.1.

Working in a different reference frame is also easy; the only modification is in the relation (1.1) between the amplitude and the cross section. Or one can simply perform a Lorentz transformation on the CM result, boosting it to a different frame.

When the spin states of the initial and/or final particles are known and we still wish to retain the muon mass, the calculation becomes somewhat cumbersome but no more difficult in principle. The trace technology can be generalized to this case, but it is often easier to evaluate expression (1.10) directly, using the explicit values of the spinors u and v.

Next one could compute cross sections for different processes. The process $e^+e^- \to e^+e^-$, known as *Bhabha scattering*, is more difficult because there is a second allowed diagram (see Fig. 1.6). The amplitudes for the two diagrams must first be added, then squared.

Other processes contain photons in the initial and/or final states. The

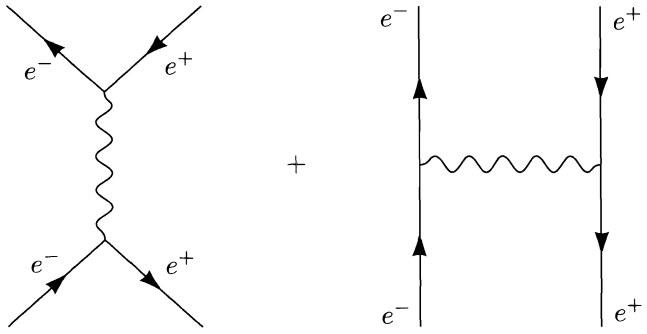


Figure 1.6. The two lowest-order diagrams for Bhabha scattering, $e^+e^- \to e^+e^-$.

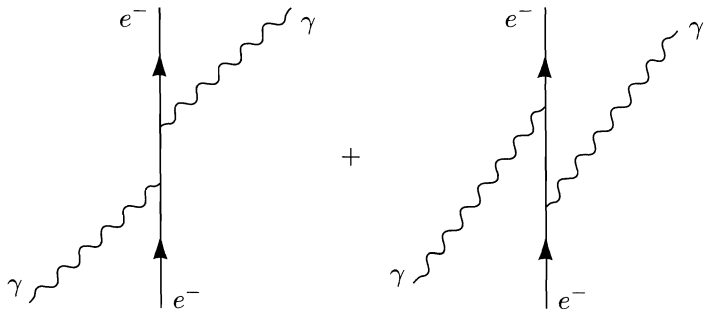


Figure 1.7. The two lowest-order diagrams for Compton scattering.

paradigm example is Compton scattering, for which the two lowest-order diagrams are shown in Fig. 1.7. The Feynman rules for external photon lines and for internal electron lines are no more complicated than those we have already seen. We discuss Compton scattering in detail in Section 5.5.

Finally we could compute higher-order terms in the perturbation series. Thanks to Feynman, the diagrams are at least easy to draw; we have seen those that contribute to the next term in the $e^+e^- \to \mu^+\mu^-$ cross section in Fig. 1.4. Remarkably, the algorithm that assigns algebraic factors to pieces of the diagrams holds for all higher-order contributions, and allows one to evaluate such diagrams in a straightforward, if tedious, way. The computation of the full set of nine diagrams is a serious chore, at the level of a research paper.

In this book, starting in Chapter 6, we will analyze much of the physics that arises from higher-order Feynman diagrams such as those in Fig. 1.4. We will see that the last four of these diagrams, which involve an additional photon in the final state, are necessary because no detector is sensitive enough to notice the presence of extremely low-energy photons. Thus a final state containing such a photon cannot be distinguished from our desired final state of just a muon pair.

The other five diagrams in Fig. 1.4 involve intermediate states of several virtual particles rather than just a single virtual photon. In each of these diagrams there will be one virtual particle whose momentum is not determined by conservation of momentum at the vertices. Since perturbation theory requires us to sum over all possible intermediate states, we must integrate over all possible values of this momentum. At this step, however, a new difficulty appears: The loop-momentum integrals in the first three diagrams, when performed naively, turn out to be infinite. We will provide a fix for this problem, so that we get finite results, by the end of Part I. But the question of the physical origin of these divergences cannot be dismissed so lightly; that will be the main subject of Part II of this book.

We have discussed Feynman diagrams as an algorithm for performing computations. The chapters that follow should amply illustrate the power of this tool. As we expose more applications of the diagrams, though, they begin to take on a life and significance of their own. They indicate unsuspected relations between different physical processes, and they suggest intuitive arguments that might later be verified by calculation. We hope that this book will enable you, the reader, to take up this tool and apply it in novel and enlightening ways.

Chapter 2

The Klein-Gordon Field

2.1 The Necessity of the Field Viewpoint

Quantum field theory is the application of quantum mechanics to dynamical systems of *fields*, in the same sense that the basic course in quantum mechanics is concerned mainly with the quantization of dynamical systems of *particles*. It is a subject that is absolutely essential for understanding the current state of elementary particle physics. With some modification, the methods we will discuss also play a crucial role in the most active areas of atomic, nuclear, and condensed-matter physics. In Part I of this book, however, our primary concern will be with elementary particles, and hence *relativistic* fields.

Given that we wish to understand processes that occur at very small (quantum-mechanical) scales and very large (relativistic) energies, one might still ask why we must study the quantization of *fields*. Why can't we just quantize relativistic particles the way we quantized nonrelativistic particles?

This question can be answered on a number of levels. Perhaps the best approach is to write down a single-particle relativistic wave equation (such as the Klein-Gordon equation or the Dirac equation) and see that it gives rise to negative-energy states and other inconsistencies. Since this discussion usually takes place near the end of a graduate-level quantum mechanics course, we will not repeat it here. It is easy, however, to understand why such an approach cannot work. We have no right to assume that any relativistic process can be explained in terms of a single particle, since the Einstein relation $E = mc^2$ allows for the creation of particle-antiparticle pairs. Even when there is not enough energy for pair creation, multiparticle states appear, for example, as intermediate states in second-order perturbation theory. We can think of such states as existing only for a very short time, according to the uncertainty principle $\Delta E \cdot \Delta t = \hbar$. As we go to higher orders in perturbation theory, arbitrarily many such "virtual" particles can be created.

The necessity of having a multiparticle theory also arises in a less obvious way, from considerations of causality. Consider the amplitude for a free particle to propagate from $\mathbf{x}_0$ to $\mathbf{x}$:

$$U(t) = \langle \mathbf{x} | \, e^{-iHt} \, | \mathbf{x}_0 \rangle \, .$$

In nonrelativistic quantum mechanics we have $E = \mathbf{p}^2/2m$, so

$$\begin{aligned} U(t) &= \langle \mathbf{x}|\, e^{-i(\mathbf{p}^2/2m)t}\, |\mathbf{x}_0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3}\, \langle \mathbf{x}|\, e^{-i(\mathbf{p}^2/2m)t}\, |\mathbf{p}\rangle\, \langle \mathbf{p}|\, \mathbf{x}_0\rangle \\ &= \frac{1}{(2\pi)^3} \int d^3p\, e^{-i(\mathbf{p}^2/2m)t} \cdot e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)} \\ &= \left(\frac{m}{2\pi i t}\right)^{3/2} e^{im(\mathbf{x}-\mathbf{x}_0)^2/2t}. \end{aligned}$$

This expression is nonzero for all x and t, indicating that a particle can propagate between any two points in an arbitrarily short time. In a relativistic theory, this conclusion would signal a violation of causality. One might hope that using the relativistic expression $E = \sqrt{p^2 + m^2}$ would help, but it does not. In analogy with the nonrelativistic case, we have

$$\begin{aligned} U(t) &= \langle \mathbf{x}|\, e^{-it\sqrt{\mathbf{p}^2+m^2}}\, |\mathbf{x}_0\rangle \\ &= \frac{1}{(2\pi)^3} \int d^3p\, e^{-it\sqrt{\mathbf{p}^2+m^2}} \cdot e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)} \\ &= \frac{1}{2\pi^2|\mathbf{x}-\mathbf{x}_0|} \int_0^\infty dp\, p\, \sin(p|\mathbf{x}-\mathbf{x}_0|) e^{-it\sqrt{p^2+m^2}}. \end{aligned}$$

This integral can be evaluated explicitly in terms of Bessel functions.* We will content ourselves with looking at its asymptotic behavior for $x^2 \gg t^2$ (well outside the light-cone), using the method of stationary phase. The phase function $px - t\sqrt{p^2+m^2}$ has a stationary point at $p = imx/\sqrt{x^2-t^2}$. We may freely push the contour upward so that it goes through this point. Plugging in this value for p, we find that, up to a rational function of x and t,

$$U(t) \sim e^{-m\sqrt{x^2-t^2}}.$$

Thus the propagation amplitude is small but nonzero outside the light-cone, and causality is still violated.

Quantum field theory solves the causality problem in a miraculous way, which we will discuss in Section 2.4. We will find that, in the multiparticle field theory, the propagation of a particle across a spacelike interval is indistinguishable from the propagation of an *antiparticle* in the opposite direction (see Fig. 2.1). When we ask whether an observation made at point x_0 can affect an observation made at point x, we will find that the amplitudes for particle and antiparticle propagation exactly cancel—so causality is preserved.

Quantum field theory provides a natural way to handle not only multiparticle states, but also transitions between states of different particle number. It solves the causality problem by introducing antiparticles, then goes on to

*See Gradshteyn and Ryzhik (1980), #3.914.

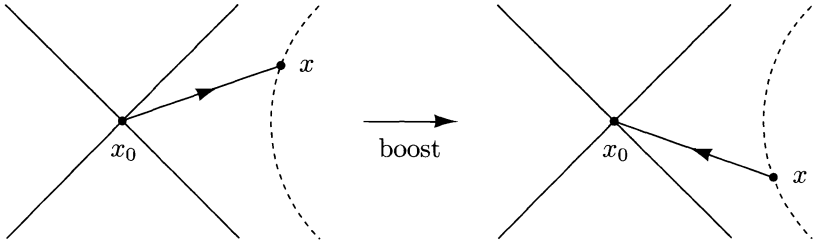


Figure 2.1. Propagation from x_0 to x in one frame looks like propagation from x to x_0 in another frame.

explain the relation between spin and statistics. But most important, it provides the tools necessary to calculate innumerable scattering cross sections, particle lifetimes, and other observable quantities. The experimental confirmation of these predictions, often to an unprecedented level of accuracy, is our real reason for studying quantum field theory.

2.2 Elements of Classical Field Theory

In this section we review some of the formalism of classical field theory that will be necessary in our subsequent discussion of quantum field theory.

Lagrangian Field Theory

The fundamental quantity of classical mechanics is the action, S, the time integral of the Lagrangian, L. In a local field theory the Lagrangian can be written as the spatial integral of a Lagrangian density, denoted by $\mathcal{L}$, which is a function of one or more fields $\phi(x)$ and their derivatives $\partial_\mu \phi$. Thus we have

$$S = \int L\, dt = \int \mathcal{L}(\phi, \partial_\mu \phi)\, d^4x. \tag{2.1}$$

Since this is a book on field theory, we will refer to $\mathcal{L}$ simply as the Lagrangian.

The principle of least action states that when a system evolves from one given configuration to another between times t_1 and t_2, it does so along the "path" in configuration space for which S is an extremum (normally a minimum). We can write this condition as

$$\begin{aligned} 0 &= \delta S \\ &= \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta(\partial_\mu \phi) \right\} \\ &= \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \delta\phi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right) \right\}. \end{aligned} \tag{2.2}$$

The last term can be turned into a surface integral over the boundary of the four-dimensional spacetime region of integration. Since the initial and final field configurations are assumed given, $\delta\phi$ is zero at the temporal beginning

and end of this region. If we restrict our consideration to deformations $\delta\phi$ that vanish on the spatial boundary of the region as well, then the surface term is zero. Factoring out the $\delta\phi$ from the first two terms, we note that, since the integral must vanish for arbitrary $\delta\phi$, the quantity that multiplies $\delta\phi$ must vanish at all points. Thus we arrive at the Euler-Lagrange equation of motion for a field,

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{2.3}$$

If the Lagrangian contains more than one field, there is one such equation for each.

Hamiltonian Field Theory

The Lagrangian formulation of field theory is particularly suited to relativistic dynamics because all expressions are explicitly Lorentz invariant. Nevertheless we will use the Hamiltonian formulation throughout the first part of this book, since it will make the transition to quantum mechanics easier. Recall that for a discrete system one can define a conjugate momentum $p \equiv \partial L/\partial \dot{q}$ (where $\dot{q} = \partial q/\partial t$) for each dynamical variable q. The Hamiltonian is then $H \equiv \sum p\dot{q} - L$. The generalization to a continuous system is best understood by pretending that the spatial points $\mathbf{x}$ are discretely spaced. We can define

$$\begin{aligned} p(\mathbf{x}) &\equiv \frac{\partial L}{\partial \dot{\phi}(\mathbf{x})} = \frac{\partial}{\partial \dot{\phi}(\mathbf{x})} \int \mathcal{L}\big(\phi(\mathbf{y}), \dot{\phi}(\mathbf{y})\big)\, d^3y \\ &\sim \frac{\partial}{\partial \dot{\phi}(\mathbf{x})} \sum_{\mathbf{y}} \mathcal{L}\big(\phi(\mathbf{y}), \dot{\phi}(\mathbf{y})\big) d^3y \\ &= \pi(\mathbf{x}) d^3x, \end{aligned}$$

where

$$\pi(\mathbf{x}) \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\mathbf{x})} \tag{2.4}$$

is called the *momentum density* conjugate to $\phi(\mathbf{x})$. Thus the Hamiltonian can be written

$$H = \sum_{\mathbf{x}} p(\mathbf{x})\dot{\phi}(\mathbf{x}) - L.$$

Passing to the continuum, this becomes

$$H = \int d^3x \left[\pi(\mathbf{x})\dot{\phi}(\mathbf{x}) - \mathcal{L}\right] \equiv \int d^3x\, \mathcal{H}. \tag{2.5}$$

We will rederive this expression for the Hamiltonian density $\mathcal{H}$ near the end of this section, using a different method.

As a simple example, consider the theory of a single field $\phi(x)$, governed by the Lagrangian

$$\begin{aligned} \mathcal{L} &= \tfrac{1}{2}\dot{\phi}^2 - \tfrac{1}{2}(\nabla\phi)^2 - \tfrac{1}{2}m^2\phi^2 \\ &= \tfrac{1}{2}(\partial_\mu \phi)^2 - \tfrac{1}{2}m^2\phi^2. \end{aligned} \tag{2.6}$$

For now we take ϕ to be a real-valued field. The quantity m will be interpreted as a mass in Section 2.3, but for now just think of it as a parameter. From this Lagrangian the usual procedure gives the equation of motion

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\phi = 0 \qquad \text{or} \qquad (\partial^\mu \partial_\mu + m^2)\phi = 0, \tag{2.7}$$

which is the well-known Klein-Gordon equation. (In this context it is a *classical* field equation, like Maxwell's equations—not a quantum-mechanical wave equation.) Noting that the canonical momentum density conjugate to $\phi(x)$ is $\pi(x) = \dot{\phi}(x)$, we can also construct the Hamiltonian:

$$H = \int d^3x\, \mathcal{H} = \int d^3x \left[\tfrac{1}{2}\pi^2 + \tfrac{1}{2}(\nabla\phi)^2 + \tfrac{1}{2}m^2\phi^2\right]. \tag{2.8}$$

We can think of the three terms, respectively, as the energy cost of "moving" in time, the energy cost of "shearing" in space, and the energy cost of having the field around at all. We will investigate this Hamiltonian much further in Sections 2.3 and 2.4.

Noether's Theorem

Next let us discuss the relationship between symmetries and conservation laws in classical field theory, summarized in *Noether's theorem.* This theorem concerns continuous transformations on the fields ϕ, which in infinitesimal form can be written

$$\phi(x) \to \phi'(x) = \phi(x) + \alpha\Delta\phi(x), \tag{2.9}$$

where α is an infinitesimal parameter and $\Delta\phi$ is some deformation of the field configuration. We call this transformation a symmetry if it leaves the equations of motion invariant. This is insured if the action is invariant under (2.9). More generally, we can allow the action to change by a surface term, since the presence of such a term would not affect our derivation of the Euler-Lagrange equations of motion (2.3). The Lagrangian, therefore, must be invariant under (2.9) up to a 4-divergence:

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha\partial_\mu \mathcal{J}^\mu(x), \tag{2.10}$$

for some $\mathcal{J}^\mu$. Let us compare this expectation for $\Delta\mathcal{L}$ to the result obtained by varying the fields:

$$\begin{aligned} \alpha\Delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi}(\alpha\Delta\phi) + \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\partial_\mu(\alpha\Delta\phi) \\ &= \alpha\,\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Delta\phi\right) + \alpha\left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right]\Delta\phi. \end{aligned} \tag{2.11}$$

The second term vanishes by the Euler-Lagrange equation (2.3). We set the remaining term equal to $\alpha\partial_\mu \mathcal{J}^\mu$ and find

$$\partial_\mu j^\mu(x) = 0, \qquad \text{for} \quad j^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\Delta\phi - \mathcal{J}^\mu. \tag{2.12}$$

(If the symmetry involves more than one field, the first term of this expression for $j^\mu(x)$ should be replaced by a sum of such terms, one for each field.) This result states that the current $j^\mu(x)$ is conserved. For each continuous symmetry of $\mathcal{L}$, we have such a conservation law.

The conservation law can also be expressed by saying that the charge

$$Q \equiv \int_{\text{all space}} j^0 \, d^3x \tag{2.13}$$

is a constant in time. Note, however, that the formulation of field theory in terms of a local Lagrangian density leads directly to the local form of the conservation law, Eq. (2.12).

The easiest example of such a conservation law arises from a Lagrangian with only a kinetic term: $\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2$. The transformation $\phi \to \phi + \alpha$, where α is a constant, leaves $\mathcal{L}$ unchanged, so we conclude that the current $j^\mu = \partial^\mu \phi$ is conserved. As a less trivial example, consider the Lagrangian

$$\mathcal{L} = |\partial_\mu \phi|^2 - m^2|\phi|^2, \tag{2.14}$$

where ϕ is now a *complex*-valued field. You can easily show that the equation of motion for this Lagrangian is again the Klein-Gordon equation, (2.7). This Lagrangian is invariant under the transformation $\phi \to e^{i\alpha}\phi$; for an infinitesimal transformation we have

$$\alpha\Delta\phi = i\alpha\phi; \qquad \alpha\Delta\phi^* = -i\alpha\phi^*. \tag{2.15}$$

(We treat ϕ and ϕ^* as independent fields. Alternatively, we could work with the real and imaginary parts of ϕ.) It is now a simple matter to show that the conserved Noether current is

$$j^\mu = i\big[(\partial^\mu \phi^*)\phi - \phi^*(\partial^\mu \phi)\big]. \tag{2.16}$$

(The overall constant has been chosen arbitrarily.) You can check directly that the divergence of this current vanishes by using the Klein-Gordon equation. Later we will add terms to this Lagrangian that couple ϕ to an electromagnetic field. We will then interpret j^μ as the electromagnetic current density carried by the field, and the spatial integral of j^0 as its electric charge.

Noether's theorem can also be applied to spacetime transformations such as translations and rotations. We can describe the infinitesimal translation

$$x^\mu \to x^\mu - a^\mu$$

alternatively as a transformation of the field configuration

$$\phi(x) \to \phi(x + a) \doteq \phi(x) + a^\mu \partial_\mu \phi(x).$$

The Lagrangian is also a scalar, so it must transform in the same way:

$$\mathcal{L} \to \mathcal{L} + a^\mu \partial_\mu \mathcal{L} = \mathcal{L} + a^\nu \partial_\mu (\delta^\mu{}_\nu \mathcal{L}).$$

Comparing this equation to (2.10), we see that we now have a nonzero $\mathcal{J}^\mu$. Taking this into account, we can apply the theorem to obtain four separately conserved currents:

$$T^\mu{}_\nu \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \mathcal{L} \delta^\mu{}_\nu. \tag{2.17}$$

This is precisely the *stress-energy tensor*, also called the *energy-momentum tensor*, of the field ϕ. The conserved charge associated with time translations is the Hamiltonian:

$$H = \int T^{00}\, d^3x = \int \mathcal{H}\, d^3x. \tag{2.18}$$

By computing this quantity for the Klein-Gordon field, one can recover the result (2.8). The conserved charges associated with spatial translations are

$$P^i = \int T^{0i}\, d^3x = -\int \pi \partial_i \phi \; d^3x, \tag{2.19}$$

and we naturally interpret this as the (physical) momentum carried by the field (not to be confused with the canonical momentum).

2.3 The Klein-Gordon Field as Harmonic Oscillators

We begin our discussion of *quantum* field theory with a rather formal treatment of the simplest type of field: the real Klein-Gordon field. The idea is to start with a classical field theory (the theory of a classical scalar field governed by the Lagrangian (2.6)) and then "quantize" it, that is, reinterpret the dynamical variables as operators that obey canonical commutation relations.[†] We will then "solve" the theory by finding the eigenvalues and eigenstates of the Hamiltonian, using the harmonic oscillator as an analogy.

The classical theory of the real Klein-Gordon field was discussed briefly (but sufficiently) in the previous section; the relevant expressions are given in Eqs. (2.6), (2.7), and (2.8). To quantize the theory, we follow the same procedure as for any other dynamical system: We promote ϕ and π to operators, and impose suitable commutation relations. Recall that for a discrete system of one or more particles the commutation relations are

$$[q_i, p_j] = i\delta_{ij};$$
$$[q_i, q_j] = [p_i, p_j] = 0.$$

[†]This procedure is sometimes called *second quantization*, to distinguish the resulting Klein-Gordon equation (in which ϕ is an operator) from the old one-particle Klein-Gordon equation (in which ϕ was a wavefunction). In this book we never adopt the latter point of view; we start with a classical equation (in which ϕ is a classical field) and quantize it exactly once.

For a continuous system the generalization is quite natural; since $\pi(\mathbf{x})$ is the momentum *density*, we get a Dirac delta function instead of a Kronecker delta:

$$\begin{aligned}\big[\phi(\mathbf{x}),\pi(\mathbf{y})\big] &= i\delta^{(3)}(\mathbf{x}-\mathbf{y});\\ \big[\phi(\mathbf{x}),\phi(\mathbf{y})\big] &= \big[\pi(\mathbf{x}),\pi(\mathbf{y})\big] = 0.\end{aligned} \tag{2.20}$$

(For now we work in the Schrödinger picture where ϕ and π do not depend on time. When we switch to the Heisenberg picture in the next section, these "equal time" commutation relations will still hold provided that both operators are considered at the same time.)

The Hamiltonian, being a function of ϕ and π, also becomes an operator. Our next task is to find the spectrum from the Hamiltonian. Since there is no obvious way to do this, let us seek guidance by writing the Klein-Gordon equation in Fourier space. If we expand the classical Klein-Gordon field as

$$\phi(\mathbf{x},t) = \int \frac{d^3p}{(2\pi)^3}\, e^{i\mathbf{p}\cdot\mathbf{x}}\, \phi(\mathbf{p},t)$$

(with $\phi^*(\mathbf{p}) = \phi(-\mathbf{p})$ so that $\phi(\mathbf{x})$ is real), the Klein-Gordon equation (2.7) becomes

$$\left[\frac{\partial^2}{\partial t^2} + \big(|\mathbf{p}|^2 + m^2\big)\right]\phi(\mathbf{p},t) = 0. \tag{2.21}$$

This is the same as the equation of motion for a simple harmonic oscillator with frequency

$$\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}. \tag{2.22}$$

The simple harmonic oscillator is a system whose spectrum we already know how to find. Let us briefly recall how it is done. We write the Hamiltonian as

$$H_{\text{SHO}} = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2\phi^2.$$

To find the eigenvalues of H_{SHO}, we write ϕ and p in terms of ladder operators:

$$\phi = \frac{1}{\sqrt{2\omega}}(a + a^\dagger); \qquad p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger). \tag{2.23}$$

The canonical commutation relation $[\phi, p] = i$ is equivalent to

$$\big[a, a^\dagger\big] = 1. \tag{2.24}$$

The Hamiltonian can now be rewritten

$$H_{\text{SHO}} = \omega(a^\dagger a + \tfrac{1}{2}).$$

The state $|0\rangle$ such that $a\,|0\rangle = 0$ is an eigenstate of H with eigenvalue $\frac{1}{2}\omega$, the zero-point energy. Furthermore, the commutators

$$\big[H_{\text{SHO}}, a^\dagger\big] = \omega a^\dagger, \qquad \big[H_{\text{SHO}}, a\big] = -\omega a$$

make it easy to verify that the states

$$|n\rangle \equiv (a^\dagger)^n\,|0\rangle$$

are eigenstates of H_{SHO} with eigenvalues $(n + \frac{1}{2})\omega$. These states exhaust the spectrum.

We can find the spectrum of the Klein-Gordon Hamiltonian using the same trick, but now each Fourier mode of the field is treated as an independent oscillator with its own a and $a^\dagger$. In analogy with (2.23) we write

$$\phi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} + a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}\right); \tag{2.25}$$

$$\pi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} (-i)\sqrt{\frac{\omega_{\mathbf{p}}}{2}} \left(a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} - a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}\right). \tag{2.26}$$

The inverse expressions for $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$ in terms of ϕ and π are easy to derive but rarely needed. In the calculations below we will find it useful to rearrange (2.25) and (2.26) as follows:

$$\phi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} (a_{\mathbf{p}} + a_{-\mathbf{p}}^\dagger) e^{i\mathbf{p}\cdot\mathbf{x}}; \tag{2.27}$$

$$\pi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} (-i)\sqrt{\frac{\omega_{\mathbf{p}}}{2}} (a_{\mathbf{p}} - a_{-\mathbf{p}}^\dagger) e^{i\mathbf{p}\cdot\mathbf{x}}. \tag{2.28}$$

The commutation relation (2.24) becomes

$$[a_{\mathbf{p}}, a_{\mathbf{p}'}^\dagger] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}'), \tag{2.29}$$

from which you can verify that the commutator of ϕ and π works out correctly:

$$\begin{aligned}[\phi(\mathbf{x}), \pi(\mathbf{x}')] &= \int \frac{d^3p\, d^3p'}{(2\pi)^6} \frac{-i}{2} \sqrt{\frac{\omega_{\mathbf{p}'}}{\omega_{\mathbf{p}}}} \left([a_{-\mathbf{p}}^\dagger, a_{\mathbf{p}'}] - [a_{\mathbf{p}}, a_{-\mathbf{p}'}^\dagger]\right) e^{i(\mathbf{p}\cdot\mathbf{x} + \mathbf{p}'\cdot\mathbf{x}')} \\ &= i\delta^{(3)}(\mathbf{x} - \mathbf{x}'). \end{aligned} \tag{2.30}$$

(If computations such as this one and the next are unfamiliar to you, please work them out carefully; they are quite easy after a little practice, and are fundamental to the formalism of the next two chapters.)

We are now ready to express the Hamiltonian in terms of ladder operators. Starting from its expression (2.8) in terms of ϕ and π, we have

$$\begin{aligned} H = \int d^3x \int \frac{d^3p\, d^3p'}{(2\pi)^6} e^{i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} \Bigg\{ &- \frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}{4} (a_{\mathbf{p}} - a_{-\mathbf{p}}^\dagger)(a_{\mathbf{p}'} - a_{-\mathbf{p}'}^\dagger) \\ &+ \frac{-\mathbf{p}\cdot\mathbf{p}' + m^2}{4\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}} (a_{\mathbf{p}} + a_{-\mathbf{p}}^\dagger)(a_{\mathbf{p}'} + a_{-\mathbf{p}'}^\dagger) \Bigg\} \\ = \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} &\left(a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \tfrac{1}{2}[a_{\mathbf{p}}, a_{\mathbf{p}}^\dagger]\right). \end{aligned} \tag{2.31}$$

The second term is proportional to $\delta(0)$, an infinite c-number. It is simply the sum over all modes of the zero-point energies $\omega_{\mathbf{p}}/2$, so its presence is completely expected, if somewhat disturbing. Fortunately, this infinite energy

shift cannot be detected experimentally, since experiments measure only energy *differences* from the ground state of H. We will therefore ignore this infinite constant term in all of our calculations. It is possible that this energy shift of the ground state could create a problem at a deeper level in the theory; we will discuss this matter in the Epilogue.

Using this expression for the Hamiltonian in terms of $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$, it is easy to evaluate the commutators

$$[H, a_{\mathbf{p}}^\dagger] = \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger; \qquad [H, a_{\mathbf{p}}] = -\omega_{\mathbf{p}} a_{\mathbf{p}}. \tag{2.32}$$

We can now write down the spectrum of the theory, just as for the harmonic oscillator. The state $|0\rangle$ such that $a_{\mathbf{p}}|0\rangle = 0$ for all $\mathbf{p}$ is the ground state or *vacuum*, and has $E = 0$ after we drop the infinite constant in (2.31). All other energy eigenstates can be built by acting on $|0\rangle$ with creation operators. In general, the state $a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger \cdots |0\rangle$ is an eigenstate of H with energy $\omega_{\mathbf{p}} + \omega_{\mathbf{q}} + \cdots$. These states exhaust the spectrum.

Having found the spectrum of the Hamiltonian, let us try to interpret its eigenstates. From (2.19) and a calculation similar to (2.31) we can write down the total momentum operator,

$$\mathbf{P} = -\int d^3x\, \pi(\mathbf{x}) \nabla\phi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\, \mathbf{p}\, a_{\mathbf{p}}^\dagger a_{\mathbf{p}}. \tag{2.33}$$

So the operator $a_{\mathbf{p}}^\dagger$ creates momentum $\mathbf{p}$ and energy $\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}$. Similarly, the state $a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger \cdots |0\rangle$ has momentum $\mathbf{p} + \mathbf{q} + \cdots$. It is quite natural to call these excitations *particles*, since they are discrete entities that have the proper relativistic energy-momentum relation. (By a *particle* we do not mean something that must be localized in space; $a_{\mathbf{p}}^\dagger$ creates particles in momentum eigenstates.) From now on we will refer to $\omega_{\mathbf{p}}$ as $E_{\mathbf{p}}$ (or simply E), since it really is the energy of a particle. Note, by the way, that the energy is always positive: $E_{\mathbf{p}} = +\sqrt{|\mathbf{p}|^2 + m^2}$.

This formalism also allows us to determine the statistics of our particles. Consider the two-particle state $a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger |0\rangle$. Since $a_{\mathbf{p}}^\dagger$ and $a_{\mathbf{q}}^\dagger$ commute, this state is identical to the state $a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle$ in which the two particles are interchanged. Moreover, a single mode $\mathbf{p}$ can contain arbitrarily many particles (just as a simple harmonic oscillator can be excited to arbitrarily high levels). Thus we conclude that Klein-Gordon particles obey *Bose-Einstein statistics.*

We naturally choose to normalize the vacuum state so that $\langle 0|0\rangle = 1$. The one-particle states $|\mathbf{p}\rangle \propto a_{\mathbf{p}}^\dagger |0\rangle$ will also appear quite often, and it is worthwhile to adopt a convention for their normalization. The simplest normalization $\langle \mathbf{p}|\mathbf{q}\rangle = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q})$ (which many books use) is not Lorentz invariant, as we can demonstrate by considering the effect of a boost in the 3-direction. Under such a boost we have $p_3' = \gamma(p_3 + \beta E)$, $E' = \gamma(E + \beta p_3)$. Using the delta function identity

$$\delta\big(f(x) - f(x_0)\big) = \frac{1}{|f'(x_0)|}\delta(x - x_0), \tag{2.34}$$

we can compute

$$\begin{aligned}\delta^{(3)}(\mathbf{p}-\mathbf{q}) &= \delta^{(3)}(\mathbf{p}'-\mathbf{q}')\cdot\frac{dp_3'}{dp_3}\\ &= \delta^{(3)}(\mathbf{p}'-\mathbf{q}')\gamma\Big(1+\beta\frac{dE}{dp_3}\Big)\\ &= \delta^{(3)}(\mathbf{p}'-\mathbf{q}')\frac{\gamma}{E}(E+\beta p_3)\\ &= \delta^{(3)}(\mathbf{p}'-\mathbf{q}')\frac{E'}{E}.\end{aligned}$$

The problem is that volumes are not invariant under boosts; a box whose volume is V in its rest frame has volume V/γ in a boosted frame, due to Lorentz contraction. But from the above calculation, we see that the quantity $E_{\mathbf{p}}\delta^{(3)}(\mathbf{p}-\mathbf{q})$ *is* Lorentz invariant. We therefore define

$$|\mathbf{p}\rangle = \sqrt{2E_{\mathbf{p}}}\, a_{\mathbf{p}}^\dagger\,|0\rangle\,, \tag{2.35}$$

so that

$$\langle\mathbf{p}|\mathbf{q}\rangle = 2E_{\mathbf{p}}(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q}). \tag{2.36}$$

(The factor of 2 is unnecessary, but is convenient because of the factor of 2 in Eq. (2.25).)

On the Hilbert space of quantum states, a Lorentz transformation Λ will be implemented as some unitary operator $U(\Lambda)$. Our normalization condition (2.35) then implies that

$$U(\Lambda)\,|\mathbf{p}\rangle = |\Lambda\mathbf{p}\rangle\,. \tag{2.37}$$

If we prefer to think of this transformation as acting on the operator $a_{\mathbf{p}}^\dagger$, we can also write

$$U(\Lambda)\,a_{\mathbf{p}}^\dagger\,U^{-1}(\Lambda) = \sqrt{\frac{E_{\Lambda\mathbf{p}}}{E_{\mathbf{p}}}}\,a_{\Lambda\mathbf{p}}^\dagger. \tag{2.38}$$

With this normalization we must divide by $2E_{\mathbf{p}}$ in other places. For example, the completeness relation for the one-particle states is

$$(\mathbf{1})_{\text{1-particle}} = \int\frac{d^3p}{(2\pi)^3}\,|\mathbf{p}\rangle\,\frac{1}{2E_{\mathbf{p}}}\,\langle\mathbf{p}|\,, \tag{2.39}$$

where the operator on the left is simply the identity within the subspace of one-particle states, and zero in the rest of the Hilbert space. Integrals of this form will occur quite often; in fact, the integral

$$\int\frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} = \int\frac{d^4p}{(2\pi)^4}(2\pi)\delta(p^2-m^2)\Big|_{p^0>0} \tag{2.40}$$

is a Lorentz-invariant 3-momentum integral, in the sense that if $f(p)$ is Lorentz-invariant, so is $\int d^3p\, f(p)/(2E_{\mathbf{p}})$. The integration can be thought of

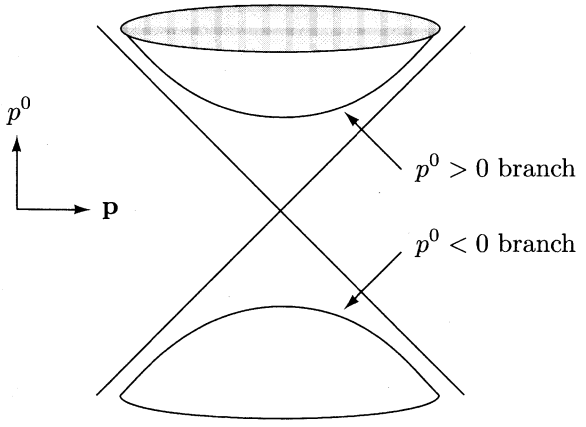


Figure 2.2. The Lorentz-invariant 3-momentum integral is over the upper branch of the hyperboloid $p^2 = m^2$.

as being over the $p^0 > 0$ branch of the hyperboloid $p^2 = m^2$ in 4-momentum space (see Fig. 2.2).

Finally let us consider the interpretation of the state $\phi(\mathbf{x})\,|0\rangle$. From the expansion (2.25) we see that

$$\phi(\mathbf{x})\,|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} e^{-i\mathbf{p}\cdot\mathbf{x}}\,|\mathbf{p}\rangle \tag{2.41}$$

is a linear superposition of single-particle states that have well-defined momentum. Except for the factor $1/2E_{\mathbf{p}}$, this is the same as the familiar nonrelativistic expression for the eigenstate of position $|\mathbf{x}\rangle$; in fact the extra factor is nearly constant for small (nonrelativistic) $\mathbf{p}$. We will therefore put forward the same interpretation, and claim that the operator $\phi(\mathbf{x})$, acting on the vacuum, *creates a particle at position* $\mathbf{x}$. This interpretation is further confirmed when we compute

$$\begin{aligned}\langle 0|\,\phi(\mathbf{x})\,|\mathbf{p}\rangle &= \langle 0|\int \frac{d^3p'}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}'}}}\left(a_{\mathbf{p}'}e^{i\mathbf{p}'\cdot\mathbf{x}} + a^\dagger_{\mathbf{p}'}e^{-i\mathbf{p}'\cdot\mathbf{x}}\right)\sqrt{2E_{\mathbf{p}}}\,a^\dagger_{\mathbf{p}}\,|0\rangle \\ &= e^{i\mathbf{p}\cdot\mathbf{x}}.\end{aligned} \tag{2.42}$$

We can interpret this as the position-space representation of the single-particle wavefunction of the state $|\mathbf{p}\rangle$, just as in nonrelativistic quantum mechanics $\langle\mathbf{x}|\mathbf{p}\rangle \propto e^{i\mathbf{p}\cdot\mathbf{x}}$ is the wavefunction of the state $|\mathbf{p}\rangle$.

2.4 The Klein-Gordon Field in Space-Time

In the previous section we quantized the Klein-Gordon field in the Schrödinger picture, and interpreted the resulting theory in terms of relativistic particles. In this section we will switch to the Heisenberg picture, where it will be easier to discuss time-dependent quantities and questions of causality. After a few preliminaries, we will return to the question of acausal propagation raised in Section 2.1. We will also derive an expression for the *Klein-Gordon propagator*, a crucial part of the Feynman rules to be developed in Chapter 4.

In the Heisenberg picture, we make the operators ϕ and π time-dependent in the usual way:

$$\phi(x) = \phi(\mathbf{x}, t) = e^{iHt}\phi(\mathbf{x})e^{-iHt}, \tag{2.43}$$

and similarly for $\pi(x) = \pi(\mathbf{x}, t)$. The Heisenberg equation of motion,

$$i\frac{\partial}{\partial t}\mathcal{O} = [\mathcal{O}, H], \tag{2.44}$$

allows us to compute the time dependence of ϕ and π:

$$\begin{aligned}
i\frac{\partial}{\partial t}\phi(\mathbf{x},t) &= \Big[\phi(\mathbf{x},t), \int d^3x' \Big\{\tfrac{1}{2}\pi^2(\mathbf{x}',t) + \tfrac{1}{2}\big(\nabla\phi(\mathbf{x}',t)\big)^2 + \tfrac{1}{2}m^2\phi^2(\mathbf{x}',t)\Big\}\Big] \\
&= \int d^3x' \Big(i\delta^{(3)}(\mathbf{x}-\mathbf{x}')\pi(\mathbf{x}',t)\Big) \\
&= i\pi(\mathbf{x},t); \\
i\frac{\partial}{\partial t}\pi(\mathbf{x},t) &= \Big[\pi(\mathbf{x},t), \int d^3x' \Big\{\tfrac{1}{2}\pi^2(\mathbf{x}',t) + \tfrac{1}{2}\phi(\mathbf{x}',t)\big(-\nabla^2+m^2\big)\phi(\mathbf{x}',t)\Big\}\Big] \\
&= \int d^3x' \Big(-i\delta^{(3)}(\mathbf{x}-\mathbf{x}')\big(-\nabla^2+m^2\big)\phi(\mathbf{x}',t)\Big) \\
&= -i\big(-\nabla^2+m^2\big)\phi(\mathbf{x},t).
\end{aligned}$$

Combining the two results gives

$$\frac{\partial^2}{\partial t^2}\phi = \big(\nabla^2 - m^2\big)\phi, \tag{2.45}$$

which is just the Klein-Gordon equation.

We can better understand the time dependence of $\phi(x)$ and $\pi(x)$ by writing them in terms of creation and annihilation operators. First note that

$$Ha_{\mathbf{p}} = a_{\mathbf{p}}(H - E_{\mathbf{p}}),$$

and hence

$$H^n a_{\mathbf{p}} = a_{\mathbf{p}}(H - E_{\mathbf{p}})^n,$$

for any n. A similar relation (with $-$ replaced by $+$) holds for $a^\dagger_{\mathbf{p}}$. Thus we have derived the identities

$$e^{iHt}a_{\mathbf{p}}e^{-iHt} = a_{\mathbf{p}}e^{-iE_{\mathbf{p}}t}, \qquad e^{iHt}a^\dagger_{\mathbf{p}}e^{-iHt} = a^\dagger_{\mathbf{p}}e^{iE_{\mathbf{p}}t}, \tag{2.46}$$

which we can use on expression (2.25) for $\phi(\mathbf{x})$ to find the desired expression for the Heisenberg operator $\phi(x)$, according to (2.43). (We will always use the symbols $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$ to represent the time-independent, Schrödinger-picture ladder operators.) The result is

$$\begin{aligned}\phi(\mathbf{x},t) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\left(a_{\mathbf{p}}e^{-ip\cdot x} + a_{\mathbf{p}}^\dagger e^{ip\cdot x}\right)\Big|_{p^0=E_{\mathbf{p}}};\\ \pi(\mathbf{x},t) &= \frac{\partial}{\partial t}\phi(\mathbf{x},t).\end{aligned} \tag{2.47}$$

It is worth mentioning that we can perform the same manipulations with $\mathbf{P}$ instead of H to relate $\phi(\mathbf{x})$ to $\phi(0)$. In analogy with (2.46), one can show

$$e^{-i\mathbf{P}\cdot\mathbf{x}}a_{\mathbf{p}}e^{i\mathbf{P}\cdot\mathbf{x}} = a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}}, \qquad e^{-i\mathbf{P}\cdot\mathbf{x}}a_{\mathbf{p}}^\dagger e^{i\mathbf{P}\cdot\mathbf{x}} = a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}, \tag{2.48}$$

and therefore

$$\begin{aligned}\phi(x) &= e^{i(Ht-\mathbf{P}\cdot\mathbf{x})}\phi(0)e^{-i(Ht-\mathbf{P}\cdot\mathbf{x})}\\ &= e^{iP\cdot x}\phi(0)e^{-iP\cdot x},\end{aligned} \tag{2.49}$$

where $P^\mu = (H,\mathbf{P})$. (The notation here is confusing but standard. Remember that $\mathbf{P}$ is the momentum operator, whose eigenvalue is the total momentum of the system. On the other hand, $\mathbf{p}$ is the momentum of a single Fourier mode of the field, which we interpret as the momentum of a particle in that mode. For a one-particle state of well-defined momentum, $\mathbf{p}$ is the eigenvalue of $\mathbf{P}$.)

Equation (2.47) makes explicit the dual particle and wave interpretations of the quantum field $\phi(x)$. On the one hand, $\phi(x)$ is written as a Hilbert space operator, which creates and destroys the particles that are the quanta of field excitation. On the other hand, $\phi(x)$ is written as a linear combination of solutions ($e^{ip\cdot x}$ and $e^{-ip\cdot x}$) of the Klein-Gordon equation. Both signs of the time dependence in the exponential appear: We find both e^{-ip^0t} and e^{+ip^0t}, although p^0 is always positive. If these were single-particle wavefunctions, they would correspond to states of positive and negative energy; let us refer to them more generally as *positive-* and *negative-frequency* modes. The connection between the particle creation operators and the waveforms displayed here is always valid for free quantum fields: A positive-frequency solution of the field equation has as its coefficient the operator that *destroys* a particle in that single-particle wavefunction. A negative-frequency solution of the field equation, being the Hermitian conjugate of a positive-frequency solution, has as its coefficient the operator that *creates* a particle in that positive-energy single-particle wavefunction. In this way, the fact that relativistic wave equations have both positive- and negative-frequency solutions is reconciled with the requirement that a sensible quantum theory contain only positive excitation energies.

Causality

Now let us return to the question of causality raised at the beginning of this chapter. In our present formalism, still working in the Heisenberg picture, the amplitude for a particle to propagate from y to x is $\langle 0|\,\phi(x)\phi(y)\,|0\rangle$. We will call this quantity $D(x-y)$. Each operator ϕ is a sum of a and $a^\dagger$ operators, but only the term $\langle 0|\,a_{\mathbf{p}}a^\dagger_{\mathbf{q}}\,|0\rangle = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})$ survives in this expression. It is easy to check that we are left with

$$D(x-y) = \langle 0|\,\phi(x)\phi(y)\,|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}e^{-ip\cdot(x-y)}. \tag{2.50}$$

We have already argued in (2.40) that integrals of this form are Lorentz invariant. Let us now evaluate this integral for some particular values of $x-y$.

First consider the case where the difference $x-y$ is purely in the time-direction: $x^0-y^0=t$, $\mathbf{x}-\mathbf{y}=0$. (If the interval from y to x is timelike, there is always a frame in which this is the case.) Then we have

$$\begin{aligned} D(x-y) &= \frac{4\pi}{(2\pi)^3}\int_0^\infty dp\,\frac{p^2}{2\sqrt{p^2+m^2}}e^{-i\sqrt{p^2+m^2}t} \\ &= \frac{1}{4\pi^2}\int_m^\infty dE\,\sqrt{E^2-m^2}\,e^{-iEt} \\ &\underset{t\to\infty}{\sim} e^{-imt}. \end{aligned} \tag{2.51}$$

Next consider the case where $x-y$ is purely spatial: $x^0-y^0=0$, $\mathbf{x}-\mathbf{y}=\mathbf{r}$. The amplitude is then

$$\begin{aligned} D(x-y) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}\,e^{i\mathbf{p}\cdot\mathbf{r}} \\ &= \frac{2\pi}{(2\pi)^3}\int_0^\infty dp\,\frac{p^2}{2E_{\mathbf{p}}}\,\frac{e^{ipr}-e^{-ipr}}{ipr} \\ &= \frac{-i}{2(2\pi)^2 r}\int_{-\infty}^\infty dp\,\frac{p\,e^{ipr}}{\sqrt{p^2+m^2}}. \end{aligned}$$

The integrand, considered as a complex function of p, has branch cuts on the imaginary axis starting at $\pm im$ (see Fig. 2.3). To evaluate the integral we push the contour up to wrap around the upper branch cut. Defining $\rho=-ip$, we obtain

$$\frac{1}{4\pi^2 r}\int_m^\infty d\rho\,\frac{\rho\,e^{-\rho r}}{\sqrt{\rho^2-m^2}} \underset{r\to\infty}{\sim} e^{-mr}. \tag{2.52}$$

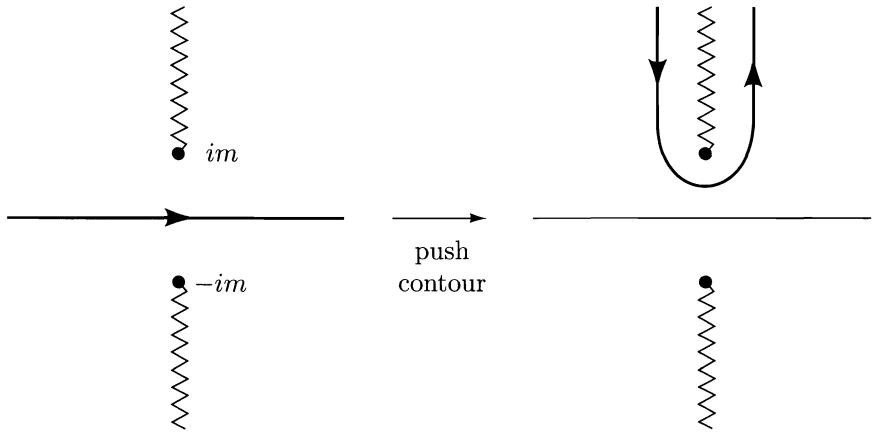


Figure 2.3. Contour for evaluating propagation amplitude $D(x-y)$ over a spacelike interval.

So again we find that outside the light-cone, the propagation amplitude is exponentially vanishing but nonzero.

To really discuss causality, however, we should ask not whether particles can propagate over spacelike intervals, but whether a *measurement* performed at one point can affect a measurement at another point whose separation from the first is spacelike. The simplest thing we could try to measure is the field $\phi(x)$, so we should compute the commutator $[\phi(x), \phi(y)]$; if this commutator vanishes, one measurement cannot affect the other. In fact, if the commutator vanishes for $(x-y)^2 < 0$, causality is preserved quite generally, since commutators involving any function of $\phi(x)$, including $\pi(x) = \partial\phi/\partial t$, would also have to vanish. Of course we know from Eq. (2.20) that the commutator vanishes for $x^0 = y^0$; now let's do the more general computation:

$$\begin{aligned}
\big[\phi(x), \phi(y)\big] &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \int \frac{d^3q}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{q}}}} \\
&\qquad \times \Big[\big(a_{\mathbf{p}} e^{-ip\cdot x} + a^\dagger_{\mathbf{p}} e^{ip\cdot x}\big), \big(a_{\mathbf{q}} e^{-iq\cdot y} + a^\dagger_{\mathbf{q}} e^{iq\cdot y}\big)\Big] \\
&= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \big(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\big) \\
&= D(x-y) - D(y-x). \qquad (2.53)
\end{aligned}$$

When $(x-y)^2 < 0$, we can perform a Lorentz transformation on the second term (since each term is separately Lorentz invariant), taking $(x-y) \to -(x-y)$, as shown in Fig. 2.4. The two terms are therefore equal and cancel to give zero; causality is preserved. Note that if $(x-y)^2 > 0$ there is no continuous Lorentz transformation that takes $(x-y) \longrightarrow -(x-y)$. In this case, by Eq. (2.51), the amplitude is (fortunately) nonzero, roughly $(e^{-imt} - e^{imt})$ for the special case $\mathbf{x} - \mathbf{y} = 0$. Thus we conclude that no measurement in the

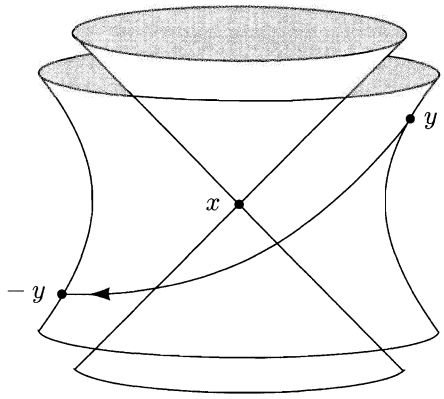


Figure 2.4. When $x - y$ is spacelike, a continuous Lorentz transformation can take $(x - y)$ to $-(x - y)$.

Klein-Gordon theory can affect another measurement outside the light-cone.

Causality is maintained in the Klein-Gordon theory just as suggested at the end of Section 2.1. To understand this mechanism properly, however, we should broaden the context of our discussion to include a *complex* Klein-Gordon field, which has distinct particle and antiparticle excitations. As was mentioned in the discussion of Eq. (2.15), we can add a conserved charge to the Klein-Gordon theory by considering the field $\phi(x)$ to be complex- rather than real-valued. When the complex scalar field theory is quantized (see Problem 2.2), $\phi(x)$ will create positively charged particles and destroy negatively charged ones, while $\phi^\dagger(x)$ will perform the opposite operations. Then the commutator $[\phi(x), \phi^\dagger(y)]$ will have nonzero contributions, which must delicately cancel outside the light-cone to preserve causality. The two contributions have the spacetime interpretation of the two terms in (2.53), but with charges attached. The first term will represent the propagation of a negatively charged particle from y to x. The second term will represent the propagation of a positively charged particle from x to y. In order for these two processes to be present and give canceling amplitudes, both of these particles must exist, and they must have the same mass. In quantum field theory, then, causality requires that every particle have a corresponding antiparticle with the same mass and opposite quantum numbers (in this case electric charge). For the real-valued Klein-Gordon field, the particle is its own antiparticle.

The Klein-Gordon Propagator

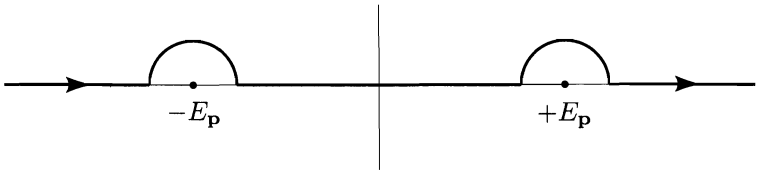
Let us study the commutator $[\phi(x), \phi(y)]$ a little further. Since it is a c-number, we can write $[\phi(x), \phi(y)] = \langle 0| \, [\phi(x), \phi(y)] \, |0\rangle$. This can be rewritten as a four-dimensional integral as follows, assuming for now that $x^0 > y^0$:

$$\langle 0| \, [\phi(x), \phi(y)] \, |0\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right)$$

$$= \int \frac{d^3p}{(2\pi)^3} \left\{ \frac{1}{2E_{\mathbf{p}}} e^{-ip\cdot(x-y)} \Big|_{p^0=E_{\mathbf{p}}} + \frac{1}{-2E_{\mathbf{p}}} e^{-ip\cdot(x-y)} \Big|_{p^0=-E_{\mathbf{p}}} \right\}$$

$$\underset{x^0>y^0}{=} \int \frac{d^3p}{(2\pi)^3} \int \frac{dp^0}{2\pi i} \frac{-1}{p^2 - m^2} e^{-ip\cdot(x-y)}. \tag{2.54}$$

In the last step the p^0 integral is to be performed along the following contour:



For $x^0 > y^0$ we can close the contour below, picking up both poles to obtain the previous line of (2.54). For $x^0 < y^0$ we may close the contour above, giving zero. Thus the last line of (2.54), together with the prescription for going around the poles, is an expression for what we will call

$$D_R(x-y) \equiv \theta(x^0 - y^0) \langle 0 | [\phi(x), \phi(y)] | 0 \rangle . \tag{2.55}$$

To understand this quantity better, let's do another computation:

$$\begin{aligned}(\partial^2 + m^2) D_R(x-y) &= \big(\partial^2 \theta(x^0 - y^0)\big) \langle 0 | [\phi(x), \phi(y)] | 0 \rangle \\ &\quad + 2\big(\partial_\mu \theta(x^0 - y^0)\big)\big(\partial^\mu \langle 0 | [\phi(x), \phi(y)] | 0 \rangle\big) \\ &\quad + \theta(x^0 - y^0) (\partial^2 + m^2) \langle 0 | [\phi(x), \phi(y)] | 0 \rangle \\ &= -\delta(x^0 - y^0) \langle 0 | [\pi(x), \phi(y)] | 0 \rangle \\ &\quad + 2\delta(x^0 - y^0) \langle 0 | [\pi(x), \phi(y)] | 0 \rangle + 0 \\ &= -i\delta^{(4)}(x-y). \end{aligned} \tag{2.56}$$

This says that $D_R(x-y)$ is a Green's function of the Klein-Gordon operator. Since it vanishes for $x^0 < y^0$, it is the *retarded* Green's function.

If we had not already derived expression (2.54), we could find it by Fourier transformation. Writing

$$D_R(x-y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \widetilde{D}_R(p), \tag{2.57}$$

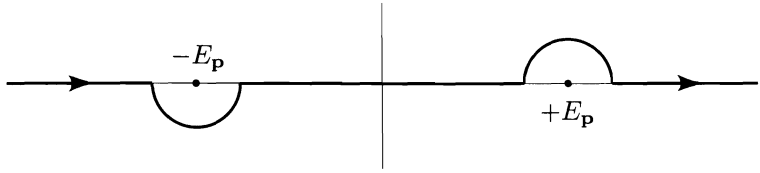
we obtain an algebraic expression for $\widetilde{D}_R(p)$:

$$(-p^2 + m^2) \widetilde{D}_R(p) = -i.$$

Thus we immediately arrive at the result

$$D_R(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{-ip\cdot(x-y)}. \tag{2.58}$$

The p^0-integral of (2.58) can be evaluated according to four different contours, of which that used in (2.54) is only one. In Chapter 4 we will find that a different pole prescription,



is extremely useful; it is called the *Feynman prescription.* A convenient way to remember it is to write

$$D_F(x-y) \equiv \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)}, \tag{2.59}$$

since the poles are then at $p^0 = \pm(E_{\mathbf{p}} - i\epsilon)$, displaced properly above and below the real axis. When $x^0 > y^0$ we can perform the p^0 integral by closing the contour below, obtaining exactly the propagation amplitude $D(x-y)$ (2.50). When $x^0 < y^0$ we close the contour above, obtaining the same expression but with x and y interchanged. Thus we have

$$\begin{aligned} D_F(x-y) &= \begin{cases} D(x-y) & \text{for } x^0 > y^0 \\ D(y-x) & \text{for } x^0 < y^0 \end{cases} \\ &= \theta(x^0 - y^0) \langle 0| \phi(x)\phi(y) |0\rangle + \theta(y^0 - x^0) \langle 0| \phi(y)\phi(x) |0\rangle \\ &\equiv \langle 0| T\phi(x)\phi(y) |0\rangle . \end{aligned} \tag{2.60}$$

The last line defines the "time-ordering" symbol T, which instructs us to place the operators that follow in order with the latest to the left. By applying $(\partial^2 + m^2)$ to the last line, you can verify directly that D_F is a Green's function of the Klein-Gordon operator.

Equations (2.59) and (2.60) are, from a practical point of view, the most important results of this chapter. The Green's function $D_F(x-y)$ is called the *Feynman propagator* for a Klein-Gordon particle, since it is, after all, a propagation amplitude. Indeed, the Feynman propagator will turn out to be part of the Feynman rules: $D_F(x-y)$ (or $\widetilde{D}_F(p)$) is the expression that we will attach to internal lines of Feynman diagrams, representing the propagation of virtual particles.

Nevertheless we are still a long way from being able to do any real calculations, since so far we have talked only about the *free* Klein-Gordon theory, where the field equation is linear and there are no interactions. Individual particles live in their isolated modes, oblivious to each others' existence and to the existence of any other species of particles. In such a theory there is no hope of making any observations, by scattering or any other means. On the other hand, the formalism we have developed is extremely important, since the free theory forms the basis for doing perturbative calculations in the interacting theory.

Particle Creation by a Classical Source

There is one type of interaction, however, that we are already equipped to handle. Consider a Klein-Gordon field coupled to an external, classical source field $j(x)$. That is, consider the field equation

$$(\partial^2 + m^2)\phi(x) = j(x), \tag{2.61}$$

where $j(x)$ is some fixed, known function of space and time that is nonzero only for a finite time interval. If we start in the vacuum state, what will we find after $j(x)$ has been turned on and off again?

The field equation (2.61) follows from the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2 + j(x)\phi(x). \tag{2.62}$$

But if $j(x)$ is turned on for only a finite time, it is easiest to solve the problem using the field equation directly. Before $j(x)$ is turned on, $\phi(x)$ has the form

$$\phi_0(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\left(a_{\mathbf{p}}e^{-ip\cdot x} + a_{\mathbf{p}}^\dagger e^{ip\cdot x}\right).$$

If there were no source, this would be the solution for all time. With a source, the solution of the equation of motion can be constructed using the retarded Green's function:

$$\begin{aligned}\phi(x) &= \phi_0(x) + i\int d^4y\, D_R(x-y)j(y)\\ &= \phi_0(x) + i\int d^4y \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}\,\theta(x^0-y^0)\\ &\qquad\qquad \times \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right)j(y).\end{aligned} \tag{2.63}$$

If we wait until all of j is in the past, the theta function equals 1 in the whole domain of integration. Then $\phi(x)$ involves only the Fourier transform of j,

$$\tilde{j}(p) = \int d^4y\, e^{ip\cdot y}j(y),$$

evaluated at 4-momenta p such that $p^2 = m^2$. It is natural to group the positive-frequency terms together with $a_{\mathbf{p}}$ and the negative-frequency terms with $a_{\mathbf{p}}^\dagger$; this yields the expression

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\left\{\left(a_{\mathbf{p}} + \frac{i}{\sqrt{2E_{\mathbf{p}}}}\tilde{j}(p)\right)e^{-ip\cdot x} + \text{h.c.}\right\}. \tag{2.64}$$

You can now guess (or compute) the form of the Hamiltonian after $j(x)$ has acted: Just replace $a_{\mathbf{p}}$ with $\left(a_{\mathbf{p}} + i\tilde{j}(p)/\sqrt{2E_{\mathbf{p}}}\right)$ to obtain

$$H = \int \frac{d^3p}{(2\pi)^3}E_{\mathbf{p}}\left(a_{\mathbf{p}}^\dagger - \frac{i}{\sqrt{2E_{\mathbf{p}}}}\tilde{j}^*(p)\right)\left(a_{\mathbf{p}} + \frac{i}{\sqrt{2E_{\mathbf{p}}}}\tilde{j}(p)\right).$$

The energy of the system after the source has been turned off is

$$\langle 0|\, H\, |0\rangle = \int \frac{d^3p}{(2\pi)^3}\, \frac{1}{2} |\tilde{\jmath}(p)|^2, \tag{2.65}$$

where $|0\rangle$ still denotes the ground state of the free theory. We can interpret these results in terms of particles by identifying $|\tilde{\jmath}(p)|^2/2E_{\mathbf{p}}$ as the probability density for creating a particle in the mode p. Then the total number of particles produced is

$$\int dN = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} |\tilde{\jmath}(p)|^2. \tag{2.66}$$

Only those Fourier components of $j(x)$ that are in resonance with on-mass-shell (i.e., $p^2 = m^2$) Klein-Gordon waves are effective at creating particles.

We will return to this subject in Problem 4.1. In Chapter 6 we will study the analogous problem of photon creation by an accelerated electron (bremsstrahlung).

Problems

2.1 Classical electromagnetism (with no sources) follows from the action

$$S = \int d^4x \left(-\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu}\right), \qquad \text{where } F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

(a) Derive Maxwell's equations as the Euler-Lagrange equations of this action, treating the components $A_\mu(x)$ as the dynamical variables. Write the equations in standard form by identifying $E^i = -F^{0i}$ and $\epsilon^{ijk} B^k = -F^{ij}$.

(b) Construct the energy-momentum tensor for this theory. Note that the usual procedure does not result in a symmetric tensor. To remedy that, we can add to $T^{\mu\nu}$ a term of the form $\partial_\lambda K^{\lambda\mu\nu}$, where $K^{\lambda\mu\nu}$ is antisymmetric in its first two indices. Such an object is automatically divergenceless, so

$$\widehat{T}^{\mu\nu} = T^{\mu\nu} + \partial_\lambda K^{\lambda\mu\nu}$$

is an equally good energy-momentum tensor with the same globally conserved energy and momentum. Show that this construction, with

$$K^{\lambda\mu\nu} = F^{\mu\lambda} A^\nu,$$

leads to an energy-momentum tensor $\widehat{T}$ that is symmetric and yields the standard formulae for the electromagnetic energy and momentum densities:

$$\mathcal{E} = \tfrac{1}{2}(E^2 + B^2); \qquad \mathbf{S} = \mathbf{E} \times \mathbf{B}.$$

2.2 The complex scalar field. Consider the field theory of a complex-valued scalar field obeying the Klein-Gordon equation. The action of this theory is

$$S = \int d^4x\, (\partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi).$$

It is easiest to analyze this theory by considering $\phi(x)$ and $\phi^*(x)$, rather than the real and imaginary parts of $\phi(x)$, as the basic dynamical variables.

(a) Find the conjugate momenta to $\phi(x)$ and $\phi^*(x)$ and the canonical commutation relations. Show that the Hamiltonian is

$$H = \int d^3x \, (\pi^*\pi + \nabla\phi^* \cdot \nabla\phi + m^2\phi^*\phi).$$

Compute the Heisenberg equation of motion for $\phi(x)$ and show that it is indeed the Klein-Gordon equation.

(b) Diagonalize H by introducing creation and annihilation operators. Show that the theory contains two sets of particles of mass m.

(c) Rewrite the conserved charge

$$Q = \int d^3x \, \frac{i}{2}(\phi^*\pi^* - \pi\phi)$$

in terms of creation and annihilation operators, and evaluate the charge of the particles of each type.

(d) Consider the case of two complex Klein-Gordon fields with the same mass. Label the fields as $\phi_a(x)$, where $a = 1, 2$. Show that there are now four conserved charges, one given by the generalization of part (c), and the other three given by

$$Q^i = \int d^3x \, \frac{i}{2}(\phi_a^*(\sigma^i)_{ab}\pi_b^* - \pi_a(\sigma^i)_{ab}\phi_b),$$

where σ^i are the Pauli sigma matrices. Show that these three charges have the commutation relations of angular momentum ($SU(2)$). Generalize these results to the case of n identical complex scalar fields.‡

2.3 Evaluate the function

$$\langle 0|\, \phi(x)\phi(y)\, |0\rangle = D(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} e^{-ip\cdot(x-y)},$$

for $(x-y)$ spacelike so that $(x-y)^2 = -r^2$, explicitly in terms of Bessel functions.

‡With some additional work you can show that there are actually six conserved charges in the case of two complex fields, and $n(2n-1)$ in the case of n fields, corresponding to the generators of the rotation group in four and $2n$ dimensions, respectively. The extra symmetries often do not survive when nonlinear interactions of the fields are included.

Chapter 3

The Dirac Field

Having exhaustively treated the simplest relativistic field equation, we now move on to the second simplest, the Dirac equation. You may already be familiar with the Dirac equation in its original incarnation, that is, as a single-particle quantum-mechanical wave equation.* In this chapter our viewpoint will be quite different. First we will rederive the Dirac equation as a *classical* relativistic field equation, with special emphasis on its relativistic invariance. Then, in Section 3.5, we will quantize the Dirac field in a manner similar to that used for the Klein-Gordon field.

3.1 Lorentz Invariance in Wave Equations

First we must address a question that we swept over in Chapter 2: What do we mean when we say that an equation is "relativistically invariant"? A reasonable definition is the following: If ϕ is a field or collection of fields and $\mathcal{D}$ is some differential operator, then the statement "$\mathcal{D}\phi = 0$ is relativistically invariant" means that if $\phi(x)$ satisfies this equation, and we perform a rotation or boost to a different frame of reference, then the transformed field, in the new frame of reference, satisfies the same equation. Equivalently, we can imagine physically rotating or boosting all particles or fields by a common angle or velocity; again, the equation $\mathcal{D}\phi = 0$ should be true after the transformation. We will adopt this "active" point of view toward transformations in the following analysis.

The Lagrangian formulation of field theory makes it especially easy to discuss Lorentz invariance. An equation of motion is automatically Lorentz invariant by the above definition if it follows from a Lagrangian that is a Lorentz *scalar*. This is an immediate consequence of the principle of least action: If boosts leave the Lagrangian unchanged, the boost of an extremum in the action will be another extremum.

*This subject is covered, for example, in Schiff (1968), Chapter 13; Baym (1969), Chapter 23; Sakurai (1967), Chapter 3. Although the present chapter is self-contained, we recommend that you also study the single-particle Dirac equation at some point.

As an example, consider the Klein-Gordon theory. We can write an arbitrary Lorentz transformation as

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{3.1}$$

for some 4×4 matrix Λ. What happens to the Klein-Gordon field $\phi(x)$ under this transformation? Think of the field ϕ as measuring the local value of some quantity that is distributed through space. If there is an accumulation of this quantity at $x = x_0$, $\phi(x)$ will have a maximum at x_0. If we now transform the original distribution by a boost, the new distribution will have a maximum at $x = \Lambda x_0$. This is illustrated in Fig. 3.1(a). The corresponding transformation of the field is

$$\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x). \tag{3.2}$$

That is, the transformed field, evaluated at the boosted point, gives the same value as the original field evaluated at the point before boosting.

We should check that this transformation leaves the form of the Klein-Gordon Lagrangian unchanged. According to (3.2), the mass term $\frac{1}{2}m^2\phi^2(x)$ is simply shifted to the point $(\Lambda^{-1}x)$. The transformation of $\partial_\mu\phi(x)$ is

$$\partial_\mu\phi(x) \to \partial_\mu\big(\phi(\Lambda^{-1}x)\big) = (\Lambda^{-1})^\nu{}_\mu(\partial_\nu\phi)(\Lambda^{-1}x). \tag{3.3}$$

Since the metric tensor $g^{\mu\nu}$ is Lorentz invariant, the matrices Λ^{-1} obey the identity

$$(\Lambda^{-1})^\rho{}_\mu\,(\Lambda^{-1})^\sigma{}_\nu\, g^{\mu\nu} = g^{\rho\sigma}. \tag{3.4}$$

Using this relation, we can compute the transformation law of the kinetic term of the Klein-Gordon Lagrangian:

$$\begin{aligned}(\partial_\mu\phi(x))^2 &\to g^{\mu\nu}\big(\partial_\mu\phi'(x)\big)\big(\partial_\nu\phi'(x)\big)\\ &= g^{\mu\nu}\big[(\Lambda^{-1})^\rho{}_\mu\partial_\rho\phi\big]\big[(\Lambda^{-1})^\sigma{}_\nu\partial_\sigma\phi\big](\Lambda^{-1}x)\\ &= g^{\rho\sigma}\big(\partial_\rho\phi\big)\big(\partial_\sigma\phi\big)(\Lambda^{-1}x)\\ &= (\partial_\mu\phi)^2(\Lambda^{-1}x).\end{aligned}$$

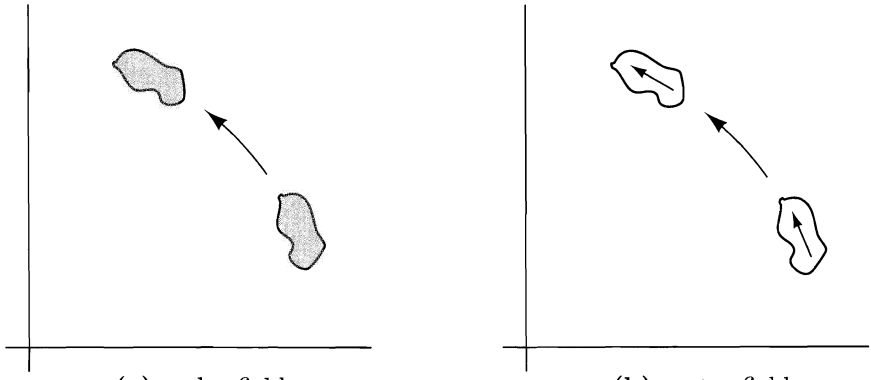
Thus, the whole Lagrangian is simply transformed as a scalar:

$$\mathcal{L}(x) \to \mathcal{L}(\Lambda^{-1}x). \tag{3.5}$$

The action S, formed by integrating $\mathcal{L}$ over spacetime, is Lorentz invariant. A similar calculation shows that the equation of motion is invariant:

$$\begin{aligned}(\partial^2 + m^2)\phi'(x) &= \big[(\Lambda^{-1})^\nu{}_\mu\partial_\nu(\Lambda^{-1})^{\sigma\mu}\partial_\sigma + m^2\big]\,\phi(\Lambda^{-1}x)\\ &= (g^{\nu\sigma}\partial_\nu\partial_\sigma + m^2)\phi(\Lambda^{-1}x)\\ &= 0.\end{aligned}$$

The transformation law (3.2) used for ϕ is the simplest possible transformation law for a field. It is the only possibility for a field that has just one component. But we know examples of multiple-component fields that transform in more complicated ways. The most familiar case is that of a vector field,



(a) scalar field **(b)** vector field

Figure 3.1. When a rotation is performed on a vector field, it affects the *orientation* of the vector as well as the location of the region containing the configuration.

such as the 4-current density $j^\mu(x)$ or the vector potential $A^\mu(x)$. In this case, the quantity that is distributed in spacetime also carries an orientation, which must be rotated or boosted. As shown in Fig. 3.1(b), the orientation must be rotated forward as the point of evaluation of the field is changed:

$$\begin{aligned} &\text{under 3-dimensional rotations,} && V^i(x) \to R^{ij}V^j(R^{-1}x); \\ &\text{under Lorentz transformations,} && V^\mu(x) \to \Lambda^\mu{}_\nu V^\nu(\Lambda^{-1}x). \end{aligned}$$

Tensors of arbitrary rank can be built out of vectors by adding more indices, with correspondingly more factors of Λ in the transformation law. Using such vector and tensor fields we can write a variety of Lorentz-invariant equations, for example, Maxwell's equations,

$$\partial^\mu F_{\mu\nu} = 0 \qquad \text{or} \qquad \partial^2 A_\nu - \partial_\nu \partial^\mu A_\mu = 0, \tag{3.6}$$

which follow from the Lagrangian

$$\mathcal{L}_{\text{Maxwell}} = -\tfrac{1}{4}(F_{\mu\nu})^2 = -\tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2. \tag{3.7}$$

In general, any equation in which each term has the same set of uncontracted Lorentz indices will naturally be invariant under Lorentz transformations.

This method of tensor notation yields a large class of Lorentz-invariant equations, but it turns out that there are still more. How do we find them? We could try to systematically find all possible transformation laws for a field. Then it would not be hard to write invariant Lagrangians. For simplicity, we will restrict our attention to linear transformations, so that, if Φ_a is an n component multiplet, the Lorentz transformation law is given by an $n \times n$ matrix $M(\Lambda)$:

$$\Phi_a(x) \to M_{ab}(\Lambda)\Phi_b(\Lambda^{-1}x). \tag{3.8}$$

It can be shown that the most general nonlinear transformation laws can be built from these linear transformations, so there is no advantage in considering transformations more general than (3.8). In the following discussion, we will suppress the change in the field argument and write the transformation (3.8) in the form

$$\Phi \to M(\Lambda)\Phi. \tag{3.9}$$

What are the possible allowed forms for the matrices $M(\Lambda)$? The basic restriction on $M(\Lambda)$ is found by imagining two successive transformations, Λ and Λ'. The net result must be a new Lorentz transformation Λ''; that is, the Lorentz transformations form a *group*. This gives a consistency condition that must be satisfied by the matrices $M(\Lambda)$: Under the sequence of two transformations,

$$\Phi \to M(\Lambda')M(\Lambda)\Phi = M(\Lambda'')\Phi, \tag{3.10}$$

for $\Lambda'' = \Lambda'\Lambda$. Thus the correspondence between the matrices M and the transformations Λ must be preserved under multiplication. In mathematical language, we say that the matrices M must form an n-dimensional *representation* of the Lorentz group. So our question now is rephrased in mathematical language: What are the (finite-dimensional) matrix representations of the Lorentz group?

Before answering this question for the Lorentz group, let us consider a simpler group, the rotation group in three dimensions. This group has representations of every dimensionality n, familiar in quantum mechanics as the matrices that rotate the n-component wavefunctions of particles of different spins. The dimensionality is related to the spin quantum number s by $n = 2s + 1$. The most important nontrivial representation is the two-dimensional representation, corresponding to spin 1/2. The matrices of this representation are the 2×2 unitary matrices with determinant 1, which can be expressed as

$$U = e^{-i\theta^i \sigma^i/2}, \tag{3.11}$$

where θ^i are three arbitrary parameters and σ^i are the Pauli sigma matrices.

For any continuous group, the transformations that lie infinitesimally close to the identity define a vector space, called the *Lie algebra* of the group. The basis vectors for this vector space are called the *generators* of the Lie algebra, or of the group. For the rotation group, the generators are the angular momentum operators J^i, which satisfy the commutation relations

$$\left[J^i, J^j\right] = i\epsilon^{ijk}J^k. \tag{3.12}$$

The finite rotation operations are formed by exponentiating these operators: In quantum mechanics, the operator

$$R = \exp\left[-i\theta^i J^i\right] \tag{3.13}$$

gives the rotation by an angle $|\theta|$ about the axis $\hat{\theta}$. The commutation relations of the operators J^i determine the multiplication laws of these rotation

operators. Thus, a set of matrices satisfying the commutation relations (3.12) produces, through exponentiation as in (3.13), a representation of the rotation group. In the example given in the previous paragraph, the representation of the angular momentum operators

$$J^i \to \frac{\sigma^i}{2} \tag{3.14}$$

produces the representation of the rotation group given in Eq. (3.11). It is generally true that one can find matrix representations of a continuous group by finding matrix representations of the generators of the group (which must satisfy the proper commutation relations), then exponentiating these infinitesimal transformations.

For our present problem, we need to know the commutation relations of the generators of the group of Lorentz transformations. For the rotation group, one can work out the commutation relations by writing the generators as differential operators; from the expression

$$\mathbf{J} = \mathbf{x} \times \mathbf{p} = \mathbf{x} \times (-i\nabla), \tag{3.15}$$

the angular momentum commutation relations (3.12) follow straightforwardly. The use of the cross product in (3.15) is special to the case of three dimensions. However, we can also write the operators as an antisymmetric tensor,

$$J^{ij} = -i(x^i \nabla^j - x^j \nabla^i),$$

so that $J^3 = J^{12}$ and so on. The generalization to four-dimensional Lorentz transformations is now quite natural:

$$J^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu). \tag{3.16}$$

We will soon see that these six operators generate the three boosts and three rotations of the Lorentz group.

To determine the commutation rules of the Lorentz algebra, we can now simply compute the commutators of the differential operators (3.16). The result is

$$[J^{\mu\nu}, J^{\rho\sigma}] = i\big(g^{\nu\rho} J^{\mu\sigma} - g^{\mu\rho} J^{\nu\sigma} - g^{\nu\sigma} J^{\mu\rho} + g^{\mu\sigma} J^{\nu\rho}\big). \tag{3.17}$$

Any matrices that are to represent this algebra must obey these same commutation rules.

Just to see that we have this right, let us look at one particular representation (which we will simply pull out of a hat). Consider the 4×4 matrices

$$(\mathcal{J}^{\mu\nu})_{\alpha\beta} = i(\delta^\mu{}_\alpha \delta^\nu{}_\beta - \delta^\mu{}_\beta \delta^\nu{}_\alpha). \tag{3.18}$$

(Here μ and ν label which of the six matrices we want, while α and β label components of the matrices.) You can easily verify that these matrices satisfy the commutation relations (3.17). In fact, they are nothing but the

matrices that act on ordinary Lorentz 4-vectors. To see this, parametrize an infinitesimal transformation as follows:

$$V^\alpha \to \left(\delta^\alpha{}_\beta - \frac{i}{2}\omega_{\mu\nu}(\mathcal{J}^{\mu\nu})^\alpha{}_\beta\right)V^\beta, \tag{3.19}$$

where V is a 4-vector and $\omega_{\mu\nu}$, an antisymmetric tensor, gives the infinitesimal angles. For example, consider the case $\omega_{12} = -\omega_{21} = \theta$, with all other components of ω equal to zero. Then Eq. (3.19) becomes

$$V \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\theta & 0 \\ 0 & \theta & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} V, \tag{3.20}$$

which is just an infinitesimal rotation in the xy-plane. You can also verify that setting $\omega_{01} = -\omega_{10} = \beta$ gives

$$V \to \begin{pmatrix} 1 & \beta & 0 & 0 \\ \beta & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} V, \tag{3.21}$$

an infinitesimal boost in the x-direction. The other components of ω generate the remaining boosts and rotations in a similar manner.

3.2 The Dirac Equation

Now that we have seen one finite-dimensional representation of the Lorentz group, the logical next step would be to develop the formalism for finding all other representations. Although this is not very difficult to do (see Problem 3.1), it is hardly necessary for our purposes, since we are mainly interested in the representation(s) corresponding to spin 1/2.

We can find such a representation using a trick due to Dirac: Suppose that we had a set of four $n \times n$ matrices γ^μ satisfying the anticommutation relations

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \times \mathbf{1}_{n\times n} \qquad \text{(Dirac algebra)}. \tag{3.22}$$

Then we could immediately write down an n-dimensional representation of the Lorentz algebra. Here it is:

$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]. \tag{3.23}$$

By repeated use of (3.22), it is easy to verify that these matrices satisfy the commutation relations (3.17).

This computation goes through in any dimensionality, with Lorentz or Euclidean metric. In particular, it should work in three-dimensional Euclidean

space, and in fact we can simply write

$$\gamma^j \equiv i\sigma^j \qquad \text{(Pauli sigma matrices)},$$

$$\text{so that} \qquad \{\gamma^i, \gamma^j\} = -2\delta^{ij}.$$

The factor of i in the first line and the minus sign in the second line are purely conventional. The matrices representing the Lorentz algebra are then

$$S^{ij} = \tfrac{1}{2}\epsilon^{ijk}\sigma^k, \tag{3.24}$$

which we recognize as the two-dimensional representation of the rotation group.

Now let us find Dirac matrices γ^μ for four-dimensional Minkowski space. It turns out that these matrices must be at least 4×4. (There is no fourth 2×2 matrix, for example, that anticommutes with the three Pauli sigma matrices.) Further, all 4×4 representations of the Dirac algebra are unitarily equivalent.† We thus need only write one explicit realization of the Dirac algebra. One representation, in 2×2 block form, is

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \tag{3.25}$$

This representation is called the *Weyl* or *chiral* representation. We will find it an especially convenient choice, and we will use it exclusively throughout this book. (Be careful, however, since many field theory textbooks choose a different representation, in which γ^0 is diagonal. Furthermore, books that use chiral representations often make a different choice of sign conventions.)

In our representation, the boost and rotation generators are

$$S^{0i} = \frac{i}{4}[\gamma^0, \gamma^i] = -\frac{i}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix}, \tag{3.26}$$

and

$$S^{ij} = \frac{i}{4}[\gamma^i, \gamma^j] = \frac{1}{2}\epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \equiv \frac{1}{2}\epsilon^{ijk}\Sigma^k. \tag{3.27}$$

A four-component field ψ that transforms under boosts and rotations according to (3.26) and (3.27) is called a *Dirac spinor*. Note that the rotation generator S^{ij} is just the three-dimensional spinor transformation matrix (3.24) replicated twice. The boost generators S^{0i} are not Hermitian, and thus our implementation of boosts is not unitary (this was also true of the vector representation (3.18)). In fact the Lorentz group, being "noncompact", has no faithful, finite-dimensional representations that are unitary. But that does not matter to us, since ψ is not a wavefunction; it is a classical field.

†This statement and the preceding one follow from the general theory of the representations of the Lorentz group derived in Problem 3.1.

Now that we have the transformation law for ψ, we should look for an appropriate field equation. One possibility is simply the Klein-Gordon equation:

$$(\partial^2 + m^2)\psi = 0. \tag{3.28}$$

This works because the spinor transformation matrices (3.26) and (3.27) operate only in the "internal" space; they go right through the differential operator. But it is possible to write a stronger, first-order equation, which implies (3.28) but contains additional information. To do this we need to know one more property of the γ matrices. With a short computation you can verify that

$$[\gamma^\mu, S^{\rho\sigma}] = (\mathcal{J}^{\rho\sigma})^\mu{}_\nu \gamma^\nu,$$

or equivalently,

$$\big(1 + \tfrac{i}{2}\omega_{\rho\sigma}S^{\rho\sigma}\big)\gamma^\mu\big(1 - \tfrac{i}{2}\omega_{\rho\sigma}S^{\rho\sigma}\big) = \big(1 - \tfrac{i}{2}\omega_{\rho\sigma}\mathcal{J}^{\rho\sigma}\big)^\mu{}_\nu \gamma^\nu.$$

This equation is just the infinitesimal form of

$$\Lambda_{\frac{1}{2}}^{-1}\gamma^\mu\Lambda_{\frac{1}{2}} = \Lambda^\mu{}_\nu\gamma^\nu, \tag{3.29}$$

where

$$\Lambda_{\frac{1}{2}} = \exp\big(-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}\big) \tag{3.30}$$

is the spinor representation of the Lorentz transformation Λ (compare (3.19)). Equation (3.29) says that the γ matrices are invariant under simultaneous rotations of their vector and spinor indices (just like the σ^i under spatial rotations). In other words, we can "take the vector index μ on γ^μ seriously," and dot γ^μ into ∂_μ to form a Lorentz-invariant differential operator.

We are now ready to write down the Dirac equation. Here it is:

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = 0. \tag{3.31}$$

To show that it is Lorentz invariant, write down the Lorentz-transformed version of the left-hand side and calculate:

$$\begin{aligned}
\big[i\gamma^\mu\partial_\mu - m\big]\psi(x) &\to \big[i\gamma^\mu(\Lambda^{-1})^\nu{}_\mu\partial_\nu - m\big]\Lambda_{\frac{1}{2}}\psi(\Lambda^{-1}x) \\
&= \Lambda_{\frac{1}{2}}\Lambda_{\frac{1}{2}}^{-1}\big[i\gamma^\mu(\Lambda^{-1})^\nu{}_\mu\partial_\nu - m\big]\Lambda_{\frac{1}{2}}\psi(\Lambda^{-1}x) \\
&= \Lambda_{\frac{1}{2}}\big[i\Lambda_{\frac{1}{2}}^{-1}\gamma^\mu\Lambda_{\frac{1}{2}}(\Lambda^{-1})^\nu{}_\mu\partial_\nu - m\big]\psi(\Lambda^{-1}x) \\
&= \Lambda_{\frac{1}{2}}\big[i\Lambda^\mu{}_\sigma\gamma^\sigma(\Lambda^{-1})^\nu{}_\mu\partial_\nu - m\big]\psi(\Lambda^{-1}x) \\
&= \Lambda_{\frac{1}{2}}\big[i\gamma^\nu\partial_\nu - m\big]\psi(\Lambda^{-1}x) \\
&= 0.
\end{aligned}$$

To see that the Dirac equation implies the Klein-Gordon equation, act on the left with $(-i\gamma^\mu\partial_\mu - m)$:

$$\begin{aligned} 0 &= (-i\gamma^\mu\partial_\mu - m)(i\gamma^\nu\partial_\nu - m)\psi \\ &= (\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + m^2)\psi \\ &= (\tfrac{1}{2}\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu + m^2)\psi \\ &= (\partial^2 + m^2)\psi. \end{aligned}$$

To write down a Lagrangian for the Dirac theory, we must figure out how to multiply two Dirac spinors to form a Lorentz scalar. The obvious guess, $\psi^\dagger\psi$, does not work. Under a Lorentz boost this becomes $\psi^\dagger\Lambda_{\frac{1}{2}}^\dagger\Lambda_{\frac{1}{2}}\psi$; if the boost matrix were unitary, we would have $\Lambda_{\frac{1}{2}}^\dagger = \Lambda_{\frac{1}{2}}^{-1}$ and everything would be fine. But $\Lambda_{\frac{1}{2}}$ is not unitary, because the generators (3.26) are not Hermitian.

The solution is to define

$$\bar{\psi} \equiv \psi^\dagger\gamma^0. \tag{3.32}$$

Under an infinitesimal Lorentz transformation parametrized by $\omega_{\mu\nu}$, we have $\bar{\psi} \to \psi^\dagger(1 + \frac{i}{2}\omega_{\mu\nu}(S^{\mu\nu})^\dagger)\gamma^0$. The sum over μ and ν has six distinct nonzero terms. In the rotation terms, where μ and ν are both nonzero, $(S^{\mu\nu})^\dagger = S^{\mu\nu}$ and $S^{\mu\nu}$ commutes with γ^0. In the boost terms, where μ or ν is 0, $(S^{\mu\nu})^\dagger = -(S^{\mu\nu})$ but $S^{\mu\nu}$ anticommutes with γ^0. Passing the γ^0 to the left therefore removes the dagger from $S^{\mu\nu}$, yielding the transformation law

$$\bar{\psi} \to \bar{\psi}\Lambda_{\frac{1}{2}}^{-1}, \tag{3.33}$$

and therefore the quantity $\bar{\psi}\psi$ is a Lorentz scalar. Similarly you can show (with the aid of (3.29)) that $\bar{\psi}\gamma^\mu\psi$ is a Lorentz vector.

The correct, Lorentz-invariant Dirac Lagrangian is therefore

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi. \tag{3.34}$$

The Euler-Lagrange equation for $\bar{\psi}$ (or $\psi^\dagger$) immediately yields the Dirac equation in the form (3.31); the Euler-Lagrange equation for ψ gives the same equation, in Hermitian-conjugate form:

$$-i\partial_\mu\bar{\psi}\gamma^\mu - m\bar{\psi} = 0. \tag{3.35}$$

Weyl Spinors

From the block-diagonal form of the generators (3.26) and (3.27), it is apparent that the Dirac representation of the Lorentz group is *reducible.*[‡] We can form two 2-dimensional representations by considering each block separately, and writing

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \tag{3.36}$$

[‡]If we had used a different representation of the gamma matrices, the reducibility would not be manifest; this is essentially the reason for using the chiral representation.

The two-component objects ψ_L and ψ_R are called left-handed and right-handed *Weyl spinors*. You can easily verify that their transformation laws, under infinitesimal rotations $\boldsymbol{\theta}$ and boosts $\boldsymbol{\beta}$, are

$$\begin{aligned} \psi_L &\to (1 - i\boldsymbol{\theta} \cdot \tfrac{\boldsymbol{\sigma}}{2} - \boldsymbol{\beta} \cdot \tfrac{\boldsymbol{\sigma}}{2})\psi_L; \\ \psi_R &\to (1 - i\boldsymbol{\theta} \cdot \tfrac{\boldsymbol{\sigma}}{2} + \boldsymbol{\beta} \cdot \tfrac{\boldsymbol{\sigma}}{2})\psi_R. \end{aligned} \tag{3.37}$$

These transformation laws are connected by complex conjugation; using the identity

$$\sigma^2 \boldsymbol{\sigma}^* = -\boldsymbol{\sigma}\sigma^2, \tag{3.38}$$

it is not hard to show that the quantity $\sigma^2 \psi_L^*$ transforms like a right-handed spinor.

In terms of ψ_L and ψ_R, the Dirac equation is

$$(i\gamma^\mu \partial_\mu - m)\psi = \begin{pmatrix} -m & i(\partial_0 + \boldsymbol{\sigma} \cdot \boldsymbol{\nabla}) \\ i(\partial_0 - \boldsymbol{\sigma} \cdot \boldsymbol{\nabla}) & -m \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = 0. \tag{3.39}$$

The two Lorentz group representations ψ_L and ψ_R are mixed by the mass term in the Dirac equation. But if we set $m = 0$, the equations for ψ_L and ψ_R decouple:

$$\begin{aligned} i(\partial_0 - \boldsymbol{\sigma} \cdot \boldsymbol{\nabla})\psi_L &= 0; \\ i(\partial_0 + \boldsymbol{\sigma} \cdot \boldsymbol{\nabla})\psi_R &= 0. \end{aligned} \tag{3.40}$$

These are called the *Weyl equations*; they are especially important when treating neutrinos and the theory of weak interactions.

It is possible to clean up this notation slightly. Define

$$\sigma^\mu \equiv (1, \boldsymbol{\sigma}), \qquad \bar{\sigma}^\mu \equiv (1, -\boldsymbol{\sigma}), \tag{3.41}$$

so that

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}. \tag{3.42}$$

(The bar on $\bar{\sigma}$ has absolutely nothing to do with the bar on $\overline{\psi}$.) Then the Dirac equation can be written

$$\begin{pmatrix} -m & i\sigma \cdot \partial \\ i\bar{\sigma} \cdot \partial & -m \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = 0, \tag{3.43}$$

and the Weyl equations become

$$i\bar{\sigma} \cdot \partial \psi_L = 0; \qquad i\sigma \cdot \partial \psi_R = 0. \tag{3.44}$$

3.3 Free-Particle Solutions of the Dirac Equation

To get some feel for the physics of the Dirac equation, let us now discuss its plane-wave solutions. Since a Dirac field ψ obeys the Klein-Gordon equation, we know immediately that it can be written as a linear combination of plane waves:

$$\psi(x) = u(p)e^{-ip\cdot x}, \qquad \text{where } p^2 = m^2. \tag{3.45}$$

For the moment we will concentrate on solutions with positive frequency, that is, $p^0 > 0$. The column vector $u(p)$ must obey an additional constraint, found by plugging (3.45) into the Dirac equation:

$$(\gamma^\mu p_\mu - m)u(p) = 0. \tag{3.46}$$

It is easiest to analyze this equation in the rest frame, where $p = p_0 = (m, \mathbf{0})$; the solution for general p can then be found by boosting with $\Lambda_{\frac{1}{2}}$. In the rest frame, Eq. (3.46) becomes

$$(m\gamma^0 - m)u(p_0) = m \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} u(p_0) = 0,$$

and the solutions are

$$u(p_0) = \sqrt{m}\begin{pmatrix} \xi \\ \xi \end{pmatrix}, \tag{3.47}$$

for any numerical two-component spinor ξ. We conventionally normalize ξ so that $\xi^\dagger\xi = 1$; the factor $\sqrt{m}$ has been inserted for future convenience. We can interpret the spinor ξ by looking at the rotation generator (3.27): ξ transforms under rotations as an ordinary two-component spinor of the rotation group, and therefore determines the spin orientation of the Dirac solution in the usual way. For example, when $\xi = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, the particle has spin up along the 3-direction.

Notice that after applying the Dirac equation, we are free to choose only two of the four components of $u(p)$. This is just what we want, since a spin-1/2 particle has only two physical states—spin up and spin down. (Of course we are being a bit premature in talking about *particles* and *spin*. We will *prove* that the spin angular momentum of a Dirac particle is $\hbar/2$ when we quantize the Dirac theory in Section 3.5; for now, just notice that there are two possible solutions $u(p)$ for any momentum p.)

Now that we have the general form of $u(p)$ in the rest frame, we can obtain $u(p)$ in any other frame by boosting. Consider a boost along the 3-direction. First we should remind ourselves of what the boost does to the 4-momentum vector. In infinitesimal form,

$$\begin{pmatrix} E \\ p^3 \end{pmatrix} = \left[1 + \eta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right] \begin{pmatrix} m \\ 0 \end{pmatrix},$$

where η is some infinitesimal parameter. For finite η we must write

$$\begin{aligned}\begin{pmatrix} E \\ p^3 \end{pmatrix} &= \exp\left[\eta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right] \begin{pmatrix} m \\ 0 \end{pmatrix} \\ &= \left[\cosh\eta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sinh\eta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right] \begin{pmatrix} m \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} m\cosh\eta \\ m\sinh\eta \end{pmatrix}. \end{aligned} \tag{3.48}$$

The parameter η is called the *rapidity*. It is the quantity that is additive under successive boosts.

Now apply the same boost to $u(p)$. According to Eqs. (3.26) and (3.30),

$$\begin{aligned} u(p) &= \exp\left[-\tfrac{1}{2}\eta \begin{pmatrix} \sigma^3 & 0 \\ 0 & -\sigma^3 \end{pmatrix}\right] \sqrt{m} \begin{pmatrix} \xi \\ \xi \end{pmatrix} \\ &= \left[\cosh(\tfrac{1}{2}\eta) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \sinh(\tfrac{1}{2}\eta) \begin{pmatrix} \sigma^3 & 0 \\ 0 & -\sigma^3 \end{pmatrix}\right] \sqrt{m} \begin{pmatrix} \xi \\ \xi \end{pmatrix} \\ &= \begin{pmatrix} e^{\eta/2}\left(\frac{1-\sigma^3}{2}\right) + e^{-\eta/2}\left(\frac{1+\sigma^3}{2}\right) & 0 \\ 0 & e^{\eta/2}\left(\frac{1+\sigma^3}{2}\right) + e^{-\eta/2}\left(\frac{1-\sigma^3}{2}\right) \end{pmatrix} \sqrt{m} \begin{pmatrix} \xi \\ \xi \end{pmatrix} \\ &= \begin{pmatrix} \left[\sqrt{E+p^3}\left(\frac{1-\sigma^3}{2}\right) + \sqrt{E-p^3}\left(\frac{1+\sigma^3}{2}\right)\right]\xi \\ \left[\sqrt{E+p^3}\left(\frac{1+\sigma^3}{2}\right) + \sqrt{E-p^3}\left(\frac{1-\sigma^3}{2}\right)\right]\xi \end{pmatrix}. \end{aligned} \tag{3.49}$$

The last line can be simplified to give

$$u(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi \\ \sqrt{p\cdot\bar{\sigma}}\,\xi \end{pmatrix}, \tag{3.50}$$

where it is understood that in taking the square root of a matrix, we take the positive root of each eigenvalue. This expression for $u(p)$ is not only more compact, but is also valid for an arbitrary direction of $\mathbf{p}$. When working with expressions of this form, it is often useful to know the identity

$$(p\cdot\sigma)(p\cdot\bar{\sigma}) = p^2 = m^2. \tag{3.51}$$

You can then verify directly that (3.50) is a solution of the Dirac equation in the form of (3.43).

In practice it is often convenient to work with specific spinors ξ. A useful choice here would be eigenstates of σ^3. For example, if $\xi = \binom{1}{0}$ (spin up along the 3-axis), we get

$$u(p) = \begin{pmatrix} \sqrt{E-p^3}\binom{1}{0} \\ \sqrt{E+p^3}\binom{1}{0} \end{pmatrix} \underset{\text{large boost}}{\longrightarrow} \sqrt{2E} \begin{pmatrix} 0 \\ \binom{1}{0} \end{pmatrix}, \tag{3.52}$$

while for $\xi = \binom{0}{1}$ (spin down along the 3-axis) we have

$$u(p) = \begin{pmatrix} \sqrt{E+p^3}\binom{0}{1} \\ \sqrt{E-p^3}\binom{0}{1} \end{pmatrix} \underset{\text{large boost}}{\longrightarrow} \sqrt{2E}\begin{pmatrix} \binom{0}{1} \\ 0 \end{pmatrix}. \tag{3.53}$$

In the limit $\eta \to \infty$ the states degenerate into the two-component spinors of a massless particle. (We now see the reason for the factor of $\sqrt{m}$ in (3.47): It keeps the spinor expressions finite in the massless limit.)

The solutions (3.52) and (3.53) are eigenstates of the *helicity* operator,

$$h \equiv \hat{p} \cdot \mathbf{S} = \frac{1}{2}\hat{p}_i \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}. \tag{3.54}$$

A particle with $h = +1/2$ is called *right-handed*, while one with $h = -1/2$ is called *left-handed*. The helicity of a massive particle depends on the frame of reference, since one can always boost to a frame in which its momentum is in the opposite direction (but its spin is unchanged). For a massless particle, which travels at the speed of light, one cannot perform such a boost.

The extremely simple form of $u(p)$ for a massless particle in a helicity eigenstate makes the behavior of such a particle easy to understand. In Chapter 1, it enabled us to guess the form of the $e^+e^- \to \mu^+\mu^-$ cross section in the massless limit. In subsequent chapters we will often do a mindless calculation first, then look at helicity eigenstates in the high-energy limit to understand what we have done.

Incidentally, we are now ready to understand the origin of the notation ψ_L and ψ_R for Weyl spinors. The solutions of the Weyl equations are states of definite helicity, corresponding to left- and right-handed particles, respectively. The Lorentz invariance of helicity (for a massless particle) is manifest in the notation of Weyl spinors, since ψ_L and ψ_R live in different representations of the Lorentz group.

It is convenient to write the normalization condition for $u(p)$ in a Lorentz-invariant way. We saw above that $\psi^\dagger\psi$ is not Lorentz invariant. Similarly,

$$\begin{aligned} u^\dagger u &= (\xi^\dagger\sqrt{p\cdot\sigma},\, \xi^\dagger\sqrt{p\cdot\bar{\sigma}}) \cdot \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi \\ \sqrt{p\cdot\bar{\sigma}}\,\xi \end{pmatrix} \\ &= 2E_{\mathbf{p}}\xi^\dagger\xi. \end{aligned} \tag{3.55}$$

To make a Lorentz scalar we define

$$\bar{u}(p) = u^\dagger(p)\gamma^0. \tag{3.56}$$

Then by an almost identical calculation,

$$\bar{u}u = 2m\xi^\dagger\xi. \tag{3.57}$$

This will be our normalization condition, once we also require that the two-component spinor ξ be normalized as usual: $\xi^\dagger\xi = 1$. It is also conventional to choose basis spinors ξ^1 and ξ^2 (such as $\binom{1}{0}$ and $\binom{0}{1}$) that are orthogonal. For

a massless particle Eq. (3.57) is trivial, so we must write the normalization condition in the form of (3.55).

Let us summarize our discussion so far. The general solution of the Dirac equation can be written as a linear combination of plane waves. The positive-frequency waves are of the form

$$\psi(x) = u(p)e^{-ip\cdot x}, \qquad p^2 = m^2, \qquad p^0 > 0. \tag{3.58}$$

There are two linearly independent solutions for $u(p)$,

$$u^s(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix}, \qquad s = 1,2 \tag{3.59}$$

which we normalize according to

$$\bar{u}^r(p)u^s(p) = 2m\delta^{rs} \qquad \text{or} \qquad u^{r\dagger}(p)u^s(p) = 2E_{\mathbf{p}}\delta^{rs}. \tag{3.60}$$

In exactly the same way, we can find the negative-frequency solutions:

$$\psi(x) = v(p)e^{+ip\cdot x}, \qquad p^2 = m^2, \qquad p^0 > 0. \tag{3.61}$$

(Note that we have chosen to put the + sign into the exponential, rather than having $p^0 < 0$.) There are two linearly independent solutions for $v(p)$,

$$v^s(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\eta^s \\ -\sqrt{p\cdot\bar{\sigma}}\,\eta^s \end{pmatrix}, \qquad s = 1,2 \tag{3.62}$$

where η^s is another basis of two-component spinors. These solutions are normalized according to

$$\bar{v}^r(p)v^s(p) = -2m\delta^{rs} \qquad \text{or} \qquad v^{r\dagger}(p)v^s(p) = +2E_{\mathbf{p}}\delta^{rs}. \tag{3.63}$$

The u's and v's are also orthogonal to each other:

$$\bar{u}^r(p)v^s(p) = \bar{v}^r(p)u^s(p) = 0. \tag{3.64}$$

Be careful, since $u^{r\dagger}(p)v^s(p) \neq 0$ and $v^{r\dagger}(p)u^s(p) \neq 0$. However, note that

$$u^{r\dagger}(\mathbf{p})v^s(-\mathbf{p}) = v^{r\dagger}(-\mathbf{p})u^s(\mathbf{p}) = 0, \tag{3.65}$$

where we have changed the sign of the 3-momentum in one factor of each spinor product.

Spin Sums

In evaluating Feynman diagrams, we will often wish to sum over the polarization states of a fermion. We can derive the relevant completeness relations with a simple calculation:

$$\begin{aligned} \sum_{s=1,2} u^s(p)\bar{u}^s(p) &= \sum_s \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \left(\xi^{s\dagger}\sqrt{p\cdot\bar{\sigma}},\ \xi^{s\dagger}\sqrt{p\cdot\sigma}\right) \\ &= \begin{pmatrix} \sqrt{p\cdot\sigma}\sqrt{p\cdot\bar{\sigma}} & \sqrt{p\cdot\sigma}\sqrt{p\cdot\sigma} \\ \sqrt{p\cdot\bar{\sigma}}\sqrt{p\cdot\bar{\sigma}} & \sqrt{p\cdot\bar{\sigma}}\sqrt{p\cdot\sigma} \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} m & p\cdot\sigma \\ p\cdot\bar{\sigma} & m \end{pmatrix}.$$

In the second line we have used

$$\sum_{s=1,2} \xi^s \xi^{s\dagger} = \mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus we arrive at the desired formula,

$$\sum_s u^s(p)\bar{u}^s(p) = \gamma\cdot p + m. \tag{3.66}$$

Similarly,

$$\sum_s v^s(p)\bar{v}^s(p) = \gamma\cdot p - m. \tag{3.67}$$

The combination $\gamma\cdot p$ occurs so often that Feynman introduced the notation $\not{p} \equiv \gamma^\mu p_\mu$. We will use this notation frequently from now on.

3.4 Dirac Matrices and Dirac Field Bilinears

We saw in Section 3.2 that the quantity $\bar{\psi}\psi$ is a Lorentz scalar. It is also easy to show that $\bar{\psi}\gamma^\mu\psi$ is a 4-vector—we used this fact in writing down the Dirac Lagrangian (3.34). Now let us ask a more general question: Consider the expression $\bar{\psi}\Gamma\psi$, where Γ is any 4×4 constant matrix. Can we decompose this expression into terms that have definite transformation properties under the Lorentz group? The answer is yes, if we write Γ in terms of the following basis of sixteen 4×4 matrices, defined as antisymmetric combinations of γ-matrices:

1	1 of these
γ^μ	4 of these
$\gamma^{\mu\nu} = \frac{1}{2}[\gamma^\mu,\gamma^\nu] \equiv \gamma^{[\mu}\gamma^{\nu]} \equiv -i\sigma^{\mu\nu}$	6 of these
$\gamma^{\mu\nu\rho} = \gamma^{[\mu}\gamma^\nu\gamma^{\rho]}$	4 of these
$\gamma^{\mu\nu\rho\sigma} = \gamma^{[\mu}\gamma^\nu\gamma^\rho\gamma^{\sigma]}$	1 of these
	16 total

The Lorentz-transformation properties of these matrices are easy to determine. For example,

$$\begin{aligned}\bar{\psi}\gamma^{\mu\nu}\psi &\to \left(\bar{\psi}\Lambda_{\frac{1}{2}}^{-1}\right)\left(\tfrac{1}{2}[\gamma^\mu,\gamma^\nu]\right)\left(\Lambda_{\frac{1}{2}}\psi\right) \\ &= \tfrac{1}{2}\bar{\psi}\left(\Lambda_{\frac{1}{2}}^{-1}\gamma^\mu\Lambda_{\frac{1}{2}}\Lambda_{\frac{1}{2}}^{-1}\gamma^\nu\Lambda_{\frac{1}{2}} - \Lambda_{\frac{1}{2}}^{-1}\gamma^\nu\Lambda_{\frac{1}{2}}\Lambda_{\frac{1}{2}}^{-1}\gamma^\mu\Lambda_{\frac{1}{2}}\right)\psi \\ &= \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta\bar{\psi}\gamma^{\alpha\beta}\psi.\end{aligned}$$

Each set of matrices transforms as an antisymmetric tensor of successively higher rank.

The last two sets of matrices can be simplified by introducing an additional gamma matrix,

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = -\frac{i}{4!}\epsilon^{\mu\nu\rho\sigma}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma. \tag{3.68}$$

Then $\gamma^{\mu\nu\rho\sigma} = -i\epsilon^{\mu\nu\rho\sigma}\gamma^5$ and $\gamma^{\mu\nu\rho} = +i\epsilon^{\mu\nu\rho\sigma}\gamma_\sigma\gamma^5$. The matrix γ^5 has the following properties, all of which can be verified using (3.68) and the anticommutation relations (3.22):

$$(\gamma^5)^\dagger = \gamma^5; \tag{3.69}$$

$$(\gamma^5)^2 = 1; \tag{3.70}$$

$$\{\gamma^5, \gamma^\mu\} = 0. \tag{3.71}$$

This last property implies that $[\gamma^5, S^{\mu\nu}] = 0$. Thus the Dirac representation must be reducible, since eigenvectors of γ^5 whose eigenvalues are different transform without mixing (this criterion for reducibility is known as Schur's lemma). In our basis,

$$\gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.72}$$

in block-diagonal form. So a Dirac spinor with only left- (right-) handed components is an eigenstate of γ^5 with eigenvalue -1 $(+1)$, and indeed these spinors do transform without mixing, as we saw explicitly in Section 3.2.

Let us now rewrite our table of 4×4 matrices, and introduce some standard terminology:

1	scalar	1
γ^μ	vector	4
$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$	tensor	6
$\gamma^\mu\gamma^5$	pseudo-vector	4
γ^5	pseudo-scalar	1
		16

The terms *pseudo-vector* and *pseudo-scalar* arise from the fact that these quantities transform as a vector and scalar, respectively, under continuous Lorentz transformations, but with an additional sign change under parity transformations (as we will discuss in Section 3.6).

From the vector and pseudo-vector matrices we can form two currents out of Dirac field bilinears:

$$j^\mu(x) = \bar{\psi}(x)\gamma^\mu\psi(x); \qquad j^{\mu 5}(x) = \bar{\psi}(x)\gamma^\mu\gamma^5\psi(x). \tag{3.73}$$

Let us compute the divergences of these currents, assuming that ψ satisfies

the Dirac equation:

$$\begin{aligned}\partial_\mu j^\mu &= (\partial_\mu \bar{\psi})\gamma^\mu \psi + \bar{\psi}\gamma^\mu \partial_\mu \psi \\ &= (im\bar{\psi})\psi + \bar{\psi}(-im\psi) \\ &= 0.\end{aligned} \tag{3.74}$$

Thus j^μ is always conserved if $\psi(x)$ satisfies the Dirac equation. When we couple the Dirac field to the electromagnetic field, j^μ will become the electric current density. Similarly, one can compute

$$\partial_\mu j^{\mu 5} = 2im\bar{\psi}\gamma^5\psi. \tag{3.75}$$

If $m = 0$, this current (often called the *axial vector current*) is also conserved. It is then useful to form the linear combinations

$$j_L^\mu = \bar{\psi}\gamma^\mu\Big(\frac{1-\gamma^5}{2}\Big)\psi, \qquad j_R^\mu = \bar{\psi}\gamma^\mu\Big(\frac{1+\gamma^5}{2}\Big)\psi. \tag{3.76}$$

When $m = 0$, these are the electric current densities of left-handed and right-handed particles, respectively, and are separately conserved.

The two currents $j^\mu(x)$ and $j^{\mu 5}(x)$ are the Noether currents corresponding to the two transformations

$$\psi(x) \to e^{i\alpha}\psi(x) \qquad \text{and} \qquad \psi(x) \to e^{i\alpha\gamma^5}\psi(x).$$

The first of these is a symmetry of the Dirac Lagrangian (3.34). The second, called a *chiral transformation*, is a symmetry of the derivative term in $\mathcal{L}$ but not the mass term; thus, Noether's theorem confirms that the axial vector current is conserved only if $m = 0$.

Products of Dirac bilinears obey interchange relations, known as *Fierz identities*. We will discuss only the simplest of these, which will be needed several times later in the book. This simplest identity is most easily written in terms of the two-component Weyl spinors introduced in Eq. (3.36).

The core of the relation is the identity for the 2×2 matrices σ^μ defined in Eq. (3.41):

$$(\sigma^\mu)_{\alpha\beta}(\sigma_\mu)_{\gamma\delta} = 2\epsilon_{\alpha\gamma}\epsilon_{\beta\delta}. \tag{3.77}$$

(Here α, β, etc. are spinor indices, and ϵ is the antisymmetric symbol.) One can understand this relation by noting that the indices α, γ transform in the Lorentz representation of ψ_L, while β, δ transform in the separate representation of ψ_R, and the whole quantity must be a Lorentz invariant. Alternatively, one can just verify the 16 components of (3.77) explicitly.

By sandwiching identity (3.77) between the right-handed portions (i.e., lower half) of Dirac spinors u_1, u_2, u_3, u_4, we find the identity

$$\begin{aligned}(\bar{u}_{1R}\sigma^\mu u_{2R})(\bar{u}_{3R}\sigma_\mu u_{4R}) &= 2\epsilon_{\alpha\gamma}\bar{u}_{1R\alpha}\bar{u}_{3R\gamma}\epsilon_{\beta\delta}u_{2R\beta}u_{4R\delta} \\ &= -(\bar{u}_{1R}\sigma^\mu u_{4R})(\bar{u}_{3R}\sigma_\mu u_{2R}).\end{aligned} \tag{3.78}$$

This nontrivial relation says that the product of bilinears in (3.78) is antisymmetric under the interchange of the labels 2 and 4, and also under the

interchange of 1 and 3. Identity (3.77) also holds for $\bar{\sigma}^\mu$, and so we also find

$$(\bar{u}_{1L}\bar{\sigma}^\mu u_{2L})(\bar{u}_{3L}\bar{\sigma}_\mu u_{4L}) = -(\bar{u}_{1L}\bar{\sigma}^\mu u_{4L})(\bar{u}_{3L}\bar{\sigma}_\mu u_{2L}). \tag{3.79}$$

It is sometimes useful to combine the Fierz identity (3.78) with the identity linking σ^μ and $\bar{\sigma}^\mu$:

$$\epsilon_{\alpha\beta}(\sigma^\mu)_{\beta\gamma} = (\bar{\sigma}^{\mu T})_{\alpha\beta}\epsilon_{\beta\gamma}. \tag{3.80}$$

This relation is also straightforward to verify explicitly. By the use of (3.80), (3.79), and the relation

$$\bar{\sigma}^\mu\sigma_\mu = 4, \tag{3.81}$$

we can, for example, simplify horrible products of bilinears such as

$$\begin{aligned}(\bar{u}_{1L}\bar{\sigma}^\mu\sigma^\nu\bar{\sigma}^\lambda u_{2L})(\bar{u}_{3L}\bar{\sigma}_\mu\sigma_\nu\bar{\sigma}_\lambda u_{4L}) &= 2\epsilon_{\alpha\gamma}\bar{u}_{1L\alpha}\bar{u}_{3L\gamma}\epsilon_{\beta\delta}(\sigma^\nu\bar{\sigma}^\lambda u_{2L})_\beta(\sigma_\nu\bar{\sigma}_\lambda u_{4L})_\delta \\ &= 2\epsilon_{\alpha\gamma}\bar{u}_{1L\alpha}\bar{u}_{3L\gamma}\epsilon_{\beta\delta}u_{2L\beta}(\sigma^\lambda\bar{\sigma}^\nu\sigma_\nu\bar{\sigma}_\lambda u_{4L})_\delta \\ &= 2\cdot(4)^2\cdot\epsilon_{\alpha\gamma}\bar{u}_{1L\alpha}\bar{u}_{3L\gamma}\epsilon_{\beta\delta}u_{2L\beta}u_{4L\delta} \\ &= 16(\bar{u}_{1L}\bar{\sigma}^\mu u_{2L})(\bar{u}_{3L}\bar{\sigma}_\mu u_{4L}).\end{aligned} \tag{3.82}$$

There are also Fierz rearrangement identities for 4-component Dirac spinors and 4×4 Dirac matrices. To derive these, however, it is useful to take a more systematic approach. Problem 3.6 presents a general method and gives some examples of its application.

3.5 Quantization of the Dirac Field

We are now ready to construct the quantum theory of the free Dirac field. From the Lagrangian

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - m)\psi = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi, \tag{3.83}$$

we see that the canonical momentum conjugate to ψ is $i\psi^\dagger$, and thus the Hamiltonian is

$$H = \int d^3x\,\bar{\psi}(-i\boldsymbol{\gamma}\cdot\boldsymbol{\nabla} + m)\psi = \int d^3x\,\psi^\dagger\left[-i\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\nabla} + m\gamma^0\right]\psi. \tag{3.84}$$

If we define $\boldsymbol{\alpha} = \gamma^0\boldsymbol{\gamma}$, $\beta = \gamma^0$, you may recognize the quantity in brackets as the Dirac Hamiltonian of one-particle quantum mechanics:

$$h_D = -i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + m\beta. \tag{3.85}$$

How Not to Quantize the Dirac Field: A Lesson in Spin and Statistics

To quantize the Dirac field in analogy with the Klein-Gordon field we would impose the canonical commutation relations

$$\left[\psi_a(\mathbf{x}), \psi_b^\dagger(\mathbf{y})\right] = \delta^{(3)}(\mathbf{x}-\mathbf{y})\delta_{ab}, \qquad \text{(equal times)} \tag{3.86}$$

where a and b denote the spinor components of ψ. This already looks peculiar: If $\psi(x)$ were real-valued, the left-hand side would be antisymmetric under $\mathbf{x} \leftrightarrow \mathbf{y}$, while the right-hand side is symmetric. But ψ is complex, so we do not have a contradiction yet. In fact, we will soon find that much worse problems arise when we impose commutation relations on the Dirac field. But it is instructive to see how far we can get, in order to better understand the relation between spin and statistics. So let us press on; just remember that the next few pages will eventually turn out to be a blind alley.

Our first task is to find a representation of the commutation relations in terms of creation and annihilation operators that diagonalizes H. From the form of the Hamiltonian (3.84), it will clearly be helpful to expand $\psi(x)$ in a basis of eigenfunctions of h_D. We know these eigenfunctions already from our calculations in Section 3.3. There we found that

$$\left[i\gamma^0\partial_0 + i\boldsymbol{\gamma}\cdot\boldsymbol{\nabla} - m\right]u^s(p)e^{-ip\cdot x} = 0,$$

so $u^s(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}$ are eigenfunctions of h_D with eigenvalues $E_\mathbf{p}$. Similarly, the functions $v^s(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}}$ (or equivalently, $v^s(-\mathbf{p})e^{+i\mathbf{p}\cdot\mathbf{x}}$) are eigenfunctions of h_D with eigenvalues $-E_\mathbf{p}$. These form a complete set of eigenfunctions, since for any $\mathbf{p}$ there are two u's and two v's, giving us four eigenvectors of the 4×4 matrix h_D.

Expanding ψ in this basis, we obtain

$$\psi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_\mathbf{p}}}e^{i\mathbf{p}\cdot\mathbf{x}}\sum_{s=1,2}\Big(a^s_\mathbf{p}u^s(\mathbf{p}) + b^{s\dagger}_{-\mathbf{p}}v^s(-\mathbf{p})\Big), \tag{3.87}$$

where $a^s_\mathbf{p}$ and $b^s_\mathbf{p}$ are operator coefficients. (For now we work in the Schrödinger picture, where ψ does not depend on time.) Postulate the commutation relations

$$\left[a^r_\mathbf{p}, a^{s\dagger}_\mathbf{q}\right] = \left[b^r_\mathbf{p}, b^{s\dagger}_\mathbf{q}\right] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})\delta^{rs}. \tag{3.88}$$

It is then easy to verify the commutation relations (3.86) for ψ and $\psi^\dagger$:

$$\begin{aligned}\left[\psi(\mathbf{x}),\psi^\dagger(\mathbf{y})\right] &= \int \frac{d^3p\,d^3q}{(2\pi)^6}\frac{1}{\sqrt{2E_\mathbf{p}\,2E_\mathbf{q}}}e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})}\\ &\quad\times\sum_{r,s}\Big(\left[a^r_\mathbf{p}, a^{s\dagger}_\mathbf{q}\right]u^r(\mathbf{p})\bar{u}^s(\mathbf{q}) + \left[b^r_{-\mathbf{p}}, b^{s\dagger}_{-\mathbf{q}}\right]v^r(-\mathbf{p})\bar{v}^s(-\mathbf{q})\Big)\gamma^0\\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_\mathbf{p}}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\\ &\quad\times\Big[\left(\gamma^0E_\mathbf{p} - \boldsymbol{\gamma}\cdot\mathbf{p} + m\right) + \left(\gamma^0E_\mathbf{p} + \boldsymbol{\gamma}\cdot\mathbf{p} - m\right)\Big]\gamma^0\\ &= \delta^{(3)}(\mathbf{x}-\mathbf{y})\times\mathbf{1}_{4\times4}.\end{aligned} \tag{3.89}$$

In the second step we have used the spin sum completeness relations (3.66) and (3.67).

We are now ready to write H in terms of the a's and b's. After another short calculation (making use of the orthogonality relations (3.60), (3.63), and (3.65)), we find

$$H = \int \frac{d^3p}{(2\pi)^3} \sum_s \Big(E_{\mathbf{p}} a^{s\dagger}_{\mathbf{p}} a^s_{\mathbf{p}} - E_{\mathbf{p}} b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} \Big). \tag{3.90}$$

Something is terribly wrong with the second term: By creating more and more particles with $b^\dagger$, we can lower the energy indefinitely. (It would not have helped to rename $b \leftrightarrow b^\dagger$, since doing so would ruin the commutation relation (3.89).)

We seem to be in rather deep trouble, but again let's press on, and investigate the causality of this theory. To do this we should compute $[\psi(x), \psi^\dagger(y)]$ (or more conveniently, $[\psi(x), \bar{\psi}(y)]$) at non-equal times and hope to get zero outside the light-cone. First we must switch to the Heisenberg picture and restore the time-dependence of ψ and $\bar{\psi}$. Using the relations

$$e^{iHt} a^s_{\mathbf{p}} e^{-iHt} = a^s_{\mathbf{p}} e^{-iE_{\mathbf{p}} t}, \qquad e^{iHt} b^s_{\mathbf{p}} e^{-iHt} = b^s_{\mathbf{p}} e^{+iE_{\mathbf{p}} t}, \tag{3.91}$$

we immediately have

$$\begin{aligned} \psi(x) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_s \Big(a^s_{\mathbf{p}} u^s(p) e^{-ip\cdot x} + b^s_{\mathbf{p}} v^s(p) e^{ip\cdot x} \Big); \\ \bar{\psi}(x) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_s \Big(a^{s\dagger}_{\mathbf{p}} \bar{u}^s(p) e^{ip\cdot x} + b^{s\dagger}_{\mathbf{p}} \bar{v}^s(p) e^{-ip\cdot x} \Big). \end{aligned} \tag{3.92}$$

We can now calculate the general commutator:

$$\begin{aligned} \big[\psi_a(x), \bar{\psi}_b(y)\big] &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \sum_s \Big(u^s_a(p) \bar{u}^s_b(p) e^{-ip\cdot(x-y)} + v^s_a(p) \bar{v}^s_b(p) e^{ip\cdot(x-y)} \Big) \\ &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \Big((\not{p} + m)_{ab} e^{-ip\cdot(x-y)} + (\not{p} - m)_{ab} e^{ip\cdot(x-y)} \Big) \\ &= \big(i\not{\partial}_x + m\big)_{ab} \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \Big(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)} \Big) \\ &= \big(i\not{\partial}_x + m\big)_{ab} \big[\phi(x), \phi(y)\big]. \end{aligned}$$

Since $\big[\phi(x), \phi(y)\big]$ (the commutator of a real Klein-Gordon field) vanishes outside the light-cone, this quantity does also.

There is something odd, however, about this solution to the causality problem. Let $|0\rangle$ be the state that is annihilated by all the $a^s_{\mathbf{p}}$ and $b^s_{\mathbf{p}}$: $a^s_{\mathbf{p}} |0\rangle = b^s_{\mathbf{p}} |0\rangle = 0$. Then

$$\begin{aligned} \big[\psi_a(x), \bar{\psi}_b(y)\big] &= \langle 0 | \left[\psi_a(x), \bar{\psi}_b(y)\right] | 0 \rangle \\ &= \langle 0 | \psi_a(x) \bar{\psi}_b(y) | 0 \rangle - \langle 0 | \bar{\psi}_b(y) \psi_a(x) | 0 \rangle , \end{aligned}$$

just as for the Klein-Gordon field. But in the Klein-Gordon case, we got one term of the commutator from each of these two pieces: the propagation of a particle from y to x was canceled by the propagation of an antiparticle from x to y outside the light-cone. Here both terms come from the first piece, $\langle 0|\,\psi(x)\bar{\psi}(y)\,|0\rangle$, since the second piece is zero. The cancellation is between positive-energy particles and negative-energy particles, both propagating from y to x.

This observation can actually lead us to a resolution of the negative-energy problem. One of the assumptions we made in quantizing the Dirac theory must have been incorrect. Let us therefore forget about the postulated commutation relations (3.86) and (3.88), and see whether we can find a way for positive-energy particles to propagate in both directions. We will also have to drop our definition of the vacuum $|0\rangle$ as the state that is annihilated by all $a^s_{\mathbf{p}}$ and $b^s_{\mathbf{p}}$. We will, however, retain the expressions (3.92) for $\psi(x)$ and $\bar{\psi}(x)$ as Heisenberg operators, since if $\psi(x)$ and $\bar{\psi}(x)$ solve the Dirac equation, they must be decomposable into such plane-wave solutions.

First consider the propagation amplitude $\langle 0|\,\psi(x)\bar{\psi}(y)\,|0\rangle$, which is to represent a positive-energy particle propagating from y to x. In this case we want the (Heisenberg) state $\bar{\psi}(y)\,|0\rangle$ to be made up of only positive-energy, or negative-frequency components (since a Heisenberg state $\Psi_H = e^{+iHt}\Psi_S$). Thus only the $a^{s\dagger}_{\mathbf{p}}$ term of $\bar{\psi}(y)$ can contribute, which means that $b^{s\dagger}_{\mathbf{p}}$ must annihilate the vacuum. Similarly $\langle 0|\,\psi(x)$ can contain only positive-frequency components. Thus we have

$$\begin{aligned}\langle 0|\,\psi(x)\bar{\psi}(y)\,|0\rangle = \langle 0|\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\sum_r a^r_{\mathbf{p}}u^r(p)e^{-ipx} \\ \times \int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{q}}}}\sum_s a^{s\dagger}_{\mathbf{q}}\bar{u}^s(q)e^{iqy}\,|0\rangle\,.\end{aligned} \tag{3.93}$$

We can say something about the matrix element $\langle 0|\,a^r_{\mathbf{p}}a^{s\dagger}_{\mathbf{q}}\,|0\rangle$ even without knowing how to interchange $a^r_{\mathbf{p}}$ and $a^{s\dagger}_{\mathbf{q}}$, by using translational and rotational invariance. If the ground state $|0\rangle$ is to be invariant under translations, we must have $|0\rangle = e^{i\mathbf{P}\cdot\mathbf{x}}\,|0\rangle$. Furthermore, since $a^{s\dagger}_{\mathbf{q}}$ creates momentum $\mathbf{q}$, we can use Eq. (2.48) to compute

$$\begin{aligned}\langle 0|\,a^r_{\mathbf{p}}a^{s\dagger}_{\mathbf{q}}\,|0\rangle &= \langle 0|\,a^r_{\mathbf{p}}a^{s\dagger}_{\mathbf{q}}e^{i\mathbf{P}\cdot\mathbf{x}}\,|0\rangle \\ &= e^{i(\mathbf{p}-\mathbf{q})\cdot\mathbf{x}}\,\langle 0|\,e^{i\mathbf{P}\cdot\mathbf{x}}a^r_{\mathbf{p}}a^{s\dagger}_{\mathbf{q}}\,|0\rangle \\ &= e^{i(\mathbf{p}-\mathbf{q})\cdot\mathbf{x}}\,\langle 0|\,a^r_{\mathbf{p}}a^{s\dagger}_{\mathbf{q}}\,|0\rangle\,.\end{aligned}$$

This says that if $\langle 0|\,a^r_{\mathbf{p}}a^{s\dagger}_{\mathbf{q}}\,|0\rangle$ is to be nonzero, $\mathbf{p}$ must equal $\mathbf{q}$. Similarly, it can be shown that rotational invariance of $|0\rangle$ implies $r = s$. (This should be intuitively clear, and can be checked after we discuss the angular momentum operator later in this section.) From these considerations we conclude that

the matrix element can be written

$$\langle 0|\, a^r_{\mathbf{p}} a^{s\dagger}_{\mathbf{q}}\, |0\rangle = (2\pi)^3 \delta^{(3)}(\mathbf{p}-\mathbf{q})\delta^{rs} \cdot A(\mathbf{p}),$$

where $A(\mathbf{p})$ is so far undetermined. Note, however, that if the norm of a state is always positive (as it should be in any self-respecting Hilbert space), $A(\mathbf{p})$ must be greater than zero. We can now go back to (3.93), and write

$$\begin{aligned}\langle 0|\, \psi(x)\bar{\psi}(y)\, |0\rangle &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} \sum_s u^s(p)\bar{u}^s(p) A(\mathbf{p}) e^{-ip(x-y)} \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} (\not{p} + m) A(\mathbf{p}) e^{-ip(x-y)}.\end{aligned}$$

This expression is properly invariant under boosts only if $A(\mathbf{p})$ is a Lorentz scalar, i.e., $A(\mathbf{p})$=$A(p^2)$. Since $p^2 = m^2$, A must be a constant. So finally we obtain

$$\langle 0|\, \psi_a(x)\bar{\psi}_b(y)\, |0\rangle = \left(i\not{\partial}_x + m\right)_{ab} \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} e^{-ip(x-y)} \cdot A. \tag{3.94}$$

Similarly, in the amplitude $\langle 0|\, \bar{\psi}(y)\psi(x)\, |0\rangle$, we want the only contributions to be from the positive-frequency terms of $\bar{\psi}(y)$ and the negative-frequency terms of $\psi(x)$. So $a^s_{\mathbf{p}}$ still annihilates the vacuum, but $b^s_{\mathbf{p}}$ does not. Then by arguments identical to those given above, we have

$$\langle 0|\, \bar{\psi}_b(y)\psi_a(x)\, |0\rangle = -\left(i\not{\partial}_x + m\right)_{ab} \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} e^{ip(x-y)} \cdot B, \tag{3.95}$$

where B is another positive constant. The minus sign is important; it comes from the completeness relation (3.67) for $\sum v\bar{v}$ and the sign of x in the exponential factor. It implies that we cannot have $\langle 0|\, [\psi(x), \bar{\psi}(y)]\, |0\rangle = 0$ outside the light-cone: The two terms (3.94) and (3.95) would indeed cancel if $A = -B$, but this is impossible since A and B must both be positive.

The solution, however, is now at hand. By setting $A = B = 1$, it is easy to obtain (outside the light-cone)

$$\langle 0|\, \psi_a(x)\bar{\psi}_b(y)\, |0\rangle = -\langle 0|\, \bar{\psi}_b(y)\psi_a(x)\, |0\rangle\,.$$

That is, the spinor fields *anti*commute at spacelike separation. This is enough to preserve causality, since all reasonable observables (such as energy, charge, and particle number) are built out of an *even* number of spinor fields; for any such observables $\mathcal{O}_1$ and $\mathcal{O}_2$, we still have $[\mathcal{O}_1(x), \mathcal{O}_2(y)] = 0$ for $(x-y)^2 < 0$.

And remarkably, postulating *anti*commutation relations for the Dirac field solves the negative energy problem. The equal-time anticommutation relations will be

$$\begin{aligned}\left\{\psi_a(\mathbf{x}), \psi^\dagger_b(\mathbf{y})\right\} &= \delta^{(3)}(\mathbf{x}-\mathbf{y})\delta_{ab}; \\ \left\{\psi_a(\mathbf{x}), \psi_b(\mathbf{y})\right\} &= \left\{\psi^\dagger_a(\mathbf{x}), \psi^\dagger_b(\mathbf{y})\right\} = 0.\end{aligned} \tag{3.96}$$

We can expand $\psi(\mathbf{x})$ in terms of $a^s_{\mathbf{p}}$ and $b^s_{\mathbf{p}}$ as before (Eq. (3.87)). The creation and annihilation operators must now obey

$$\{a^r_{\mathbf{p}}, a^{s\dagger}_{\mathbf{q}}\} = \{b^r_{\mathbf{p}}, b^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})\delta^{rs} \tag{3.97}$$

(with all other anticommutators equal to zero) in order that (3.96) be satisfied. Another computation gives the Hamiltonian,

$$H = \int \frac{d^3p}{(2\pi)^3} \sum_s \Big(E_{\mathbf{p}} a^{s\dagger}_{\mathbf{p}} a^s_{\mathbf{p}} - E_{\mathbf{p}} b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}}\Big),$$

which is the same as before; $b^{s\dagger}_{\mathbf{p}}$ still creates negative energy. However, the relation $\{b^r_{\mathbf{p}}, b^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})\delta^{rs}$ is symmetric between $b^r_{\mathbf{p}}$ and $b^{s\dagger}_{\mathbf{q}}$. So let us simply redefine

$$\tilde{b}^s_{\mathbf{p}} \equiv b^{s\dagger}_{\mathbf{p}}; \qquad \tilde{b}^{s\dagger}_{\mathbf{p}} \equiv b^s_{\mathbf{p}}. \tag{3.98}$$

These of course obey exactly the same anticommutation relations, but now the second term in the Hamiltonian is

$$-E_{\mathbf{p}} b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} = +E_{\mathbf{p}} \tilde{b}^{s\dagger}_{\mathbf{p}} \tilde{b}^s_{\mathbf{p}} - (\text{const}).$$

If we choose $|0\rangle$ to be the state that is annihilated by $a^s_{\mathbf{p}}$ and $\tilde{b}^s_{\mathbf{p}}$, then all excitations of $|0\rangle$ have positive energy.

What happened? To better understand this trick, let us abandon the field theory for a moment and consider a theory with a single pair of b and $b^\dagger$ operators obeying $\{b, b^\dagger\} = 1$ and $\{b, b\} = \{b^\dagger, b^\dagger\} = 0$. Choose a state $|0\rangle$ such that $b|0\rangle = 0$. Then $b^\dagger|0\rangle$ is a new state; call it $|1\rangle$. This state satisfies $b|1\rangle = |0\rangle$ and $b^\dagger|1\rangle = 0$. So b and $b^\dagger$ act on a Hilbert space of only two states, $|0\rangle$ and $|1\rangle$. We might say that $|0\rangle$ represents an "empty" state, and that $b^\dagger$ "fills" the state. But we could equally well call $|1\rangle$ the empty state and say that $b = \tilde{b}^\dagger$ fills it. The two descriptions are completely equivalent, until we specify some observable that allows us to distinguish the states physically. In our case the correct choice is to take the state of lower energy to be the empty one. And it is less confusing to put the dagger on the operator that creates positive energy. That is exactly what we have done.

Note, by the way, that since $(\tilde{b}^\dagger)^2 = 0$, the state cannot be filled twice. More generally, the anticommutation relations imply that any multiparticle state is antisymmetric under the interchange of two particles: $a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{q}}|0\rangle = -a^\dagger_{\mathbf{q}} a^\dagger_{\mathbf{p}}|0\rangle$. Thus we conclude that if the ladder operators obey *anticommutation* relations, the corresponding particles obey *Fermi-Dirac* statistics.

We have just shown that in order to insure that the vacuum has only positive-energy excitations, we must quantize the Dirac field with anticommutation relations; under these conditions the particles associated with the Dirac field obey Fermi-Dirac statistics. This conclusion is part of a more gen-

eral result, first derived by Pauli*: Lorentz invariance, positive energies, positive norms, and causality together imply that particles of integer spin obey Bose-Einstein statistics, while particles of half-odd-integer spin obey Fermi-Dirac statistics.

The Quantized Dirac Field

Let us now summarize the results of the quantized Dirac theory in a systematic way. Since the dust has settled, we should clean up our notation: From now on we will write $\tilde{b}_{\mathbf{p}}$ (the operator that lowers the energy of a state) simply as $b_{\mathbf{p}}$, and $\tilde{b}^{\dagger}_{\mathbf{p}}$ as $b^{\dagger}_{\mathbf{p}}$. All the expressions we will need in our later work are listed below; corresponding expressions above, where they differ, should be forgotten.

First we write the field operators:

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_s \left(a^s_{\mathbf{p}} u^s(p) e^{-ip\cdot x} + b^{s\dagger}_{\mathbf{p}} v^s(p) e^{ip\cdot x} \right); \tag{3.99}$$

$$\bar{\psi}(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_s \left(b^s_{\mathbf{p}} \bar{v}^s(p) e^{-ip\cdot x} + a^{s\dagger}_{\mathbf{p}} \bar{u}^s(p) e^{ip\cdot x} \right). \tag{3.100}$$

The creation and annihilation operators obey the anticommutation rules

$$\{a^r_{\mathbf{p}}, a^{s\dagger}_{\mathbf{q}}\} = \{b^r_{\mathbf{p}}, b^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \delta^{rs}, \tag{3.101}$$

with all other anticommutators equal to zero. The equal-time anticommutation relations for ψ and $\psi^{\dagger}$ are then

$$\begin{aligned} \{\psi_a(\mathbf{x}), \psi^{\dagger}_b(\mathbf{y})\} &= \delta^{(3)}(\mathbf{x} - \mathbf{y}) \delta_{ab}; \\ \{\psi_a(\mathbf{x}), \psi_b(\mathbf{y})\} &= \{\psi^{\dagger}_a(\mathbf{x}), \psi^{\dagger}_b(\mathbf{y})\} = 0. \end{aligned} \tag{3.102}$$

The vacuum $|0\rangle$ is defined to be the state such that

$$a^s_{\mathbf{p}} |0\rangle = b^s_{\mathbf{p}} |0\rangle = 0. \tag{3.103}$$

The Hamiltonian can be written

$$H = \int \frac{d^3p}{(2\pi)^3} \sum_s E_{\mathbf{p}} \left(a^{s\dagger}_{\mathbf{p}} a^s_{\mathbf{p}} + b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} \right), \tag{3.104}$$

where we have dropped the infinite constant term that comes from anticommuting $b^s_{\mathbf{p}}$ and $b^{s\dagger}_{\mathbf{p}}$. From this we see that the vacuum is the state of lowest energy, as desired. The momentum operator is

$$\mathbf{P} = \int d^3x \, \psi^{\dagger}(-i\boldsymbol{\nabla})\psi = \int \frac{d^3p}{(2\pi)^3} \sum_s \mathbf{p} \left(a^{s\dagger}_{\mathbf{p}} a^s_{\mathbf{p}} + b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} \right). \tag{3.105}$$

*W. Pauli, *Phys. Rev.* **58**, 716 (1940), reprinted in Schwinger (1958). A rigorous treatment is given by R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (Benjamin/Cummings, Reading, Mass., 1964).

Thus both $a_{\mathbf{p}}^{s\dagger}$ and $b_{\mathbf{p}}^{s\dagger}$ create particles with energy $+E_{\mathbf{p}}$ and momentum $\mathbf{p}$. We will refer to the particles created by $a_{\mathbf{p}}^{s\dagger}$ as *fermions* and to those created by $b_{\mathbf{p}}^{s\dagger}$ as *antifermions*.

The one-particle states

$$|\mathbf{p}, s\rangle \equiv \sqrt{2E_{\mathbf{p}}}\, a_{\mathbf{p}}^{s\dagger}\, |0\rangle \tag{3.106}$$

are defined so that their inner product

$$\langle \mathbf{p}, r|\mathbf{q}, s\rangle = 2E_{\mathbf{p}}(2\pi)^3 \delta^{(3)}(\mathbf{p}-\mathbf{q})\delta^{rs} \tag{3.107}$$

is Lorentz invariant. This implies that the operator $U(\Lambda)$ that implements Lorentz transformations on the states of the Hilbert space is unitary, even though for boosts, $\Lambda_{\frac{1}{2}}$ is not unitary.

It will be reassuring to do a consistency check, to see that $U(\Lambda)$ implements the right transformation on $\psi(x)$. So calculate

$$U\psi(x)U^{-1} = U \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_s \left(a_{\mathbf{p}}^s u^s(p) e^{-ipx} + b_{\mathbf{p}}^{s\dagger} v^s(p) e^{ipx} \right) U^{-1}. \tag{3.108}$$

We can concentrate on the first term; the second is completely analogous. Equation (3.106) implies that $a_{\mathbf{p}}^s$ transforms according to

$$U(\Lambda) a_{\mathbf{p}}^s U^{-1}(\Lambda) = \sqrt{\frac{E_{\Lambda\mathbf{p}}}{E_{\mathbf{p}}}}\, a_{\Lambda\mathbf{p}}^s, \tag{3.109}$$

assuming that the axis of spin quantization is parallel to the boost or rotation axis. To use this relation to evaluate (3.108), rewrite the integral as

$$\int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} a_{\mathbf{p}}^s = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \cdot \sqrt{2E_{\mathbf{p}}}\, a_{\mathbf{p}}^s.$$

The second factor is transformed in a simple way by U, and the first is a Lorentz-invariant integral. Thus, if we apply (3.109) and make the substitution $\tilde{p} = \Lambda p$, Eq. (3.108) becomes

$$U(\Lambda)\psi(x)U^{-1}(\Lambda) = \int \frac{d^3\tilde{p}}{(2\pi)^3} \frac{1}{2E_{\tilde{\mathbf{p}}}} \sum_s u^s(\Lambda^{-1}\tilde{p}) \sqrt{2E_{\tilde{\mathbf{p}}}}\, a_{\tilde{\mathbf{p}}}^s e^{-i\tilde{p}\cdot\Lambda x} + \cdots.$$

But $u^s(\Lambda^{-1}\tilde{p}) = \Lambda_{\frac{1}{2}}^{-1} u^s(\tilde{p})$, so indeed we have

$$\begin{aligned} U(\Lambda)\psi(x)U^{-1}(\Lambda) &= \int \frac{d^3\tilde{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\tilde{\mathbf{p}}}}} \sum_s \Lambda_{\frac{1}{2}}^{-1} u^s(\tilde{p}) a_{\tilde{\mathbf{p}}}^s e^{-i\tilde{p}\cdot\Lambda x} + \cdots \\ &= \Lambda_{\frac{1}{2}}^{-1} \psi(\Lambda x). \end{aligned} \tag{3.110}$$

This result says that the transformed field creates and destroys particles at the point Λx, as it must. Note, however, that this transformation appears to be in the wrong direction compared to Eq. (3.2), where the transformed

field ϕ was evaluated at $\Lambda^{-1}x$. The difference is that in Section 3.1 we imagined that we transformed a pre-existing field distribution that was measured by $\phi(x)$. Here, we are transforming the action of $\phi(x)$ in creating or destroying particles. These two ways of implementing the Lorentz transformation work in opposite directions. Notice, though, that the matrix acting on ψ and the transformation of the coordinate x have the correct relative orientation, consistent with Eq. (3.8).

Next we should discuss the spin of a Dirac particle. We expect Dirac fermions to have spin 1/2; now we can demonstrate this property from our formalism. We have already shown that the particles created by $a^{s\dagger}_{\mathbf{p}}$ and $b^{s\dagger}_{\mathbf{p}}$ each come in two "spin" states: $s = 1, 2$. But we haven't proved yet that this "spin" has anything to do with angular momentum. To do this, we must write down the angular momentum operator.

Recall that we found the linear momentum operator in Section 2.2 by looking for the conserved quantity associated with translational invariance. We can find the angular momentum operator in a similar way as a consequence of rotational invariance. Under a rotation (or any Lorentz transformation), the Dirac field ψ transforms (in our original convention) according to

$$\psi(x) \to \psi'(x) = \Lambda_{\frac{1}{2}}\psi(\Lambda^{-1}x).$$

To apply Noether's theorem we must compute the change in the field at a fixed point, that is,

$$\delta\psi = \psi'(x) - \psi(x) = \Lambda_{\frac{1}{2}}\psi(\Lambda^{-1}x) - \psi(x).$$

Consider for definiteness an infinitesimal rotation of coordinates by an angle θ about the z-axis. The parametrization of this transformation is given just below Eq. (3.19): $\omega_{12} = -\omega_{21} = \theta$. Using the same parameters in Eq. (3.30), we find

$$\Lambda_{\frac{1}{2}} \approx 1 - \tfrac{i}{2}\omega_{\mu\nu}S^{\mu\nu} = 1 - \tfrac{i}{2}\theta\Sigma^3.$$

We can now compute

$$\begin{aligned}\delta\psi(x) &= \big(1 - \tfrac{i}{2}\theta\Sigma^3\big)\psi(t, x + \theta y, y - \theta x, z) - \psi(x)\\ &= -\theta\big(x\partial_y - y\partial_x + \tfrac{i}{2}\Sigma^3\big)\psi(x) \equiv \theta\Delta\psi.\end{aligned}$$

The time-component of the conserved Noether current is then

$$j^0 = \frac{\partial\mathcal{L}}{\partial(\partial_0\psi)}\Delta\psi = -i\bar{\psi}\gamma^0\big(x\partial_y - y\partial_x + \tfrac{i}{2}\Sigma^3\big)\psi.$$

Similar expressions hold for rotations about the x- and y-axes, so the angular momentum operator is

$$\mathbf{J} = \int d^3x\, \psi^\dagger\Big(\mathbf{x} \times (-i\boldsymbol{\nabla}) + \tfrac{1}{2}\boldsymbol{\Sigma}\Big)\psi. \tag{3.111}$$

For nonrelativistic fermions, the first term of (3.111) gives the orbital angular momentum. The second term therefore gives the spin angular momentum.

Unfortunately, the division of (3.111) into spin and orbital parts is not so straightforward for relativistic fermions, so it is not simple to write a general expression for this quantity in terms of ladder operators.

To prove that a Dirac particle has spin 1/2, however, it suffices to consider particles at rest. We would like to apply J_z to the state $a_0^{s\dagger}|0\rangle$ and show that this state is an eigenvector. This is most easily done using a trick: Since J_z must annihilate the vacuum, $J_z a_0^{s\dagger}|0\rangle = [J_z, a_0^{s\dagger}]|0\rangle$. The commutator is nonzero only for the terms in J_z that have annihiliation operators at $\mathbf{p} = 0$. For these terms, the orbital part of (3.111) does not contribute. To write the spin term of (3.111) in terms of ladder operators, use expansions (3.99) and (3.100), evaluated at $t = 0$:

$$J_z = \int d^3x \int \frac{d^3p\, d^3p'}{(2\pi)^6} \frac{1}{\sqrt{2E_{\mathbf{p}}\, 2E_{\mathbf{p}'}}} e^{-i\mathbf{p}'\cdot\mathbf{x}} e^{i\mathbf{p}\cdot\mathbf{x}}$$
$$\times \sum_{r,r'} \Big(a_{\mathbf{p}'}^{r'\dagger} u^{r'\dagger}(\mathbf{p}') + b_{-\mathbf{p}'}^{r'} v^{r'\dagger}(-\mathbf{p}') \Big) \frac{\Sigma^3}{2} \Big(a_{\mathbf{p}}^{r} u^{r}(\mathbf{p}) + b_{-\mathbf{p}}^{r\dagger} v^{r}(-\mathbf{p}) \Big).$$

Taking the commutator with $a_0^{s\dagger}$, the only nonzero term has the structure $[a_{\mathbf{p}}^{r\dagger} a_{\mathbf{p}}^{r}, a_0^{s\dagger}] = (2\pi)^3 \delta^{(3)}(\mathbf{p}) a_0^{r\dagger} \delta^{rs}$; the other three terms in the commutator either vanish or annihilate the vacuum. Thus we find

$$J_z a_0^{s\dagger}|0\rangle = \frac{1}{2m} \sum_r \Big(u^{s\dagger}(0) \frac{\Sigma^3}{2} u^r(0) \Big) a_0^{r\dagger}|0\rangle = \sum_r \Big(\xi^{s\dagger} \frac{\sigma^3}{2} \xi^r \Big) a_0^{r\dagger}|0\rangle\,,$$

where we have used the explicit form (3.47) of $u(0)$ to obtain the last expression. The sum over r is accomplished most easily by choosing the spinors ξ^r to be eigenstates of σ^3. We then find that for $\xi^s = \binom{1}{0}$, the one-particle state is an eigenstate of J_z with eigenvalue $+1/2$, while for $\xi^s = \binom{0}{1}$, it is an eigenstate of J_z with eigenvalue $-1/2$. This result is exactly what we expect for electrons.

An analogous calculation determines the spin of a zero-momentum antifermion. But in this case, since the order of the b and $b^\dagger$ terms in J_z is reversed, we get an extra minus sign from evaluating $[b_{\mathbf{p}} b_{\mathbf{p}}^\dagger, b_0^\dagger] = -[b_{\mathbf{p}}^\dagger b_{\mathbf{p}}, b_0^\dagger]$. Thus for positrons, the association between the spinors η^s and the spin angular momentum is reversed: $\binom{1}{0}$ corresponds to spin $-1/2$, while $\binom{0}{1}$ corresponds to spin $+1/2$. This reversal of sign agrees with the prediction of Dirac hole theory. From that viewpoint, a positron is the absence of a negative-energy electron. If the missing electron had positive J_z, its absence has negative J_z.

In summary, the angular momentum of zero-momentum fermions is given by

$$J_z a_0^{s\dagger}|0\rangle = \pm\tfrac{1}{2} a_0^{s\dagger}|0\rangle\,, \qquad J_z b_0^{s\dagger}|0\rangle = \mp\tfrac{1}{2} b_0^{s\dagger}|0\rangle\,, \tag{3.112}$$

where the upper sign is for $\xi^s = \binom{1}{0}$ and the lower sign is for $\xi^s = \binom{0}{1}$.

There is one more important conserved quantity in the Dirac theory. In Section 3.4 we saw that the current $j^\mu = \bar{\psi}\gamma^\mu\psi$ is conserved. The charge associated with this current is

$$Q = \int d^3x\, \psi^\dagger(x)\psi(x) = \int \frac{d^3p}{(2\pi)^3} \sum_s \left(a^{s\dagger}_{\mathbf{p}} a^s_{\mathbf{p}} + b^s_{-\mathbf{p}} b^{s\dagger}_{-\mathbf{p}} \right),$$

or, if we ignore another infinite constant,

$$Q = \int \frac{d^3p}{(2\pi)^3} \sum_s \left(a^{s\dagger}_{\mathbf{p}} a^s_{\mathbf{p}} - b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} \right). \tag{3.113}$$

So $a^{s\dagger}_{\mathbf{p}}$ creates fermions with charge $+1$, while $b^{s\dagger}_{\mathbf{p}}$ creates antifermions with charge -1. When we couple the Dirac field to the electromagnetic field, we will see that Q is none other than the electric charge (up to a constant factor that depends on which type of particle we wish to describe; e.g., for electrons, the electric charge is Qe).

In Quantum Electrodynamics we will use the spinor field ψ to describe electrons and positrons. The particles created by $a^{s\dagger}_{\mathbf{p}}$ are electrons; they have energy $E_{\mathbf{p}}$, momentum $\mathbf{p}$, spin $1/2$ with polarization appropriate to ξ^s, and charge $+1$ (in units of e). The particles created by $b^{s\dagger}_{\mathbf{p}}$ are positrons; they have energy $E_{\mathbf{p}}$, momentum $\mathbf{p}$, spin $1/2$ with polarization opposite to that of ξ^s, and charge -1. The state $\psi_\alpha(x)\,|0\rangle$ contains a positron at position x, whose polarization corresponds to the spinor component chosen. Similarly, $\bar{\psi}_\alpha(x)\,|0\rangle$ is a state of one electron at position x.

The Dirac Propagator

Calculating propagation amplitudes for the Dirac field is by now a straightforward exercise:

$$\begin{aligned}
\langle 0|\, \psi_a(x)\bar{\psi}_b(y)\, |0\rangle &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \sum_s u^s_a(p)\bar{u}^s_b(p) e^{-ip\cdot(x-y)} \\
&= \left(i\partial\!\!\!/_x + m\right)_{ab} \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} e^{-ip\cdot(x-y)}, \qquad (3.114) \\
\langle 0|\, \bar{\psi}_b(y)\psi_a(x)\, |0\rangle &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \sum_s v^s_a(p)\bar{v}^s_b(p) e^{-ip\cdot(y-x)} \\
&= -\left(i\partial\!\!\!/_x + m\right)_{ab} \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} e^{-ip\cdot(y-x)}. \qquad (3.115)
\end{aligned}$$

Just as we did for the Klein-Gordon equation, we can construct Green's functions for the Dirac equation obeying various boundary conditions. For example, the retarded Green's function is

$$S^{ab}_R(x-y) \equiv \theta(x^0 - y^0)\, \langle 0|\, \{\psi_a(x), \bar{\psi}_b(y)\}\, |0\rangle\,. \tag{3.116}$$

It is easy to verify that

$$S_R(x-y) = \left(i\not{\partial}_x + m\right) D_R(x-y), \tag{3.117}$$

since on the right-hand side the term involving $\partial_0 \theta(x^0 - y^0)$ vanishes. Using (3.117) and the fact that $\not{\partial}\not{\partial} = \partial^2$, we see that S_R is a Green's function of the Dirac operator:

$$\left(i\not{\partial}_x - m\right) S_R(x-y) = i\delta^{(4)}(x-y) \cdot \mathbf{1}_{4\times 4}. \tag{3.118}$$

The Green's function of the Dirac operator can also be found by Fourier transformation. Expanding $S_R(x-y)$ as a Fourier integral and acting on both sides with $(i\not{\partial}_x - m)$, we find

$$i\delta^{(4)}(x-y) = \int \frac{d^4p}{(2\pi)^4} (\not{p} - m) e^{-ip\cdot(x-y)} \widetilde{S}_R(p), \tag{3.119}$$

and hence

$$\widetilde{S}_R(p) = \frac{i}{\not{p} - m} = \frac{i(\not{p} + m)}{p^2 - m^2}. \tag{3.120}$$

To obtain the retarded Green's function, we must evaluate the p^0 integral in (3.120) along the contour shown on page 30. For $x^0 > y^0$ we close the contour below, picking up both poles to obtain the sum of (3.114) and (3.115). For $x^0 < y^0$ we close the contour above and get zero.

The Green's function with Feynman boundary conditions is defined by the contour shown on page 31:

$$\begin{aligned} S_F(x-y) &= \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)} \\ &= \begin{cases} \langle 0|\, \psi(x)\bar{\psi}(y)\, |0\rangle & \text{for } x^0 > y^0 \text{ (close contour below)} \\ -\langle 0|\, \bar{\psi}(y)\psi(x)\, |0\rangle & \text{for } x^0 < y^0 \text{ (close contour above)} \end{cases} \\ &\equiv \langle 0|\, T\psi(x)\bar{\psi}(y)\, |0\rangle\,, \end{aligned} \tag{3.121}$$

where we have chosen to define the time-ordered product of spinor fields with an additional minus sign when the operators are interchanged. This minus sign is extremely important in the quantum field theory of fermions; we will meet it again in Section 4.7.

As with the Klein-Gordon theory, the expression (3.121) for the Feynman propagator is the most useful result of this chapter. When we do perturbative calculations with Feynman diagrams, we will associate the factor $\widetilde{S}_F(p)$ with each internal fermion line.

3.6 Discrete Symmetries of the Dirac Theory

In the last section we discussed the implementation of continuous Lorentz transformations on the Hilbert space of the Dirac theory. We found that for each transformation Λ there was a unitary operator $U(\Lambda)$, which induced the correct transformation on the fields:

$$U(\Lambda)\psi(x)U^{-1}(\Lambda) = \Lambda_{\frac{1}{2}}^{-1}\psi(\Lambda x). \tag{3.122}$$

In this section we will discuss the analogous operators that implement various discrete symmetries on the Dirac field.

In addition to continuous Lorentz transformations, there are two other spacetime operations that are potential symmetries of the Lagrangian: *parity* and *time reversal*. Parity, denoted by P, sends $(t, \mathbf{x}) \to (t, -\mathbf{x})$, reversing the handedness of space. Time reversal, denoted by T, sends $(t, \mathbf{x}) \to (-t, \mathbf{x})$, interchanging the forward and backward light-cones. Neither of these operations can be achieved by a continuous Lorentz transformation starting from the identity. Both, however, preserve the Minkowski interval $x^2 = t^2 - \mathbf{x}^2$. In standard terminology, the continuous Lorentz transformations are referred to as the proper, orthochronous Lorentz group, $\mathbf{L}_+^\uparrow$. Then the full Lorentz group breaks up into four disconnected subsets, as shown below.

$$\begin{array}{ccc} \mathbf{L}_+^\uparrow & \overset{P}{\longleftrightarrow} & \mathbf{L}_-^\uparrow = P\mathbf{L}_+^\uparrow \quad \text{“orthochronous”} \\ T\updownarrow & & \updownarrow T \\ \mathbf{L}_+^\downarrow = T\mathbf{L}_+^\uparrow & \underset{P}{\longleftrightarrow} & \mathbf{L}_-^\downarrow = PT\mathbf{L}_+^\uparrow \quad \text{“nonorthochronous”} \\ \text{“proper”} & & \text{“improper”} \end{array}$$

At the same time that we discuss P and T, it will be convenient to discuss a third (non-spacetime) discrete operation: *charge conjugation*, denoted by C. Under this operation, particles and antiparticles are interchanged.

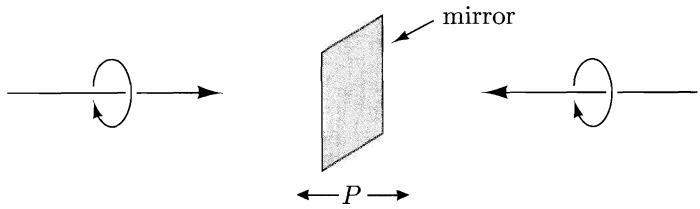
Although any relativistic field theory must be invariant under $\mathbf{L}_+^\uparrow$, it need not be invariant under P, T, or C. What is the status of these symmetry operations in the real world? From experiment, we know that three of the forces of Nature— the gravitational, electromagnetic, and strong interactions—are symmetric with respect to P, C, and T. The weak interactions violate C and P separately, but preserve CP and T. But certain rare processes (all so far observed involve neutral K mesons) also show CP and T violation. All observations indicate that the combination CPT is a perfect symmetry of Nature.

The currently accepted theoretical model of the weak interactions is the Glashow-Weinberg-Salam gauge theory, described in Chapter 20. This theory violates C and P in the strongest possible way. It is actually a surprise (though not quite an accident) that C and P happen to be quite good symmetries in the most readily observable processes. On the other hand, no one knows a really beautiful theory that violates CP. In the current theory, when there are three (or more) fermion generations, there is room for a parameter that, if nonzero,

causes CP violation. But the value of this parameter is no better understood than the value of the electron mass; the physical origin of CP violation remains a mystery. We will discuss this question further in Section 20.3.

Parity

With this introduction, let us now discuss the action of P, T, and C on Dirac particles and fields. First consider parity. The operator P should reverse the momentum of a particle without flipping its spin:



Mathematically, this means that P should be implemented by a unitary operator (properly called $U(P)$, but we'll just call it P) which, for example, transforms the state $a^{s\dagger}_{\mathbf{p}}|0\rangle$ into $a^{s\dagger}_{-\mathbf{p}}|0\rangle$. In other words, we want

$$Pa^s_{\mathbf{p}}P = \eta_a a^s_{-\mathbf{p}} \qquad \text{and} \qquad Pb^s_{\mathbf{p}}P = \eta_b b^s_{-\mathbf{p}}, \tag{3.123}$$

where η_a and η_b are possible phases. These phases are restricted by the condition that two applications of the parity operator should return observables to their original values. Since observables are built from an even number of fermion operators, this requires η_a^2, $\eta_b^2 = \pm 1$.

Just as a continuous Lorentz transformation is implemented on the Dirac field as the 4×4 constant matrix $\Lambda_{\frac{1}{2}}$, the parity transformation should also be represented by a 4×4 constant matrix. To find this matrix, and to determine η_a and η_b, we compute the action of P on $\psi(x)$. Using (3.123), we have

$$P\psi(x)P = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\sum_s \left(\eta_a a^s_{-\mathbf{p}}u^s(p)e^{-ipx} + \eta_b^* b^{s\dagger}_{-\mathbf{p}}v^s(p)e^{ipx}\right). \tag{3.124}$$

Now change variables to $\tilde{p} = (p^0, -\mathbf{p})$. Note that $p\cdot x = \tilde{p}\cdot(t,-\mathbf{x})$. Also $\tilde{p}\cdot\sigma = p\cdot\bar{\sigma}$ and $\tilde{p}\cdot\bar{\sigma} = p\cdot\sigma$. This allows us to write

$$u(p) = \begin{pmatrix}\sqrt{p\cdot\sigma}\,\xi\\ \sqrt{p\cdot\bar{\sigma}}\,\xi\end{pmatrix} = \begin{pmatrix}\sqrt{\tilde{p}\cdot\bar{\sigma}}\,\xi\\ \sqrt{\tilde{p}\cdot\sigma}\,\xi\end{pmatrix} = \gamma^0 u(\tilde{p});$$

$$v(p) = \begin{pmatrix}\sqrt{p\cdot\sigma}\,\xi\\ -\sqrt{p\cdot\bar{\sigma}}\,\xi\end{pmatrix} = \begin{pmatrix}\sqrt{\tilde{p}\cdot\bar{\sigma}}\,\xi\\ -\sqrt{\tilde{p}\cdot\sigma}\,\xi\end{pmatrix} = -\gamma^0 v(\tilde{p}).$$

So (3.124) becomes

$$P\psi(x)P = \int \frac{d^3\tilde{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\tilde{p}}}}\sum_s \Big(\eta_a a^s_{\tilde{p}}\gamma^0 u^s(\tilde{p})e^{-i\tilde{p}(t,-\mathbf{x})} - \eta_b^* b^{s\dagger}_{\tilde{p}}\gamma^0 v^s(\tilde{p})e^{i\tilde{p}(t,-\mathbf{x})}\Big).$$

This should equal some constant matrix times $\psi(t,-\mathbf{x})$, and indeed it works if we make $\eta_b^* = -\eta_a$. This implies

$$\eta_a \eta_b = -\eta_a \eta_a^* = -1. \tag{3.125}$$

Thus we have the parity transformation of $\psi(x)$ in its final form,

$$P\psi(t,\mathbf{x})P = \eta_a \gamma^0 \psi(t,-\mathbf{x}). \tag{3.126}$$

It will be very important (for example, in writing down Lagrangians) to know how the various Dirac field bilinears transform under parity. Recall that the five bilinears are

$$\bar{\psi}\psi, \qquad \bar{\psi}\gamma^\mu\psi, \qquad i\bar{\psi}\big[\gamma^\mu,\gamma^\nu\big]\psi, \qquad \bar{\psi}\gamma^\mu\gamma^5\psi, \qquad i\bar{\psi}\gamma^5\psi. \tag{3.127}$$

The factors of i have been chosen to make all these quantities Hermitian, as you can easily verify. (Any new term that we add to a Lagrangian must be real.) First we should compute

$$P\bar{\psi}(t,\mathbf{x})P = P\psi^\dagger(t,\mathbf{x})P\gamma^0 = \big(P\psi(t,\mathbf{x})P\big)^\dagger\gamma^0 = \eta_a^*\bar{\psi}(t,-\mathbf{x})\gamma^0. \tag{3.128}$$

Then the scalar bilinear transforms as

$$P\bar{\psi}\psi P = |\eta_a|^2\bar{\psi}(t,-\mathbf{x})\gamma^0\gamma^0\psi(t,-\mathbf{x}) = +\bar{\psi}\psi(t,-\mathbf{x}), \tag{3.129}$$

while for the vector we obtain

$$P\bar{\psi}\gamma^\mu\psi P = \bar{\psi}\gamma^0\gamma^\mu\gamma^0\psi(t,-\mathbf{x}) = \begin{cases} +\bar{\psi}\gamma^\mu\psi(t,-\mathbf{x}) & \text{for } \mu = 0, \\ -\bar{\psi}\gamma^\mu\psi(t,-\mathbf{x}) & \text{for } \mu = 1,2,3. \end{cases} \tag{3.130}$$

Note that the vector acquires the same minus sign on the spatial components as does the vector x^μ. Similarly, the transformations of the pseudo-scalar and pseudo-vector are

$$Pi\bar{\psi}\gamma^5\psi P = i\bar{\psi}\gamma^0\gamma^5\gamma^0\psi(t,-\mathbf{x}) = -i\bar{\psi}\gamma^5\psi(t,-\mathbf{x}); \tag{3.131}$$

$$P\bar{\psi}\gamma^\mu\gamma^5\psi P = \bar{\psi}\gamma^0\gamma^\mu\gamma^5\gamma^0\psi(t,-\mathbf{x}) = \begin{cases} -\bar{\psi}\gamma^\mu\gamma^5\psi & \text{for } \mu = 0, \\ +\bar{\psi}\gamma^\mu\gamma^5\psi & \text{for } \mu = 1,2,3. \end{cases} \tag{3.132}$$

Just as we anticipated in Section 3.4, the "pseudo" signifies an extra minus sign in the parity transformation. (The transformation properties of $i\bar{\psi}[\gamma^\mu,\gamma^\nu]\psi = 2\bar{\psi}\sigma^{\mu\nu}\psi$ are reserved for Problem 3.7.) Note that the transformation properties of fermion bilinears were independent of η_a, so there would have been no loss of generality in setting $\eta_a = -\eta_b = 1$ from the beginning.

However, the relative minus sign (3.125) between the parity transformations of a fermion and an antifermion has important consequences. Consider a fermion-antifermion state, $a_{\mathbf{p}}^{s\dagger}b_{\mathbf{q}}^{s'\dagger}|0\rangle$. Applying P, we find $P\big(a_{\mathbf{p}}^{s\dagger}b_{\mathbf{q}}^{s'\dagger}|0\rangle\big) = -\big(a_{-\mathbf{p}}^{s\dagger}b_{-\mathbf{q}}^{s'\dagger}|0\rangle\big)$. Thus a state containing a fermion-antifermion pair gets an extra (-1) under parity. This information is most useful in the context of bound states, in which the fermion and antifermion momenta are integrated with the Schrödinger wavefunction to produce a system localized in space. We consider

such states in detail in Section 5.3, but here we should remark that if the spatial wavefunction is symmetric under $\mathbf{x} \to -\mathbf{x}$, the state has *odd* parity, while if it is antisymmetric under $\mathbf{x} \to -\mathbf{x}$, the state has *even* parity. The $L = 0$ bound states, for example, have odd parity; the $J = 0$ state transforms as a pseudo-scalar, while the three $J = 1$ states transform as the spatial components of a vector. These properties show up in selection rules for decays of positronium and quark-antiquark systems (see Problem 3.8).

Time Reversal

Now let us turn to the implementation of time reversal. We would like T to take the form of a unitary operator that sends $a_{\mathbf{p}}$ to $a_{-\mathbf{p}}$ (and similarly for $b_{\mathbf{p}}$) and $\psi(t, \mathbf{x})$ to $\psi(-t, \mathbf{x})$ (times some constant matrix). These properties, however, are extremely difficult to achieve, since we saw above that sending $a_{\mathbf{p}}$ to $a_{-\mathbf{p}}$ instead sends $(t, \mathbf{x})$ to $(t, -\mathbf{x})$ in the expansion of ψ. The difficulty is even more apparent when we impose the constraint that time reversal should be a symmetry of the free Dirac theory, $[T, H] = 0$. Then

$$\begin{aligned} &\psi(t, \mathbf{x}) = e^{iHt}\psi(\mathbf{x})e^{-iHt} \\ \Rightarrow\ &T\psi(t, \mathbf{x})T = e^{iHt}\big[T\psi(\mathbf{x})T\big]e^{-iHt} \\ \Rightarrow\ &T\psi(t, \mathbf{x})T\,|0\rangle = e^{iHt}\big[T\psi(\mathbf{x})T\big]\,|0\rangle\,, \end{aligned}$$

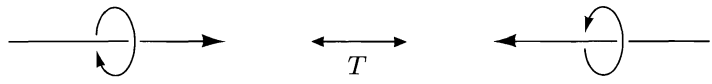
assuming that $H\,|0\rangle = 0$. The right-hand side is a sum of negative-frequency terms only. But if T is to reverse the time dependence of $\psi(t, \mathbf{x})$, then the left-hand side is (up to a constant matrix) $\psi(-t, \mathbf{x})\,|0\rangle = e^{-iHt}\psi(\mathbf{x})\,|0\rangle$, which is a sum of positive-frequency terms. Thus we have proved that T cannot be implemented as a linear unitary operator.

What can we do? The way out is to retain the unitarity condition $T^\dagger = T^{-1}$, but have T act on c-numbers as well as operators, as follows:

$$T(\text{c-number}) = (\text{c-number})^* T. \tag{3.133}$$

Then even if $[T, H] = 0$, the time dependence of all exponential factors is reversed: $Te^{+iHt} = e^{-iHt}T$. Since all time evolution in quantum mechanics is performed with such exponential factors, this effectively changes the sign of t. Note that the operation of complex conjugation is nonlinear; T is referred to as an *antilinear* or *antiunitary* operator.

In addition to reversing the momentum of a particle, T should also flip the spin:



To quantify this, we must find a mathematical operation that flips a spinor ξ.

In the earlier parts of this chapter, we denoted the spin state of a fermion by a label $s = 1, 2$. In the remainder of this section, we will associate s with the physical spin component of the fermion along a specific axis. If this axis

has polar coordinates θ, ϕ, the two-component spinors with spin up and spin down along this axis are

$$\xi(\uparrow) = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix}, \qquad \xi(\downarrow) = \begin{pmatrix} -e^{-i\phi}\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{pmatrix}.$$

Let $\xi^s = (\xi(\uparrow), \xi(\downarrow))$ for $s = 1, 2$. Also define

$$\xi^{-s} = -i\sigma^2(\xi^s)^*. \tag{3.134}$$

This quantity is the flipped spinor; from the explicit formulae,

$$\xi^{-s} = (\xi(\downarrow), -\xi(\uparrow)). \tag{3.135}$$

The form of the spin reversal relation follows more generally from the identity $\boldsymbol{\sigma}\sigma^2 = \sigma^2(-\boldsymbol{\sigma}^*)$. This equation implies that, if ξ satisfies $\mathbf{n}\cdot\boldsymbol{\sigma}\xi = +\xi$ for some axis $\mathbf{n}$, then

$$(\mathbf{n}\cdot\boldsymbol{\sigma})(-i\sigma^2\xi^*) = -i\sigma^2(-\mathbf{n}\cdot\boldsymbol{\sigma})^*\xi^* = i\sigma^2(\xi^*) = -(-i\sigma^2\xi^*).$$

Notice that, with this convention for the spin flip, two successive spin flips return a spin to (-1) times the original state.

We now associate the various fermion spin states with these spinors. The electron annihilation operator $a^s_{\mathbf{p}}$ destroys an electron whose spinor $u^s(p)$ contains ξ^s. The positron annihilation operator $b^s_{\mathbf{p}}$ destroys a positron whose spinor $v^s(p)$ contains ξ^{-s}:

$$v^s(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^{-s} \\ -\sqrt{p\cdot\bar{\sigma}}\,\xi^{-s} \end{pmatrix}. \tag{3.136}$$

As in Eq. (3.135), we define

$$a^{-s}_{\mathbf{p}} = (a^2_{\mathbf{p}}, -a^1_{\mathbf{p}}), \qquad b^{-s}_{\mathbf{p}} = (b^2_{\mathbf{p}}, -b^1_{\mathbf{p}}). \tag{3.137}$$

We can now work out the relation between the Dirac spinors u and v and their time reversals. Define $\tilde{p} = (p^0, -\mathbf{p})$. This vector satisfies the identity $\sqrt{\tilde{p}\cdot\sigma}\,\sigma^2 = \sigma^2\sqrt{p\cdot\sigma^*}$; to prove this, expand the square root as in (3.49). For some choice of spin and momentum, associated with the Dirac spinor $u^s(p)$, let $u^{-s}(\tilde{p})$ be the spinor with the reversed momentum and flipped spin. These quantities are related by

$$\begin{aligned} u^{-s}(\tilde{p}) &= \begin{pmatrix} \sqrt{\tilde{p}\cdot\sigma}\,(-i\sigma^2\xi^{s*}) \\ \sqrt{\tilde{p}\cdot\bar{\sigma}}\,(-i\sigma^2\xi^{s*}) \end{pmatrix} = \begin{pmatrix} -i\sigma^2\sqrt{p\cdot\sigma^*}\,\xi^{s*} \\ -i\sigma^2\sqrt{p\cdot\bar{\sigma}^*}\,\xi^{s*} \end{pmatrix} \\ &= -i\begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}\left[u^s(p)\right]^* = -\gamma^1\gamma^3\left[u^s(p)\right]^*. \end{aligned}$$

Similarly, for $v^s(p)$,

$$v^{-s}(\tilde{p}) = -\gamma^1\gamma^3\left[v^s(p)\right]^*;$$

in this relation, v^{-s} contains $\xi^{-(-s)} = -\xi^s$.

Using the notation of Eq. (3.137), we define the time reversal transformation of fermion annihilation operators as follows:

$$Ta^s_{\mathbf{p}}T = a^{-s}_{-\mathbf{p}}, \qquad Tb^s_{\mathbf{p}}T = b^{-s}_{-\mathbf{p}}. \tag{3.138}$$

(An additional overall phase would have no effect on the rest of our discussion and is omitted for simplicity.) Relations (3.138) allow us to compute the action of T on the fermion field $\psi(x)$:

$$\begin{aligned} T\psi(t,\mathbf{x})T &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\sum_s T\Big(a^s_{\mathbf{p}}u^s(p)e^{-ipx} + b^{s\dagger}_{\mathbf{p}}v^s(p)e^{ipx}\Big)T \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\sum_s \Big(a^{-s}_{-\mathbf{p}}\big[u^s(p)\big]^* e^{ipx} + b^{-s\dagger}_{-\mathbf{p}}\big[v^s(p)\big]^* e^{-ipx}\Big) \\ &= (\gamma^1\gamma^3)\int \frac{d^3\tilde{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\tilde{p}}}}\sum_s \Big(a^{-s}_{\tilde{p}}u^{-s}(\tilde{p})e^{i\tilde{p}(t,-\mathbf{x})} \\ &\qquad\qquad + b^{-s\dagger}_{\tilde{p}}v^{-s}(\tilde{p})e^{-i\tilde{p}(t,-\mathbf{x})}\Big) \\ &= (\gamma^1\gamma^3)\psi(-t,\mathbf{x}). \end{aligned} \tag{3.139}$$

In the last step we used $\tilde{p}\cdot(t,-\mathbf{x}) = -\tilde{p}\cdot(-t,\mathbf{x})$. Just as for parity, we have derived a simple transformation law for the fermion field $\psi(x)$. The relative minus sign in the transformation laws for particle and antiparticle is present here as well, implicit in the twice-flipped spinor in v^{-s}.

Now we can check the action of T on the various bilinears. First we need

$$T\bar{\psi}T = (T\psi T)^\dagger(\gamma^0)^* = \psi^\dagger(-t,\mathbf{x})\big[\gamma^1\gamma^3\big]^\dagger\gamma^0 = \bar{\psi}(-t,\mathbf{x})\big[-\gamma^1\gamma^3\big]. \tag{3.140}$$

Then the transformation of the scalar bilinear is

$$T\bar{\psi}\psi(t,\mathbf{x})T = \bar{\psi}(-\gamma^1\gamma^3)(\gamma^1\gamma^3)\psi(-t,\mathbf{x}) = +\bar{\psi}\psi(-t,\mathbf{x}). \tag{3.141}$$

The pseudo-scalar acquires an extra minus sign when T goes through the i:

$$Ti\bar{\psi}\gamma^5\psi T = -i\bar{\psi}(-\gamma^1\gamma^3)\gamma^5(\gamma^1\gamma^3)\psi = -i\bar{\psi}\gamma^5\psi(-t,\mathbf{x}).$$

For the vector, we must separately compute each of the four cases $\mu = 0,1,2,3$. After a bit of work you should find

$$\begin{aligned} T\bar{\psi}\gamma^\mu\psi T &= \bar{\psi}(-\gamma^1\gamma^3)(\gamma^\mu)^*(\gamma^1\gamma^3)\psi \\ &= \begin{cases} +\bar{\psi}\gamma^\mu\psi(-t,\mathbf{x}) & \text{for } \mu = 0; \\ -\bar{\psi}\gamma^\mu\psi(-t,\mathbf{x}) & \text{for } \mu = 1,2,3. \end{cases} \end{aligned} \tag{3.142}$$

This is exactly the tranformation property we want for vectors such as the current density. You can verify that the pseudo-vector transforms in exactly the same way under time-reversal.

Charge Conjugation

The last of the three discrete symmetries is the particle-antiparticle symmetry C. There will be no problem in implementing C as a unitary linear operator. Charge conjugation is conventionally defined to take a fermion with a given spin orientation into an antifermion with the same spin orientation. Thus, a convenient choice for the transformation of fermion annihilation operators is

$$Ca^s_{\mathbf{p}}C = b^s_{\mathbf{p}}; \qquad Cb^s_{\mathbf{p}}C = a^s_{\mathbf{p}}. \tag{3.143}$$

Again, we ignore possible additional phases for simplicity.

Next we want to work out the action of C on $\psi(x)$. First we need a relation between $v^s(p)$ and $u^s(p)$. Using (3.136), and (3.134),

$$\big(v^s(p)\big)^* = \begin{pmatrix} \sqrt{p\cdot\sigma}(-i\sigma^2\xi^*) \\ -\sqrt{p\cdot\bar\sigma}(-i\sigma^2\xi^*) \end{pmatrix}^* = \begin{pmatrix} -i\sigma^2\sqrt{p\cdot\bar\sigma^*}\xi^* \\ i\sigma^2\sqrt{p\cdot\sigma^*}\xi^* \end{pmatrix}^* = \begin{pmatrix} 0 & -i\sigma^2 \\ i\sigma^2 & 0 \end{pmatrix}\begin{pmatrix} \sqrt{p\cdot\sigma}\xi \\ \sqrt{p\cdot\bar\sigma}\xi \end{pmatrix},$$

where ξ stands for ξ^s. That is,

$$u^s(p) = -i\gamma^2\big(v^s(p)\big)^*, \qquad v^s(p) = -i\gamma^2\big(u^s(p)\big)^*. \tag{3.144}$$

If we substitute (3.144) into the expression for the fermion field operator, and then transform this operator with C, we find

$$\begin{aligned} C\psi(x)C &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\sum_s\Big(-i\gamma^2 b^s_{\mathbf{p}}\big(v^s(p)\big)^* e^{-ipx} - i\gamma^2 a^{s\dagger}_{\mathbf{p}}\big(u^s(p)\big)^* e^{ipx}\Big) \\ &= -i\gamma^2\psi^*(x) = -i\gamma^2(\psi^\dagger)^T = -i\big(\bar\psi\gamma^0\gamma^2\big)^T. \end{aligned} \tag{3.145}$$

Note that C is a linear unitary operator, even though it takes $\psi \to \psi^*$.

Once again, we would like to know how C acts on fermion bilinears. First we need

$$C\bar\psi(x)C = C\psi^\dagger C\gamma^0 = (-i\gamma^2\psi)^T\gamma^0 = (-i\gamma^0\gamma^2\psi)^T. \tag{3.146}$$

Working out the transformations of bilinears is a bit tricky, and it helps to write in spinor indices. For the scalar,

$$\begin{aligned} C\bar\psi\psi C &= (-i\gamma^0\gamma^2\psi)^T(-i\bar\psi\gamma^0\gamma^2)^T = -\gamma^0_{ab}\gamma^2_{bc}\psi_c\bar\psi_d\gamma^0_{de}\gamma^2_{ea} \\ &= +\bar\psi_d\gamma^0_{de}\gamma^2_{ea}\gamma^0_{ab}\gamma^2_{bc}\psi_c = -\bar\psi\gamma^2\gamma^0\gamma^0\gamma^2\psi \\ &= +\bar\psi\psi. \end{aligned} \tag{3.147}$$

(The minus sign in the third step is from fermion anticommutation.) The pseudo-scalar is no more difficult:

$$Ci\bar\psi\gamma^5\psi C = i(-i\gamma^0\gamma^2\psi)^T\gamma^5(-i\bar\psi\gamma^0\gamma^2)^T = i\bar\psi\gamma^5\psi. \tag{3.148}$$

We must do each component of the vector and pseudo-vector separately. Noting that γ^0 and γ^2 are symmetric matrices while γ^1 and γ^3 are antisymmetric,

we eventually find

$$C\bar{\psi}\gamma^\mu\psi C = -\bar{\psi}\gamma^\mu\psi; \tag{3.149}$$

$$C\bar{\psi}\gamma^\mu\gamma^5\psi C = +\bar{\psi}\gamma^\mu\gamma^5\psi. \tag{3.150}$$

Although the operator C interchanges ψ and $\bar{\psi}$, it does not actually change the order of the creation and annihilation operators. Thus, if $\bar{\psi}\gamma^0\psi$ is defined to subtract the infinite constant noted above Eq. (3.113), this constant does not reappear in the process of conjugation by C.

Summary of C, P, and T

The transformation properties of the various fermion bilinears under C, P, and T are summarized in the table below. Here we use the shorthand $(-1)^\mu \equiv 1$ for $\mu = 0$ and $(-1)^\mu \equiv -1$ for $\mu = 1, 2, 3$.

	$\bar{\psi}\psi$	$i\bar{\psi}\gamma^5\psi$	$\bar{\psi}\gamma^\mu\psi$	$\bar{\psi}\gamma^\mu\gamma^5\psi$	$\bar{\psi}\sigma^{\mu\nu}\psi$	∂_μ
P	$+1$	-1	$(-1)^\mu$	$-(-1)^\mu$	$(-1)^\mu(-1)^\nu$	$(-1)^\mu$
T	$+1$	-1	$(-1)^\mu$	$(-1)^\mu$	$-(-1)^\mu(-1)^\nu$	$-(-1)^\mu$
C	$+1$	$+1$	-1	$+1$	-1	$+1$
CPT	$+1$	$+1$	-1	-1	$+1$	-1

We have included the transformation properties of the tensor bilinear (see Problem 3.7), and also of the derivative operator.

Notice first that the free Dirac Lagrangian $\mathcal{L}_0 = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi$ is invariant under C, P, and T separately. We can build more general quantum systems that violate any of these symmetries by adding to $\mathcal{L}_0$ some perturbation $\delta\mathcal{L}$. But $\delta\mathcal{L}$ must be a Lorentz scalar, and the last line of the table shows that all Lorentz scalar combinations of $\bar{\psi}$ and ψ are invariant under the combined symmetry CPT. Actually, it is quite generally true that one cannot build a Lorentz-invariant quantum field theory with a Hermitian Hamiltonian that violates CPT.[†]

Problems

3.1 Lorentz group. Recall from Eq. (3.17) the Lorentz commutation relations,

$$[J^{\mu\nu}, J^{\rho\sigma}] = i(g^{\nu\rho}J^{\mu\sigma} - g^{\mu\rho}J^{\nu\sigma} - g^{\nu\sigma}J^{\mu\rho} + g^{\mu\sigma}J^{\nu\rho}).$$

(a) Define the generators of rotations and boosts as

$$L^i = \tfrac{1}{2}\epsilon^{ijk}J^{jk}, \qquad K^i = J^{0i},$$

[†]This theorem and the spin-statistics theorem are proved with great care in Streater and Wightman, *op. cit.*

where $i, j, k = 1, 2, 3$. An infinitesimal Lorentz transformation can then be written

$$\Phi \rightarrow (1 - i\boldsymbol{\theta} \cdot \mathbf{L} - i\boldsymbol{\beta} \cdot \mathbf{K})\Phi.$$

Write the commutation relations of these vector operators explicitly. (For example, $[L^i, L^j] = i\epsilon^{ijk}L^k$.) Show that the combinations

$$\mathbf{J}_+ = \tfrac{1}{2}(\mathbf{L} + i\mathbf{K}) \quad \text{and} \quad \mathbf{J}_- = \tfrac{1}{2}(\mathbf{L} - i\mathbf{K})$$

commute with one another and separately satisfy the commutation relations of angular momentum.

(b) The finite-dimensional representations of the rotation group correspond precisely to the allowed values for angular momentum: integers or half-integers. The result of part (a) implies that all finite-dimensional representations of the Lorentz group correspond to pairs of integers or half integers, (j_+, j_-), corresponding to pairs of representations of the rotation group. Using the fact that $\mathbf{J} = \boldsymbol{\sigma}/2$ in the spin-1/2 representation of angular momentum, write explicitly the transformation laws of the 2-component objects transforming according to the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations of the Lorentz group. Show that these correspond precisely to the transformations of ψ_L and ψ_R given in (3.37).

(c) The identity $\boldsymbol{\sigma}^T = -\sigma^2\boldsymbol{\sigma}\sigma^2$ allows us to rewrite the ψ_L transformation in the unitarily equivalent form

$$\psi' \rightarrow \psi'(1 + i\boldsymbol{\theta} \cdot \frac{\boldsymbol{\sigma}}{2} + \boldsymbol{\beta} \cdot \frac{\boldsymbol{\sigma}}{2}),$$

where $\psi' = \psi_L^T\sigma^2$. Using this law, we can represent the object that transforms as $(\frac{1}{2}, \frac{1}{2})$ as a 2×2 matrix that has the ψ_R transformation law on the left and, simultaneously, the transposed ψ_L transformation on the right. Parametrize this matrix as

$$\begin{pmatrix} V^0 + V^3 & V^1 - iV^2 \\ V^1 + iV^2 & V^0 - V^3 \end{pmatrix}.$$

Show that the object V^μ transforms as a 4-vector.

3.2 Derive the *Gordon identity*,

$$\bar{u}(p')\gamma^\mu u(p) = \bar{u}(p')\left[\frac{p'^\mu + p^\mu}{2m} + \frac{i\sigma^{\mu\nu}q_\nu}{2m}\right]u(p),$$

where $q = (p' - p)$. We will put this formula to use in Chapter 6.

3.3 Spinor products. (This problem, together with Problems 5.3 and 5.6, introduces an efficient computational method for processes involving massless particles.) Let k_0^μ, k_1^μ be fixed 4-vectors satisfying $k_0^2 = 0$, $k_1^2 = -1$, $k_0 \cdot k_1 = 0$. Define basic spinors in the following way: Let u_{L0} be the left-handed spinor for a fermion with momentum k_0. Let $u_{R0} = \not{k}_1 u_{L0}$. Then, for any p such that p is lightlike ($p^2 = 0$), define

$$u_L(p) = \frac{1}{\sqrt{2p \cdot k_0}}\not{p}u_{R0} \qquad \text{and} \qquad u_R(p) = \frac{1}{\sqrt{2p \cdot k_0}}\not{p}u_{L0}.$$

This set of conventions defines the phases of spinors unambiguously (except when p is parallel to k_0).

(a) Show that $\not{k}_0 u_{R0} = 0$. Show that, for any lightlike p, $\not{p}u_L(p) = \not{p}u_R(p) = 0$.

(b) For the choices $k_0 = (E, 0, 0, -E)$, $k_1 = (0, 1, 0, 0)$, construct u_{L0}, u_{R0}, $u_L(p)$, and $u_R(p)$ explicitly.

(c) Define the *spinor products* $s(p_1, p_2)$ and $t(p_1, p_2)$, for p_1, p_2 lightlike, by

$$s(p_1, p_2) = \bar{u}_R(p_1) u_L(p_2), \qquad t(p_1, p_2) = \bar{u}_L(p_1) u_R(p_2).$$

Using the explicit forms for the u_λ given in part (b), compute the spinor products explicitly and show that $t(p_1, p_2) = (s(p_2, p_1))^*$ and $s(p_1, p_2) = -s(p_2, p_1)$. In addition, show that

$$|s(p_1, p_2)|^2 = 2p_1 \cdot p_2.$$

Thus the spinor products are the square roots of 4-vector dot products.

3.4 Majorana fermions. Recall from Eq. (3.40) that one can write a relativistic equation for a massless 2-component fermion field that transforms as the upper two components of a Dirac spinor (ψ_L). Call such a 2-component field $\chi_a(x)$, $a = 1, 2$.

(a) Show that it is possible to write an equation for $\chi(x)$ as a massive field in the following way:

$$i\bar{\sigma} \cdot \partial \chi - im\sigma^2 \chi^* = 0.$$

That is, show, first, that this equation is relativistically invariant and, second, that it implies the Klein-Gordon equation, $(\partial^2 + m^2)\chi = 0$. This form of the fermion mass is called a Majorana mass term.

(b) Does the Majorana equation follow from a Lagrangian? The mass term would seem to be the variation of $(\sigma^2)_{ab}\chi_a^*\chi_b^*$; however, since σ^2 is antisymmetric, this expression would vanish if $\chi(x)$ were an ordinary c-number field. When we go to quantum field theory, we know that $\chi(x)$ will become an anticommuting quantum field. Therefore, it makes sense to develop its classical theory by considering $\chi(x)$ as a classical anticommuting field, that is, as a field that takes as values *Grassmann numbers* which satisfy

$$\alpha\beta = -\beta\alpha \qquad \text{for any } \alpha, \beta.$$

Note that this relation implies that $\alpha^2 = 0$. A Grassmann field $\xi(x)$ can be expanded in a basis of functions as

$$\xi(x) = \sum_n \alpha_n \phi_n(x),$$

where the $\phi_n(x)$ are orthogonal c-number functions and the α_n are a set of independent Grassmann numbers. Define the complex conjugate of a product of Grassmann numbers to reverse the order:

$$(\alpha\beta)^* \equiv \beta^*\alpha^* = -\alpha^*\beta^*.$$

This rule imitates the Hermitian conjugation of quantum fields. Show that the classical action,

$$S = \int d^4x \left[\chi^\dagger i\bar{\sigma} \cdot \partial \chi + \frac{im}{2}(\chi^T \sigma^2 \chi - \chi^\dagger \sigma^2 \chi^*)\right],$$

(where $\chi^\dagger = (\chi^*)^T$) is real ($S^* = S$), and that varying this S with respect to χ and χ^* yields the Majorana equation.

(c) Let us write a 4-component Dirac field as

$$\psi(x) = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix},$$

and recall that the lower components of ψ transform in a way equivalent by a unitary transformation to the complex conjugate of the representation ψ_L. In this way, we can rewrite the 4-component Dirac field in terms of two 2-component spinors:

$$\psi_L(x) = \chi_1(x), \qquad \psi_R(x) = i\sigma^2\chi_2^*(x).$$

Rewrite the Dirac Lagrangian in terms of χ_1 and χ_2 and note the form of the mass term.

(d) Show that the action of part (c) has a global symmetry. Compute the divergences of the currents

$$J^\mu = \chi^\dagger\bar{\sigma}^\mu\chi, \qquad J^\mu = \chi_1^\dagger\bar{\sigma}^\mu\chi_1 - \chi_2^\dagger\bar{\sigma}^\mu\chi_2,$$

for the theories of parts (b) and (c), respectively, and relate your results to the symmetries of these theories. Construct a theory of N free massive 2-component fermion fields with $O(N)$ symmetry (that is, the symmetry of rotations in an N-dimensional space).

(e) Quantize the Majorana theory of parts (a) and (b). That is, promote $\chi(x)$ to a quantum field satisfying the canonical anticommutation relation

$$\{\chi_a(\mathbf{x}), \chi_b^\dagger(\mathbf{y})\} = \delta_{ab}\delta^{(3)}(\mathbf{x}-\mathbf{y}),$$

construct a Hermitian Hamiltonian, and find a representation of the canonical commutation relations that diagonalizes the Hamiltonian in terms of a set of creation and annihilation operators. (Hint: Compare $\chi(x)$ to the top two components of the quantized Dirac field.)

3.5 Supersymmetry. It is possible to write field theories with continuous symmetries linking fermions and bosons; such transformations are called *supersymmetries.*

(a) The simplest example of a supersymmetric field theory is the theory of a free complex boson and a free Weyl fermion, written in the form

$$\mathcal{L} = \partial_\mu\phi^*\partial^\mu\phi + \chi^\dagger i\bar{\sigma}\cdot\partial\chi + F^*F.$$

Here F is an auxiliary complex scalar field whose field equation is $F = 0$. Show that this Lagrangian is invariant (up to a total divergence) under the infinitesimal tranformation

$$\begin{aligned} \delta\phi &= -i\epsilon^T\sigma^2\chi, \\ \delta\chi &= \epsilon F + \sigma\cdot\partial\phi\,\sigma^2\epsilon^*, \\ \delta F &= -i\epsilon^\dagger\bar{\sigma}\cdot\partial\chi, \end{aligned}$$

where the parameter ϵ_a is a 2-component spinor of Grassmann numbers.

(b) Show that the term

$$\Delta\mathcal{L} = [m\phi F + \tfrac{1}{2}im\chi^T\sigma^2\chi] + (\text{complex conjugate})$$

is also left invariant by the transformation given in part (a). Eliminate F from the complete Lagrangian $\mathcal{L} + \Delta\mathcal{L}$ by solving its field equation, and show that the fermion and boson fields ϕ and χ are given the same mass.

(c) It is possible to write supersymmetric nonlinear field equations by adding cubic and higher-order terms to the Lagrangian. Show that the following rather general field theory, containing the field (ϕ_i, χ_i), $i = 1, \ldots, n$, is supersymmetric:

$$\mathcal{L} = \partial_\mu \phi_i^* \partial^\mu \phi_i + \chi_i^\dagger i\bar{\sigma} \cdot \partial \chi_i + F_i^* F_i + \left(F_i \frac{\partial W[\phi]}{\partial \phi_i} + \frac{i}{2} \frac{\partial^2 W[\phi]}{\partial \phi_i \partial \phi_j} \chi_i^T \sigma^2 \chi_j + \text{c.c.} \right),$$

where $W[\phi]$ is an arbitrary function of the ϕ_i, called the *superpotential*. For the simple case $n = 1$ and $W = g\phi^3/3$, write out the field equations for ϕ and χ (after elimination of F).

3.6 Fierz transformations. Let u_i, $i = 1, \ldots, 4$, be four 4-component Dirac spinors. In the text, we proved the Fierz rearrangement formulae (3.78) and (3.79). The first of these formulae can be written in 4-component notation as

$$\bar{u}_1 \gamma^\mu \left(\frac{1+\gamma^5}{2} \right) u_2 \bar{u}_3 \gamma_\mu \left(\frac{1+\gamma^5}{2} \right) u_4 = -\bar{u}_1 \gamma^\mu \left(\frac{1+\gamma^5}{2} \right) u_4 \bar{u}_3 \gamma_\mu \left(\frac{1+\gamma^5}{2} \right) u_2.$$

In fact, there are similar rearrangement formulae for any product

$$(\bar{u}_1 \Gamma^A u_2)(\bar{u}_3 \Gamma^B u_4),$$

where Γ^A, Γ^B are any of the 16 combinations of Dirac matrices listed in Section 3.4.

(a) To begin, normalize the 16 matrices Γ^A to the convention

$$\text{tr}[\Gamma^A \Gamma^B] = 4\delta^{AB}.$$

This gives $\Gamma^A = \{1, \gamma^0, i\gamma^j, \ldots\}$; write all 16 elements of this set.

(b) Write the general Fierz identity as an equation

$$(\bar{u}_1 \Gamma^A u_2)(\bar{u}_3 \Gamma^B u_4) = \sum_{C,D} C^{AB}{}_{CD} (\bar{u}_1 \Gamma^C u_4)(\bar{u}_3 \Gamma^D u_2),$$

with unknown coefficients $C^{AB}{}_{CD}$. Using the completeness of the 16 Γ^A matrices, show that

$$C^{AB}{}_{CD} = \frac{1}{16} \text{tr}[\Gamma^C \Gamma^A \Gamma^D \Gamma^B].$$

(c) Work out explicitly the Fierz transformation laws for the products $(\bar{u}_1 u_2)(\bar{u}_3 u_4)$ and $(\bar{u}_1 \gamma^\mu u_2)(\bar{u}_3 \gamma_\mu u_4)$.

3.7 This problem concerns the discrete symmetries P, C, and T.

(a) Compute the transformation properties under P, C, and T of the antisymmetric tensor fermion bilinears, $\bar{\psi} \sigma^{\mu\nu} \psi$, with $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. This completes the table of the transformation properties of bilinears at the end of the chapter.

(b) Let $\phi(x)$ be a complex-valued Klein-Gordon field, such as we considered in Problem 2.2. Find unitary operators P, C and an antiunitary operator T (all defined

in terms of their action on the annihilation operators $a_{\mathbf{p}}$ and $b_{\mathbf{p}}$ for the Klein-Gordon particles and antiparticles) that give the following tranformations of the Klein-Gordon field:

$$\begin{aligned} P\,\phi(t,\mathbf{x})\,P &= \phi(t,-\mathbf{x}); \\ T\,\phi(t,\mathbf{x})\,T &= \phi(-t,\mathbf{x}); \\ C\,\phi(t,\mathbf{x})\,C &= \phi^*(t,\mathbf{x}). \end{aligned}$$

Find the transformation properties of the components of the current

$$J^\mu = i(\phi^*\partial^\mu\phi - \partial^\mu\phi^*\phi)$$

under P, C, and T.

(c) Show that any Hermitian Lorentz-scalar local operator built from $\psi(x)$, $\phi(x)$, and their conjugates has $CPT = +1$.

3.8 Bound states. Two spin-1/2 particles can combine to a state of total spin either 0 or 1. The wavefunctions for these states are odd and even, respectively, under the interchange of the two spins.

(a) Use this information to compute the quantum numbers under P and C of all electron-positron bound states with S, P, or D wavefunctions.

(b) Since the electron-photon coupling is given by the Hamiltonian

$$\Delta H = \int d^3x\; eA_\mu j^\mu,$$

where j^μ is the electric current, electrodynamics is invariant to P and C if the components of the vector potential have the same P and C parity as the corresponding components of j^μ. Show that this implies the following surprising fact: The spin-0 ground state of positronium can decay to 2 photons, but the spin-1 ground state must decay to 3 photons. Find the selection rules for the annihilation of higher positronium states, and for 1-photon transitions between positronium levels.

Chapter 4

Interacting Fields and Feynman Diagrams

4.1 Perturbation Theory—Philosophy and Examples

We have now discussed in some detail the quantization of two free field theories that give approximate descriptions of many of the particles found in Nature. Up to this point, however, free-particle states have been eigenstates of the Hamiltonian; we have seen no interactions and no scattering. In order to obtain a closer description of the real world, we must include new, nonlinear terms in the Hamiltonian (or Lagrangian) that will couple different Fourier modes (and the particles that occupy them) to one another. To preserve causality, we insist that the new terms may involve only products of fields at the same spacetime point: $[\phi(x)]^4$ is fine, but $\phi(x)\phi(y)$ is not allowed. Thus the terms describing the interactions will be of the form

$$H_{\text{int}} = \int d^3x \, \mathcal{H}_{\text{int}}[\phi(x)] = -\int d^3x \, \mathcal{L}_{\text{int}}[\phi(x)].$$

For now we restrict ourselves to theories in which $\mathcal{H}_{\text{int}}$ $(= -\mathcal{L}_{\text{int}})$ is a function only of the fields, not of their derivatives.

In this chapter we will discuss three important examples of interacting field theories. The first is "phi-fourth" theory,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4, \tag{4.1}$$

where λ is a dimensionless *coupling constant.* (A ϕ^3 interaction would be a bit simpler, but then the energy would not be positive-definite unless we added a higher even power of ϕ as well.) Although we are introducing this theory now for purely pedagogical reasons (since it is the simplest of all interacting quantum theories), models of the real world do contain ϕ^4 interactions; the most important example in particle physics is the self-interaction of the Higgs field in the standard electroweak theory. In Part II, we will see that ϕ^4 theory also arises in statistical mechanics. The equation of motion for ϕ^4 theory is

$$(\partial^2 + m^2)\phi = -\frac{\lambda}{3!}\phi^3, \tag{4.2}$$

which cannot be solved by Fourier analysis as the free Klein-Gordon equation could. In the quantum theory we impose the equal-time commutation relations

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

which are unaffected by $\mathcal{L}_{\text{int}}$. (Note, however, that if $\mathcal{L}_{\text{int}}$ contained $\partial_\mu \phi$, the definition of $\pi(\mathbf{x})$ would change.) It is an easy exercise to write down the Hamiltonian of this theory and find the Heisenberg equation of motion for the operator $\phi(x)$; the result is the same as the classical equation of motion (4.2), just as it was in the free theory.

Our second example of an interacting field theory will be Quantum Electrodynamics:

$$\begin{aligned} \mathcal{L}_{\text{QED}} &= \mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{Maxwell}} + \mathcal{L}_{\text{int}} \\ &= \bar{\psi}(i\partial\!\!\!/ - m)\psi - \tfrac{1}{4}(F_{\mu\nu})^2 - e\bar{\psi}\gamma^\mu\psi A_\mu, \end{aligned} \tag{4.3}$$

where A_μ is the electromagnetic vector potential, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor, and $e = -|e|$ is the electron charge. (To describe a fermion of charge Q, replace e with Q. If we wish to consider several species of charged particles at once, we simply duplicate $\mathcal{L}_{\text{Dirac}}$ and $\mathcal{L}_{\text{int}}$ for each additional species.) That such a simple Lagrangian can account for nearly all observed phenomena from macroscopic scales down to 10^{-13} cm is rather astonishing. In fact, the QED Lagrangian can be written even more simply:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(iD\!\!\!/ - m)\psi - \tfrac{1}{4}(F_{\mu\nu})^2, \tag{4.4}$$

where D_μ is the *gauge covariant derivative*,

$$D_\mu \equiv \partial_\mu + ieA_\mu(x). \tag{4.5}$$

A crucial property of the QED Lagrangian is that it is invariant under the gauge transformation

$$\psi(x) \to e^{i\alpha(x)}\psi(x), \qquad A_\mu \to A_\mu - \frac{1}{e}\partial_\mu \alpha(x), \tag{4.6}$$

which is realized on the Dirac field as a *local* phase rotation. This invariance under local phase rotations has a fundamental geometrical significance, which motivates the term *covariant derivative*. For our present purposes, though, it is sufficient just to recognize (4.6) as a symmetry of the theory.

The equations of motion follow from (4.3) by the canonical procedure. The Euler-Lagrange equation for ψ is

$$(iD\!\!\!/ - m)\psi(x) = 0, \tag{4.7}$$

which is just the Dirac equation coupled to the electromagnetic field by the *minimal coupling* prescription, $\partial \to D$. The Euler-Lagrange equation for A_ν is

$$\partial_\mu F^{\mu\nu} = e\bar{\psi}\gamma^\nu\psi = ej^\nu. \tag{4.8}$$

These are the inhomogeneous Maxwell equations, with the current density $j^\nu = \bar{\psi}\gamma^\nu\psi$ given by the conserved Dirac vector current (3.73). As with ϕ^4 theory, the equations of motion can also be obtained as the Heisenberg equations of motion for the operators $\psi(x)$ and $A_\mu(x)$. This is easy to verify for $\psi(x)$; we have not yet discussed the quantization of the electromagnetic field.

In fact, we will not discuss canonical quantization of the electromagnetic field at all in this book. It is an awkward subject, essentially because of gauge invariance. Note that since $\dot{A}^0$ does not appear in the Lagrangian (4.3), the momentum conjugate to A^0 is identically zero. This contradicts the canonical commutation relation $[A^0(\mathbf{x}), \pi^0(\mathbf{y})] = i\delta(\mathbf{x} - \mathbf{y})$. One solution is to quantize in Coulomb gauge, where $\nabla \cdot \mathbf{A} = 0$ and A^0 is a constrained, rather than dynamical, variable; but then manifest Lorentz invariance is sacrificed. Alternatively, one can quantize the field in Lorentz gauge, $\partial_\mu A^\mu = 0$. It is then possible to modify the Lagrangian, adding an $\dot{A}^0$ term. One obtains the commutation relations $[A^\mu(\mathbf{x}), \dot{A}^\nu(\mathbf{y})] = -ig^{\mu\nu}\delta(\mathbf{x} - \mathbf{y})$, essentially the same as four Klein-Gordon fields. But the extra minus sign in $[A^0, \dot{A}^0]$ leads to another (surmountable) difficulty: states created by $a^{0\dagger}_{\mathbf{p}}$ have negative norm.*

The Feynman rules for calculating scattering amplitudes that involve photons are derived more easily in the functional integral formulation of field theory, to be discussed in Chapter 9. That method has the added advantage of generalizing readily to the case of non-Abelian gauge fields, as we will see in Part III. In the present chapter we will simply guess the Feynman rules for photons. This will actually be quite easy after we derive the rules for an analogous but simpler theory, *Yukawa theory*:

$$\mathcal{L}_{\text{Yukawa}} = \mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{Klein-Gordon}} - g\bar{\psi}\psi\phi. \tag{4.9}$$

This will be our third example. It is similar to QED, but with the photon replaced by a scalar particle ϕ. The interaction term contains a dimensionless coupling constant g, analogous to the electron charge e. Yukawa originally invented this theory to describe nucleons (ψ) and pions (ϕ). In modern particle theory, the Standard Model contains Yukawa interaction terms coupling the scalar Higgs field to quarks and leptons; most of the free parameters in the Standard Model are Yukawa coupling constants.

Having written down our three paradigm interactions, let us pause a moment to discuss what other interactions could be found in Nature. At first it might seem that the list would be infinite; even for a scalar theory we could write down interactions of the form ϕ^n for any n. But remarkably, one simple and reasonable axiom eliminates all but a few of the possible interactions. That axiom is that the theory be *renormalizable*, and it arises as follows. Higher-order terms in perturbation theory, as mentioned in Chapter 1, will involve

*Excellent treatments of both quantization procedures are readily available. For Coulomb gauge quantization, see Bjorken and Drell (1965), Chapter 14; for Lorentz gauge quantization, see Mandl and Shaw (1984), Chapter 5.

integrals over the 4-momenta of intermediate ("virtual") particles. These integrals are often formally divergent, and it is generally necessary to impose some form of cut-off procedure; the simplest is just to cut off the integral at some large but finite momentum Λ. At the end of the calculation one takes the limit $\Lambda \to \infty$, and hopes that physical quantities turn out to be independent of Λ. If this is indeed the case, the theory is said to be *renormalizable*. Suppose, however, that the theory includes interactions whose coupling constants have the dimensions of mass to some *negative* power. Then to obtain a dimensionless scattering amplitude, this coupling constant must be multiplied by some quantity of positive mass dimension, and it turns out that this quantity is none other than Λ. Such a term diverges as $\Lambda \to \infty$, so the theory is not renormalizable.

We will discuss these matters in detail in Chapter 10. For now we merely note that any theory containing a coupling constant with negative mass dimension is not renormalizable. A bit of dimensional analysis then allows us to throw out nearly all candidate interactions. Since the action $S = \int \mathcal{L}\, d^4x$ is dimensionless, $\mathcal{L}$ must have dimension $(\text{mass})^4$ (or simply dimension 4). From the kinetic terms of the various free Lagrangians, we note that the scalar and vector fields ϕ and A^μ have dimension 1, while the spinor field ψ has dimension $3/2$. We can now tabulate all of the allowed renormalizable interactions.

For theories involving only scalars, the allowed interaction terms are

$$\mu\phi^3 \qquad \text{and} \qquad \lambda\phi^4.$$

The coupling constant μ has dimension 1, while λ is dimensionless. Terms of the form ϕ^n for $n > 4$ are not allowed, since their coupling constants would have dimension $4 - n$. Of course, more interesting theories can be obtained by including several scalar fields, real or complex (see Problem 4.3).

Next we can add spinor fields. Spinor self-interactions are not allowed, since ψ^3 (besides violating Lorentz invariance) already has dimension $9/2$. Thus the only allowable new interaction is the Yukawa term,

$$g\bar{\psi}\psi\phi,$$

although similar interactions can also be constructed out of Weyl and Majorana spinors.

When we add vector fields, many new interactions are possible. The most familiar is the vector-spinor interaction of QED,

$$e\bar{\psi}\gamma^\mu\psi A_\mu.$$

Again it is easy to construct similar terms out of Weyl and Majorana spinors. Less important is the *scalar QED* Lagrangian,

$$\mathcal{L} = |D_\mu\phi|^2 - m^2|\phi|^2, \qquad \text{which contains} \quad eA^\mu\phi\partial_\mu\phi^*,\ e^2|\phi|^2A^2.$$

This is our first example of a derivative interaction; quantization of this theory will be much easier with the functional integral formalism, so we postpone its

discussion until Chapter 9. Other possible Lorentz-invariant terms involving vectors are

$$A^2(\partial_\mu A^\mu) \qquad \text{and} \qquad A^4.$$

Although it is far from obvious, these terms lead to inconsistencies unless their coupling constants are precisely chosen on the basis of a special type of symmetry, which must involve several vector fields. This symmetry underlies the *non-Abelian gauge theories*, which will be the main subject of Part III. A mass term $\frac{1}{2}m^2A^2$ for vector fields is also inconsistent, except in the special case where it is added to QED; in any case, it breaks (Abelian or non-Abelian) gauge invariance.

This exhausts the list of possible Lagrangians involving scalar, spinor, and vector particles. It is interesting to note that the currently accepted models of the strong, weak, and electromagnetic interactions include *all* of the types of interactions listed above. The three paradigm interactions to be studied in this chapter cover nearly half of the possibilities; we will study the others in detail later in this book.

The assumption that realistic theories must be renormalizable is certainly convenient, since a nonrenormalizable theory would have little predictive power. However, one might still ask *why* Nature has been so kind as to use only renormalizable interactions. One might have expected that the true theory of Nature would be a quantum theory of a much more general type. But it can be shown that, however complicated a fundamental theory appears at very high energies, the low-energy approximation to this theory that we see in experiments should be a renormalizable quantum field theory. We will demonstrate this in Section 12.1.

At a more practical level, the preceding analysis highlights a great difference in methodology between nonrelativistic quantum mechanics and relativistic quantum field theory. Since the potential $V(\mathbf{r})$ that appears in the Schrödinger equation is completely arbitrary, nonrelativistic quantum mechanics puts no limits on what interactions can be found in the real world. But we have just seen that quantum field theory imposes very tight constraints on Nature (or vice versa). Taken literally, our discussion implies that the only tasks left for particle physicists are to enumerate the elementary particles that exist and to measure their masses and coupling constants. While this viewpoint is perhaps overly arrogant, the fact that it is even thinkable is surely a sign that particle physicists are on the right track toward a fundamental theory.

Given a set of particles and couplings, we must still work out the experimental consequences. How do we analyze the quantum mechanics of an interacting field theory? It would be nice if we could explicitly solve at least a few examples (that is, find the exact eigenvalues and eigenvectors as we did for the free theories) to get a feel for the properties of interacting theories. Unfortunately, this is easier said than done. No exactly solvable interacting field theories are known in more than two spacetime dimensions, and even

there the solvable models involve special symmetries and considerable technical complication.† Studying these theories would be interesting, but hardly worth the effort at this stage. Instead we will fall back on a much simpler and more generally applicable approach: Treat the interaction term H_{int} as a perturbation, compute its effects as far in perturbation theory as is practicable, and hope that the coupling constant is small enough that this gives a reasonable approximation to the exact answer. In fact, the perturbation series we obtain will turn out to be very simple in structure; through the use of *Feynman diagrams* it will be possible at least to visualize the effects of interactions to arbitrarily high order.

This simplification of the perturbation series for relativistic field theories was the great advance of Tomonaga, Schwinger, and Feynman. To achieve this simplification, each, independently, found a way to reformulate quantum mechanics to remove the special role of time, and then applied his new viewpoint to recast each term of the perturbation expansion as a spacetime process. We will develop quantum field theory from a spacetime viewpoint, using Feynman's method of *functional integration*, in Chapter 9. In the present chapter we follow a more pedestrian line of analysis, first developed by Dyson, to derive the spacetime picture of perturbation theory from the conventional machinery of quantum mechanics.‡

4.2 Perturbation Expansion of Correlation Functions

Let us then begin the study of perturbation theory for interacting fields, aiming toward a formalism that will allow us to visualize the perturbation series as spacetime processes. Although we will not need to reformulate quantum mechanics, we will rederive time-dependent perturbation theory in a form that is convenient for our purposes. Ultimately, of course, we want to calculate scattering cross sections and decay rates. For now, however, let us be less ambitious and try to calculate a simpler (but more abstract) quantity, the *two-point correlation function*, or *two-point Green's function*,

$$\langle \Omega | T\phi(x)\phi(y) | \Omega \rangle , \tag{4.10}$$

in ϕ^4 theory. We introduce the notation $|\Omega\rangle$ to denote the ground state of the interacting theory, which is generally different from $|0\rangle$, the ground state of the free theory. The time-ordering symbol T is inserted for later convenience. The correlation function can be interpreted physically as the amplitude for propagation of a particle or excitation between y and x. In the free theory, it

†A brief survey of exactly solvable quantum field theories is given in the Epilogue.

‡For a historical account of the contributions of Tomonaga, Schwinger, Feynman, and Dyson, see Schweber (1994).

is simply the Feynman propagator:

$$\langle 0|\, T\phi(x)\phi(y)\, |0\rangle_{\text{free}} = D_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i\, e^{-ip\cdot(x-y)}}{p^2 - m^2 + i\epsilon}. \tag{4.11}$$

We would like to know how this expression changes in the interacting theory. Once we have analyzed the two-point correlation function, it will be easy to generalize our results to higher correlation functions in which more than two field operators appear. In Sections 4.3 and 4.4 we will continue the analysis of correlation functions, eventually developing the formalism of Feynman diagrams for evaluating them perturbatively. Then in Sections 4.5 and 4.6 we will learn how to calculate cross sections and decay rates using the same techniques.

To attack this problem, we write the Hamiltonian of ϕ^4 theory as

$$H = H_0 + H_{\text{int}} = H_{\text{Klein-Gordon}} + \int d^3x\, \frac{\lambda}{4!}\phi^4(\mathbf{x}). \tag{4.12}$$

We want an expression for the two-point correlation function (4.10) as a power series in λ. The interaction Hamiltonian H_{int} enters (4.10) in two places: first, in the definition of the Heisenberg field,

$$\phi(x) = e^{iHt}\phi(\mathbf{x})e^{-iHt}; \tag{4.13}$$

and second, in the definition of $|\Omega\rangle$. We must express both $\phi(x)$ and $|\Omega\rangle$ in terms of quantities we know how to manipulate: free field operators and the free theory vacuum $|0\rangle$.

It is easiest to begin with $\phi(x)$. At any fixed time t_0, we can of course expand ϕ as before in terms of ladder operators:

$$\phi(t_0, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} + a^\dagger_{\mathbf{p}} e^{-i\mathbf{p}\cdot\mathbf{x}}\right).$$

Then to obtain $\phi(t, \mathbf{x})$ for $t \neq t_0$, we just switch to the Heisenberg picture as usual:

$$\phi(t, \mathbf{x}) = e^{iH(t-t_0)}\phi(t_0, \mathbf{x})e^{-iH(t-t_0)}.$$

For $\lambda = 0$, H becomes H_0 and this reduces to

$$\phi(t, \mathbf{x})\big|_{\lambda=0} = e^{iH_0(t-t_0)}\phi(t_0, \mathbf{x})e^{-iH_0(t-t_0)} \equiv \phi_I(t, \mathbf{x}). \tag{4.14}$$

When λ is small, this expression will still give the most important part of the time dependence of $\phi(x)$, and thus it is convenient to give this quantity a name: the *interaction picture* field, $\phi_I(t, \mathbf{x})$. Since we can diagonalize H_0, it is easy to construct ϕ_I explicitly:

$$\phi_I(t, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{-ip\cdot x} + a^\dagger_{\mathbf{p}} e^{ip\cdot x}\right)\Big|_{x^0 = t - t_0}. \tag{4.15}$$

This is just the familiar expression for the free field from Chapter 2.

The problem now is to express the full Heisenberg field ϕ in terms of ϕ_I. Formally, it is just

$$\begin{aligned}\phi(t,\mathbf{x}) &= e^{iH(t-t_0)}e^{-iH_0(t-t_0)}\phi_I(t,\mathbf{x})e^{iH_0(t-t_0)}e^{-iH(t-t_0)} \\ &\equiv U^\dagger(t,t_0)\phi_I(t,\mathbf{x})U(t,t_0),\end{aligned} \tag{4.16}$$

where we have defined the unitary operator

$$U(t,t_0) = e^{iH_0(t-t_0)}e^{-iH(t-t_0)}, \tag{4.17}$$

known as the interaction picture propagator or time-evolution operator. We would like to express $U(t,t_0)$ entirely in terms of ϕ_I, for which we have an explicit expression in terms of ladder operators. To do this, we note that $U(t,t_0)$ is the unique solution, with initial condition $U(t_0,t_0)=1$, of a simple differential equation (the Schrödinger equation):

$$\begin{aligned}i\frac{\partial}{\partial t}U(t,t_0) &= e^{iH_0(t-t_0)}\big(H-H_0\big)e^{-iH(t-t_0)} \\ &= e^{iH_0(t-t_0)}\big(H_{\text{int}}\big)e^{-iH(t-t_0)} \\ &= e^{iH_0(t-t_0)}\big(H_{\text{int}}\big)e^{-iH_0(t-t_0)}e^{iH_0(t-t_0)}e^{-iH(t-t_0)} \\ &= H_I(t)\,U(t,t_0),\end{aligned} \tag{4.18}$$

where

$$H_I(t) = e^{iH_0(t-t_0)}\big(H_{\text{int}}\big)e^{-iH_0(t-t_0)} = \int d^3x\,\frac{\lambda}{4!}\phi_I^4 \tag{4.19}$$

is the interaction Hamiltonian written in the interaction picture. The solution of this differential equation for $U(t,t_0)$ should look something like $U \sim \exp(-iH_I t)$; this would be our desired formula for U in terms of ϕ_I. Doing it more carefully, we will show that the actual solution is the following power series in λ:

$$\begin{aligned}U(t,t_0) = 1 &+ (-i)\int_{t_0}^{t} dt_1\, H_I(t_1) + (-i)^2\int_{t_0}^{t}dt_1\int_{t_0}^{t_1}dt_2\, H_I(t_1)H_I(t_2) \\ &+ (-i)^3\int_{t_0}^{t}dt_1\int_{t_0}^{t_1}dt_2\int_{t_0}^{t_2}dt_3\, H_I(t_1)H_I(t_2)H_I(t_3) + \cdots.\end{aligned} \tag{4.20}$$

To verify this, just differentiate: Each term gives the previous one times $-iH_I(t)$. The initial condition $U(t,t_0)=1$ for $t=t_0$ is obviously satisfied.

Note that the various factors of H_I in (4.20) stand in *time order*, later on the left. This allows us to simplify the expression considerably, using the time-ordering symbol T. The H_I^2 term, for example, can be written

$$\int_{t_0}^{t}dt_1\int_{t_0}^{t_1}dt_2\, H_I(t_1)H_I(t_2) = \frac{1}{2}\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2\; T\big\{H_I(t_1)H_I(t_2)\big\}. \tag{4.21}$$

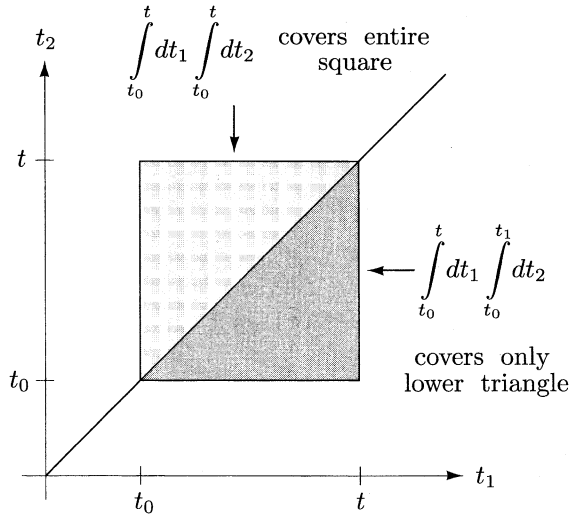


Figure 4.1. Geometric interpretation of Eq. (4.21).

The double integral on the right-hand side just counts everything twice, since in the t_1t_2-plane, the integrand $T\{H_I(t_1)H_I(t_2)\}$ is symmetric about the line $t_1 = t_2$ (see Fig. 4.1).

A similar identity holds for the higher terms:

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n \, H_I(t_1) \cdots H_I(t_n) = \frac{1}{n!} \int_{t_0}^{t} dt_1 \cdots dt_n \, T\{H_I(t_1) \cdots H_I(t_n)\}.$$

This case is a little harder to visualize, but it is not hard to convince oneself that it is true. Using this identity, we can now write $U(t, t_0)$ in an extremely compact form:

$$\begin{aligned} U(t, t_0) &= 1 + (-i) \int_{t_0}^{t} dt_1 \, H_I(t_1) + \frac{(-i)^2}{2!} \int_{t_0}^{t} dt_1 \, dt_2 \, T\{H_I(t_1) H_I(t_2)\} + \cdots \\ &\equiv T\left\{ \exp\left[-i \int_{t_0}^{t} dt' \, H_I(t') \right] \right\}, \end{aligned} \tag{4.22}$$

where the time-ordering of the exponential is just defined as the Taylor series with each term time-ordered. When we do real computations we will keep only the first few terms of the series; the time-ordered exponential is just a compact way of writing and remembering the correct expression.

We now have control over $\phi(t, \mathbf{x})$; we have written it entirely in terms of ϕ_I, as desired. Before moving on to consider $|\Omega\rangle$, however, it is convenient to generalize the definition of U, allowing its second argument to take on values

other than our "reference time" t_0. The correct definition is quite natural:

$$U(t,t') \equiv T\left\{\exp\left[-i\int_{t'}^{t} dt''\, H_I(t'')\right]\right\}. \qquad (t \geq t') \tag{4.23}$$

Several properties follow from this definition, and it is necessary to verify them. First, $U(t,t')$ satisfies the same differential equation (4.18),

$$i\frac{\partial}{\partial t}U(t,t') = H_I(t)U(t,t'), \tag{4.24}$$

but now with the initial condition $U = 1$ for $t = t'$. From this equation you can show that

$$U(t,t') = e^{iH_0(t-t_0)}e^{-iH(t-t')}e^{-iH_0(t'-t_0)}, \tag{4.25}$$

which proves that U is unitary. Finally, $U(t,t')$ satisfies the following identities (for $t_1 \geq t_2 \geq t_3$):

$$\begin{aligned} U(t_1,t_2)U(t_2,t_3) &= U(t_1,t_3); \\ U(t_1,t_3)\left[U(t_2,t_3)\right]^\dagger &= U(t_1,t_2). \end{aligned} \tag{4.26}$$

Now we can go on to discuss $|\Omega\rangle$. Since $|\Omega\rangle$ is the ground state of H, we can isolate it by the following procedure. Imagine starting with $|0\rangle$, the ground state of H_0, and evolving through time with H:

$$e^{-iHT}|0\rangle = \sum_n e^{-iE_nT}|n\rangle\langle n|0\rangle,$$

where E_n are the eigenvalues of H. We must assume that $|\Omega\rangle$ has some overlap with $|0\rangle$, that is, $\langle\Omega|0\rangle \neq 0$ (if this were not the case, H_I would in no sense be a small perturbation). Then the above series contains $|\Omega\rangle$, and we can write

$$e^{-iHT}|0\rangle = e^{-iE_0T}|\Omega\rangle\langle\Omega|0\rangle + \sum_{n\neq 0} e^{-iE_nT}|n\rangle\langle n|0\rangle,$$

where $E_0 \equiv \langle\Omega|H|\Omega\rangle$. (The zero of energy will be defined by $H_0|0\rangle = 0$.) Since $E_n > E_0$ for all $n \neq 0$, we can get rid of all the $n \neq 0$ terms in the series by sending T to ∞ in a slightly imaginary direction: $T \to \infty(1-i\epsilon)$. Then the exponential factor e^{-iE_nT} dies slowest for $n = 0$, and we have

$$|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)}\left(e^{-iE_0T}\langle\Omega|0\rangle\right)^{-1}e^{-iHT}|0\rangle. \tag{4.27}$$

Since T is now very large, we can shift it by a small constant:

$$\begin{aligned} |\Omega\rangle &= \lim_{T\to\infty(1-i\epsilon)}\left(e^{-iE_0(T+t_0)}\langle\Omega|0\rangle\right)^{-1}e^{-iH(T+t_0)}|0\rangle \\ &= \lim_{T\to\infty(1-i\epsilon)}\left(e^{-iE_0(t_0-(-T))}\langle\Omega|0\rangle\right)^{-1}e^{-iH(t_0-(-T))}e^{-iH_0(-T-t_0)}|0\rangle \\ &= \lim_{T\to\infty(1-i\epsilon)}\left(e^{-iE_0(t_0-(-T))}\langle\Omega|0\rangle\right)^{-1}U(t_0,-T)|0\rangle. \end{aligned} \tag{4.28}$$

In the second line we have used $H_0 |0\rangle = 0$. Ignoring the c-number factor in front, this expression tells us that we can get $|\Omega\rangle$ by simply evolving $|0\rangle$ from time $-T$ to time t_0 with the operator U. Similarly, we can express $\langle\Omega|$ as

$$\langle\Omega| = \lim_{T\to\infty(1-i\epsilon)} \langle 0| U(T,t_0) \big(e^{-iE_0(T-t_0)} \langle 0|\Omega\rangle\big)^{-1}. \tag{4.29}$$

Let us put together the pieces of the two-point correlation function. For the moment, assume that $x^0 > y^0 > t_0$. Then

$$\begin{aligned}\langle\Omega|\,\phi(x)\phi(y)\,|\Omega\rangle = &\lim_{T\to\infty(1-i\epsilon)} \big(e^{-iE_0(T-t_0)} \langle 0|\Omega\rangle\big)^{-1} \langle 0| U(T,t_0) \\ &\times \big[U(x^0,t_0)\big]^\dagger \phi_I(x) U(x^0,t_0) \big[U(y^0,t_0)\big]^\dagger \phi_I(y) U(y^0,t_0) \\ &\times U(t_0,-T)\,|0\rangle \big(e^{-iE_0(t_0-(-T))} \langle\Omega|\,0\rangle\big)^{-1} \\ = &\lim_{T\to\infty(1-i\epsilon)} \big(|\langle 0|\,\Omega\rangle\,|^2 e^{-iE_0(2T)}\big)^{-1} \\ &\times \langle 0| U(T,x^0)\phi_I(x)U(x^0,y^0)\phi_I(y)U(y^0,-T)\,|0\rangle\,. \end{aligned} \tag{4.30}$$

This is starting to look simple, except for the awkward factor in front. To get rid of it, divide by 1 in the form

$$1 = \langle\Omega|\,\Omega\rangle = \big(|\langle 0|\,\Omega\rangle\,|^2 e^{-iE_0(2T)}\big)^{-1} \langle 0| U(T,t_0)U(t_0,-T)\,|0\rangle\,.$$

Then our formula, still for $x^0 > y^0$, becomes

$$\langle\Omega|\,\phi(x)\phi(y)\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0| U(T,x^0)\phi_I(x)U(x^0,y^0)\phi_I(y)U(y^0,-T)\,|0\rangle}{\langle 0| U(T,-T)\,|0\rangle}.$$

Now note that all fields on both sides of this expression are in time order. If we had considered the case $y^0 > x^0$ this would still be true. Thus we arrive at our final expression, now valid for any x^0 and y^0:

$$\langle\Omega|\,T\big\{\phi(x)\phi(y)\big\}\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0|\,T\Big\{\phi_I(x)\phi_I(y)\exp\big[-i\int_{-T}^{T} dt\, H_I(t)\big]\Big\}\,|0\rangle}{\langle 0|\,T\Big\{\exp\big[-i\int_{-T}^{T} dt\, H_I(t)\big]\Big\}\,|0\rangle}. \tag{4.31}$$

The virtue of considering the time-ordered product is clear: It allows us to put everything inside one large T-operator. A similar formula holds for higher correlation functions of arbitrarily many fields; for each extra factor of ϕ on the left, put an extra factor of ϕ_I on the right. So far this expression is exact. But it is ideally suited to doing perturbative calculations; we need only retain as many terms as desired in the Taylor series expansions of the exponentials.

4.3 Wick's Theorem

We have now reduced the problem of calculating correlation functions to that of evaluating expressions of the form

$$\langle 0|\, T\{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)\}\,|0\rangle\,,$$

that is, vacuum expectation values of time-ordered products of finite (but arbitrary) numbers of free field operators. For $n = 2$ this expression is just the Feynman propagator. For higher n you could evaluate this object by brute force, plugging in the expansion of ϕ_I in terms of ladder operators. In this section and the next, however, we will see how to simplify such calculations immensely.

Consider again the case of two fields, $\langle 0|\, T\{\phi_I(x)\phi_I(y)\}\,|0\rangle$. We already know how to calculate this quantity, but now we would like to rewrite it in a form that is easy to evaluate and also generalizes to the case of more than two fields. To do this we first decompose $\phi_I(x)$ into positive- and negative-frequency parts:

$$\phi_I(x) = \phi_I^+(x) + \phi_I^-(x), \tag{4.32}$$

where

$$\phi_I^+(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\, a_{\mathbf{p}} e^{-ip\cdot x}; \qquad \phi_I^-(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\, a_{\mathbf{p}}^\dagger e^{+ip\cdot x}.$$

This decomposition can be done for any free field. It is useful because

$$\phi_I^+(x)\,|0\rangle = 0 \qquad \text{and} \qquad \langle 0|\,\phi_I^-(x) = 0.$$

For example, consider the case $x^0 > y^0$. The time-ordered product of two fields is then

$$\begin{aligned} T\phi_I(x)\phi_I(y) &\underset{x^0>y^0}{=} \phi_I^+(x)\phi_I^+(y) + \phi_I^+(x)\phi_I^-(y) + \phi_I^-(x)\phi_I^+(y) + \phi_I^-(x)\phi_I^-(y) \\ &= \phi_I^+(x)\phi_I^+(y) + \phi_I^-(y)\phi_I^+(x) + \phi_I^-(x)\phi_I^+(y) + \phi_I^-(x)\phi_I^-(y) \\ &\quad + \big[\phi_I^+(x), \phi_I^-(y)\big]. \end{aligned} \tag{4.33}$$

In every term except the commutator, all the $a_{\mathbf{p}}$'s are to the right of all the $a_{\mathbf{p}}^\dagger$'s. Such a term (e.g., $a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger a_{\mathbf{k}} a_{\mathbf{l}}$) is said to be in *normal order*, and has vanishing vacuum expectation value. Let us also define the normal ordering symbol $N()$ to place whatever operators it contains in normal order, for example,

$$N\big(a_{\mathbf{p}} a_{\mathbf{k}}^\dagger a_{\mathbf{q}}\big) \equiv a_{\mathbf{k}}^\dagger a_{\mathbf{p}} a_{\mathbf{q}}. \tag{4.34}$$

The order of $a_{\mathbf{p}}$ and $a_{\mathbf{q}}$ on the right-hand side makes no difference since they commute.*

*In the literature one often sees the notation $:\phi_1\phi_2:$ instead of $N(\phi_1\phi_2)$.

If we had instead considered the case $y^0 > x^0$, we would get the same four normal-ordered terms as in (4.33), but this time the final commutator would be $[\phi_I^+(y), \phi_I^-(x)]$. Let us therefore define one more quantity, the *contraction* of two fields, as follows:

$$\overbracket{\phi(x)\phi(y)} \equiv \begin{cases} [\phi^+(x), \phi^-(y)] & \text{for } x^0 > y^0; \\ [\phi^+(y), \phi^-(x)] & \text{for } y^0 > x^0. \end{cases} \tag{4.35}$$

This quantity is exactly the Feynman propagator:

$$\overbracket{\phi(x)\phi(y)} = D_F(x - y). \tag{4.36}$$

(From here on we will often drop the subscript I for convenience; contractions will always involve interaction-picture fields.)

The relation between time-ordering and normal-ordering is now extremely simple to express, at least for two fields:

$$T\{\phi(x)\phi(y)\} = N\{\phi(x)\phi(y) + \overbracket{\phi(x)\phi(y)}\}. \tag{4.37}$$

But now that we have all this new notation, the generalization to arbitrarily many fields is also easy to write down:

$$\begin{aligned} &T\{\phi(x_1)\phi(x_2)\cdots\phi(x_m)\} \\ &\quad = N\{\phi(x_1)\phi(x_2)\cdots\phi(x_m) + \text{all possible contractions}\}. \end{aligned} \tag{4.38}$$

This identity is known as *Wick's theorem*, and we will prove it in a moment. For $m = 2$ it is identical to (4.37). The phrase *all possible contractions* means there will be one term for each possible way of contracting the m fields in pairs. Thus for $m = 4$ we have (writing $\phi(x_a)$ as ϕ_a for brevity)

$$\begin{aligned} T\{\phi_1\phi_2\phi_3\phi_4\} = N\{&\phi_1\phi_2\phi_3\phi_4 + \overbracket{\phi_1\phi_2}\phi_3\phi_4 + \overbracket{\phi_1\phi_2\phi_3}\phi_4 + \overbracket{\phi_1\phi_2\phi_3\phi_4} \\ &+ \phi_1\overbracket{\phi_2\phi_3}\phi_4 + \phi_1\overbracket{\phi_2\phi_3\phi_4} + \phi_1\phi_2\overbracket{\phi_3\phi_4} \\ &+ \overbracket{\phi_1\phi_2}\overbracket{\phi_3\phi_4} + \overbracket{\phi_1\overbracket{\phi_2\phi_3}\phi_4} + \overbracket{\phi_1\phi_2\overbracket{\phi_3\phi_4}}\}. \end{aligned} \tag{4.39}$$

When the contraction symbol connects two operators that are not adjacent, we still define it to give a factor of D_F. For example,

$$N\{\overbracket{\phi_1\phi_2}\phi_3\phi_4\} \qquad \text{means} \qquad D_F(x_1 - x_3) \cdot N\{\phi_2\phi_4\}.$$

In the vacuum expectation value of (4.39), any term in which there remain uncontracted operators gives zero (since $\langle 0|\, N(\text{any operator})\, |0\rangle = 0$). Only the three fully contracted terms in the last line survive, and they are all c-numbers, so we have

$$\begin{aligned} \langle 0|\, T\{\phi_1\phi_2\phi_3\phi_4\}\, |0\rangle = &\ D_F(x_1 - x_2)D_F(x_3 - x_4) \\ &+ D_F(x_1 - x_3)D_F(x_2 - x_4) \\ &+ D_F(x_1 - x_4)D_F(x_2 - x_3). \end{aligned} \tag{4.40}$$

Now let us prove Wick's theorem. Naturally the proof is by induction on m, the number of fields. We have already proved the case $m = 2$. So assume the theorem is true for $m - 1$ fields, and let's try to prove it for m fields. Without loss of generality, we can restrict ourselves to the case $x_1^0 \geq x_2^0 \geq \cdots x_m^0$; if this is not the case we can just relabel the points, without affecting either side of (4.38). Then applying Wick's theorem to $\phi_2 \cdots \phi_m$, we have

$$\begin{aligned} T\{\phi_1 \cdots \phi_m\} &= \phi_1 \cdots \phi_m \\ &= \phi_1 N\Big\{\phi_2 \cdots \phi_m + \begin{pmatrix}\text{all contractions}\\ \text{not involving } \phi_1\end{pmatrix}\Big\} \\ &= (\phi_1^+ + \phi_1^-)\, N\Big\{\phi_2 \cdots \phi_m + \begin{pmatrix}\text{all contractions}\\ \text{not involving } \phi_1\end{pmatrix}\Big\}. \end{aligned} \tag{4.41}$$

We want to move the ϕ_1^+ and ϕ_1^- inside the $N\{\}$. For the ϕ_1^- term this is easy: Just move it in, since (being on the left) it is already in normal order. The term with ϕ_1^+ must be put in normal order by commuting it to the right past all the other ϕ's. Consider, for example, the term with no contractions:

$$\begin{aligned} \phi_1^+ N(\phi_2 \cdots \phi_m) &= N(\phi_2 \cdots \phi_m)\phi_1^+ + [\phi_1^+, N(\phi_2 \cdots \phi_m)] \\ &= N(\phi_1^+ \phi_2 \cdots \phi_m) \\ &\quad + N\big([\phi_1^+, \phi_2^-]\phi_3 \cdots \phi_m + \phi_2[\phi_1^+, \phi_3^-]\phi_4 \cdots \phi_m + \cdots\big) \\ &= N\big(\phi_1^+ \phi_2 \cdots \phi_m + \overbrace{\phi_1 \phi_2}\phi_3 \cdots \phi_m + \overbrace{\phi_1 \phi_2 \phi_3} \cdots + \cdots\big). \end{aligned}$$

The first term in the last line combines with part of the ϕ_1^- term from (4.41) to give $N\{\phi_1\phi_2 \cdots \phi_m\}$, so we now have the first term on the right-hand side of Wick's theorem, as well as all possible terms involving a single contraction of ϕ_1 with another field. Similarly, a term in (4.41) involving one contraction will produce all possible terms involving both that contraction and a contraction of ϕ_1 with one of the other fields. Doing this with all the terms of (4.41), we eventually get all possible contractions of all the fields, including ϕ_1. Thus the induction step is complete, and Wick's theorem is proved.

4.4 Feynman Diagrams

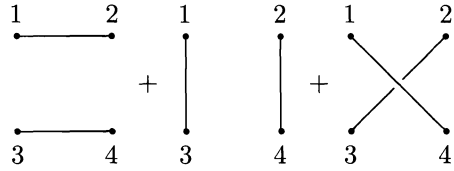
Wick's theorem allows us to turn any expression of the form

$$\langle 0|\, T\{\phi_I(x_1)\phi_I(x_2) \cdots \phi_I(x_n)\}\, |0\rangle$$

into a sum of products of Feynman propagators. Now we are ready to develop a diagrammatic interpretation of such expressions. Consider first the case of four fields, all at different spacetime points, which we worked out in Eq. (4.40). Let us represent each of the points x_1 through x_4 by a dot, and each factor

$D_F(x-y)$ by a line joining x to y. Then Eq. (4.40) can be represented as the sum of three diagrams (called *Feynman diagrams*):

$$\langle 0|\, T\{\phi_1\phi_2\phi_3\phi_4\}\,|0\rangle = \quad (4.42)$$



Although this isn't exactly a measurable quantity, the diagrams do suggest an interpretation: Particles are created at two spacetime points, each propagates to one of the other points, and then they are annihilated. This can happen in three ways, corresponding to the three ways to connect the points in pairs, as shown in the three diagrams. The total amplitude for the process is the sum of the three diagrams.

Things get more interesting when the expression contains more than one field at the same spacetime point. So let us now return to the evaluation of the two-point function $\langle\Omega|\, T\{\phi(x)\phi(y)\}\,|\Omega\rangle$, and put formula (4.31) to use. We will ignore the denominator until the very end of this section. The numerator, with the exponential expanded as a power series, is

$$\langle 0|\, T\Big\{\phi(x)\phi(y) \;+\; \phi(x)\phi(y)\Big[-i\int dt\, H_I(t)\Big] \;+\cdots\Big\}\,|0\rangle\,. \quad (4.43)$$

The first term gives the free-field result, $\langle 0|\, T\{\phi(x)\phi(y)\}\,|0\rangle = D_F(x-y)$. The second term, in ϕ^4 theory, is

$$\begin{aligned}&\langle 0|\, T\Big\{\phi(x)\phi(y)\,(-i)\int dt\int d^3z\,\frac{\lambda}{4!}\phi^4\Big\}\,|0\rangle\\ &\qquad = \langle 0|\, T\Big\{\phi(x)\phi(y)\Big(\frac{-i\lambda}{4!}\Big)\int d^4z\,\phi(z)\phi(z)\phi(z)\phi(z)\Big\}\,|0\rangle\,.\end{aligned}$$

Now apply Wick's theorem. We get one term for every way of contracting the six ϕ operators with each other in pairs. There are 15 ways to do this, but (fortunately) only two of them are really different. If we contract $\phi(x)$ with $\phi(y)$, then there are three ways to contract the four $\phi(z)$'s with each other, and all three give identical expressions. The other possibility is to contract $\phi(x)$ with one of the $\phi(z)$'s (four choices), $\phi(y)$ with one of the others (three choices), and the remaining two $\phi(z)$'s with each other (one choice). There are twelve ways to do this, and all give identical expressions. Thus we have

$$\begin{aligned}&\langle 0|\, T\Big\{\phi(x)\phi(y)\,(-i)\int dt\int d^3z\,\frac{\lambda}{4!}\phi^4\Big\}\,|0\rangle\\ &\quad = 3\cdot\Big(\frac{-i\lambda}{4!}\Big)D_F(x-y)\int d^4z\,D_F(z-z)D_F(z-z)\\ &\qquad + 12\cdot\Big(\frac{-i\lambda}{4!}\Big)\int d^4z\,D_F(x-z)D_F(y-z)D_F(z-z).\end{aligned} \quad (4.44)$$

We can understand this expression better if we represent each term as a Feynman diagram. Again we draw each contraction D_F as a line, and each point as a dot. But this time we must distinguish between the "external" points, x and y, and the "internal" point z; each internal point is associated with a factor of $(-i\lambda)\int d^4z$. We will worry about the constant factors in a moment. Using these rules, we see that the above expression (4.44) is equal to the sum of two diagrams:

We refer to the lines in these diagrams as *propagators*, since they represent the propagation amplitude D_F. Internal points, where four lines meet, are called *vertices*. Since $D_F(x-y)$ is the amplitude for a free Klein-Gordon particle to propagate between x and y, the diagrams actually interpret the analytic formula as a process of particle creation, propagation, and annihilation which takes place in spacetime.

Now let's try a more complicated contraction, from the λ^3 term in the expansion of the correlation function:

$$\langle 0|\,\phi(x)\phi(y)\,\tfrac{1}{3!}\left(\tfrac{-i\lambda}{4!}\right)^3\int d^4z\,\phi\phi\phi\phi\int d^4w\,\phi\phi\phi\phi\int d^4u\,\phi\phi\phi\phi\,|0\rangle$$

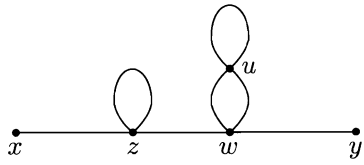
$$= \frac{1}{3!}\left(\frac{-i\lambda}{4!}\right)^3\int d^4z\,d^4w\,d^4u\;D_F(x-z)D_F(z-z)D_F(z-w)$$

$$\times\, D_F(w-y)D_F^2(w-u)D_F(u-u). \qquad (4.45)$$

The number of "different" contractions that give this same expression is large:

$$\underbrace{3!}_{\substack{\text{interchange}\\ \text{of vertices}}} \times \underbrace{4\cdot 3}_{\substack{\text{placement of}\\ \text{contractions}\\ \text{into } z \text{ vertex}}} \times \underbrace{4\cdot 3\cdot 2}_{\substack{\text{placement of}\\ \text{contractions}\\ \text{into } w \text{ vertex}}} \times \underbrace{4\cdot 3}_{\substack{\text{placement of}\\ \text{contractions}\\ \text{into } u \text{ vertex}}} \times \underbrace{1/2}_{\substack{\text{interchange}\\ \text{of } w\text{–}u\\ \text{contractions}}}$$

The product of these combinatoric factors is 10,368, roughly 1/13 of the total of 135,135 possible full contractions of the 14 operators. The structure of this particular contraction can be represented by the following "cactus" diagram:

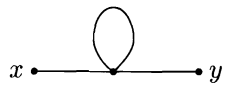


It is conventional, for obvious reasons, to let this one diagram represent the sum of all 10,368 identical terms.

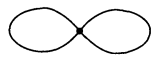
In practice one always draws the diagram first, using it as a mnemonic device for writing down the analytic expression. But then the question arises, What is the overall constant? We could, of course, work it out as above: We could associate a factor $\int d^4z(-i\lambda/4!)$ with each vertex, put in the $1/n!$ from the Taylor series, and then do the combinatorics by writing out the product of fields as in (4.45) and counting. But the $1/n!$ from the Taylor series will almost always cancel the $n!$ from interchanging the vertices, so we can just forget about both of these factors. Furthermore, the generic vertex has four lines coming in from four different places, so the various placements of these contractions into $\phi\phi\phi\phi$ generates a factor of 4! (as in the w vertex above), which cancels the denominator in $(-i\lambda/4!)$. It is therefore conventional to associate the expression $\int d^4z(-i\lambda)$ with each vertex. (This was the reason for the factor of 4! in the ϕ^4 coupling.)

In the above diagram, this scheme gives a constant that is too large by a factor of $8 = 2 \cdot 2 \cdot 2$, the *symmetry factor* of the diagram. Two factors of 2 come from lines that start and end on the same vertex: The diagram is symmetric under the interchange of the two ends of such a line. The other factor of 2 comes from the two propagators connecting w to u: The diagram is symmetric under the interchange of these two lines with each other. A third possible type of symmetry is the equivalence of two vertices. To get the correct overall constant for a diagram, we divide by its symmetry factor, which is in general the number of ways of interchanging components without changing the diagram.

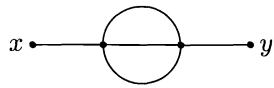
Most people never need to evaluate a diagram with a symmetry factor greater than 2, so there's no need to worry too much about these technicalities. But here are a few examples, to make some sense out of the above rules:



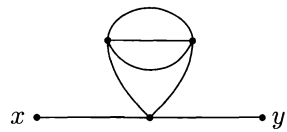
$S = 2$



$S = 2 \cdot 2 \cdot 2 = 8$



$S = 3! = 6$



$S = 3! \cdot 2 = 12$

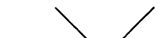
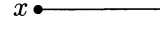
When in doubt, you can always determine the symmetry factor by counting equivalent contractions, as we did above.

We are now ready to summarize our rules for calculating the numerator

of our expression (4.31) for $\langle\Omega|\, T\phi(x)\phi(y)\,|\Omega\rangle$:

$$\langle 0|\, T\Big\{\phi_I(x)\phi_I(y)\exp\Big[-i\int dt\, H_I(t)\Big]\Big\}\,|0\rangle = \begin{pmatrix}\text{sum of all possible diagrams}\\ \text{with two external points}\end{pmatrix},$$

where each diagram is built out of propagators, vertices, and external points. The rules for associating analytic expressions with pieces of diagrams are called the *Feynman rules*. In ϕ^4 theory the rules are:

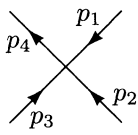
1. For each propagator,  $= D_F(x-y)$;
2. For each vertex, $= (-i\lambda)\int d^4z$;
3. For each external point,  $= 1$;
4. Divide by the symmetry factor.

One way to interpret these rules is to think of the vertex factor $(-i\lambda)$ as the amplitude for the emission and/or absorption of particles at a vertex. The integral $\int d^4z$ instructs us to sum over all points where this process can occur. This is just the *superposition* principle of quantum mechanics: When a process can happen in alternative ways, we *add* the amplitudes for each possible way. To compute each individual amplitude, the Feynman rules tell us to *multiply* the amplitudes (propagators and vertex factors) for each independent part of the process.

Since these rules are written in terms of the spacetime points x, y, etc., they are sometimes called the *position-space Feynman rules*. In most calculations, it is simpler to express the Feynman rules in terms of momenta, by introducing the Fourier expansion of each propagator:

$$D_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\,\frac{i}{p^2-m^2+i\epsilon}\,e^{-ip\cdot(x-y)}. \tag{4.46}$$

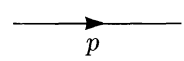
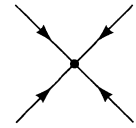
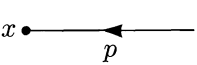
Represent this in the diagram by assigning a 4-momentum p to each propagator, indicating the direction with an arrow. (Since $D_F(x-y) = D_F(y-x)$, the direction of p is arbitrary.) Then when four lines meet at a vertex, the z-dependent factors of the diagram are



$$\longleftrightarrow \quad \int d^4z\, e^{-ip_1z}e^{-ip_2z}e^{-ip_3z}e^{+ip_4z} = (2\pi)^4\delta^{(4)}(p_1+p_2+p_3-p_4). \tag{4.47}$$

In other words, momentum is conserved at each vertex. The delta functions

from the vertices can now be used to perform some of the momentum integrals from the propagators. We are left with the following *momentum-space Feynman rules*:

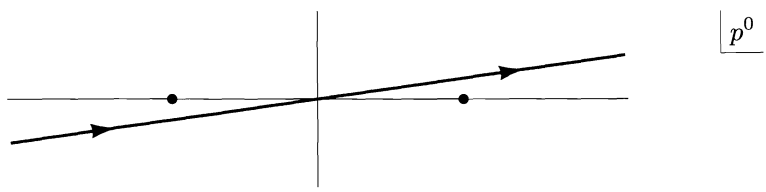
1. For each propagator,  $= \dfrac{i}{p^2 - m^2 + i\epsilon}$;
2. For each vertex,  $= -i\lambda$;
3. For each external point, x  $= e^{-ip\cdot x}$;
4. Impose momentum conservation at each vertex;
5. Integrate over each undetermined momentum: $\displaystyle\int \frac{d^4p}{(2\pi)^4}$;
6. Divide by the symmetry factor.

Again, we can interpret each factor as the amplitude for that part of the process, with the integrations coming from the superposition principle. The exponential factor for an external point is just the amplitude for a particle at that point to have the needed momentum, or, depending on the direction of the arrow, for a particle with a certain momentum to be found at that point.

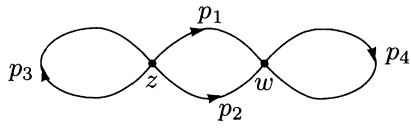
This nearly completes our discussion of the computation of correlation functions, but there are still a few loose ends. First, what happened to the large time T that was taken to $\infty(1 - i\epsilon)$? We glossed over it completely in this section, starting with Eq. (4.43). The place to put it back is Eq. (4.47), where instead of just integrating over d^4z, we should have

$$\lim_{T\to\infty(1-i\epsilon)} \int_{-T}^{T} dz^0 \int d^3z\, e^{-i(p_1+p_2+p_3-p_4)\cdot z}.$$

The exponential blows up as $z^0 \to \infty$ or $z^0 \to -\infty$ unless its argument is purely imaginary. To achieve this, we can take each p^0 to have a small imaginary part: $p^0 \propto (1 + i\epsilon)$. But this is precisely what we do in following the Feynman boundary conditions for computing D_F: We integrate along a contour that is rotated slightly away from the real axis, so that $p^0 \propto (1 + i\epsilon)$:



The explicit dependence on T seems to disappear when we take the limit $T \to \infty$ in the previous equation. But consider the diagram



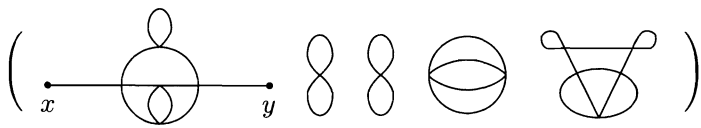
(4.48)

The delta function for the left-hand vertex is $(2\pi)^4\delta^{(4)}(p_1+p_2)$, so momentum conservation at the right-hand vertex is automatically satisfied, and we get $(2\pi)^4\delta^{(4)}(0)$ there. This awkward factor is easy to understand by going back to position space. It is simply the integral of a constant over d^4w:

$$\int d^4w\,(\text{const}) \propto (2T)\cdot(\text{volume of space}). \tag{4.49}$$

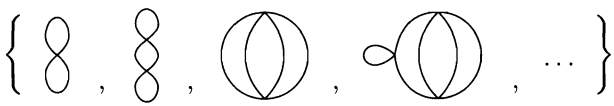
This just tells us that the spacetime process (4.48) can happen at any place in space, and at any time between $-T$ and T. Every *disconnected* piece of a diagram, that is, every piece that is not connected to an external point, will have one such $(2\pi)^4\delta^{(4)}(0) = 2T\cdot V$ factor.

The contributions to the correlation function coming from such diagrams can be better understood with the help of a very pretty identity, *the exponentiation of the disconnected diagrams.* It works as follows. A typical diagram has the form



(4.50)

with a piece connected to x and y, and several disconnected pieces. (Since each vertex has an even number of lines coming into it, x and y must be connected to each other.) Label the various possible disconnected pieces by V_i:



(4.51)

The elements V_i are connected internally, but disconnected from external points. Suppose that a diagram (such as (4.50)) has n_i pieces of the form V_i, for each i, in addition to its one piece that is connected to x and y. (In any given diagram, only finitely many of the n_i will be nonzero.) If we also let V_i denote the *value* of the piece V_i, then the value of such a diagram is

$$(\text{value of connected piece})\cdot\prod_i \frac{1}{n_i!}(V_i)^{n_i}.$$

The $1/n_i!$ is the symmetry factor coming from interchanging the n_i copies of V_i. The sum of all diagrams, representing the numerator of our formula for

the two-point correlation function, is then

$$\sum_{\substack{\text{all possible}\\ \text{connected}\\ \text{pieces}}} \sum_{\text{all } \{n_i\}} \left(\begin{matrix} \text{value of} \\ \text{connected piece} \end{matrix} \right) \times \left(\prod_i \frac{1}{n_i!} (V_i)^{n_i} \right),$$

where "all $\{n_i\}$" means "all ordered sets $\{n_1, n_2, n_3, \ldots\}$ of nonnegative integers." The sum of the connected pieces factors out of this expression, giving

$$= \Big(\sum \text{connected} \Big) \times \sum_{\text{all } \{n_i\}} \left(\prod_i \frac{1}{n_i!} (V_i)^{n_i} \right),$$

where $\big(\sum \text{connected}\big)$ is an abbreviation for the sum of the values of all connected pieces. It is not too hard to see that the rest of the expression can also be factored (try working backwards):

$$\begin{aligned}
&= \Big(\sum \text{connected} \Big) \times \Big(\sum_{n_1} \frac{1}{n_1!} V_1^{n_1} \Big) \Big(\sum_{n_2} \frac{1}{n_2!} V_2^{n_2} \Big) \Big(\sum_{n_3} \frac{1}{n_3!} V_3^{n_3} \Big) \cdots \\
&= \Big(\sum \text{connected} \Big) \times \prod_i \Big(\sum_{n_i} \frac{1}{n_i!} V_i^{n_i} \Big) \\
&= \Big(\sum \text{connected} \Big) \times \prod_i \exp(V_i) \\
&= \Big(\sum \text{connected} \Big) \times \exp\Big(\sum_i V_i \Big). \qquad (4.52)
\end{aligned}$$

We have just shown that the sum of *all* diagrams is equal to the sum of all *connected* diagrams, times the exponential of the sum of all *disconnected* diagrams. (We should really say "pieces" rather than "diagrams" on the right-hand side of the equality, but from now on we will often just call a single piece a "diagram.") Pictorially, the identity is

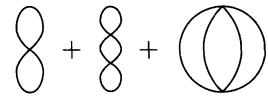
$$\lim_{T \to \infty(1 - i\epsilon)} \langle 0 | \, T \Big\{ \phi_I(x) \phi_I(y) \exp\Big[-i \int_{-T}^{T} dt \, H_I(t) \Big] \Big\} \, | 0 \rangle$$

$$= \Big(\quad + \quad + \quad + \cdots \Big)$$

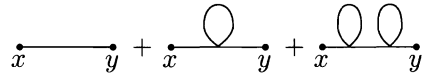
$$\times \exp \Big[\quad + \quad + \quad + \cdots \Big]. \qquad (4.53)$$

Now consider the *denominator* of our formula (4.31) for the two-point

function. By an argument identical to the above, it is just

$$\langle 0|\, T\Big\{\exp\Big[-i\int_{-T}^{T} dt\, H_I(t)\Big]\Big\}\,|0\rangle = \exp\Big[$$  $$\Big],$$

which cancels the exponential of the disconnected diagrams in the numerator. This is the final simplification of the formula, which now reads

$$\begin{aligned}&\langle \Omega|\, T\big[\phi(x)\phi(y)\big]\,|\Omega\rangle \\ &\quad = \text{ sum of all connected diagrams with two external points} \\ &\quad =\end{aligned}$$  (4.54)

We have come a long way from our original formula, Eq. (4.31).

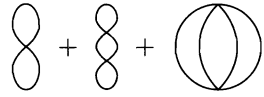
Having gotten rid of the disconnected diagrams in our formula for the correlation function, we might pause a moment to go back and interpret them physically. The place to look is Eq. (4.30), which can be written

$$\begin{aligned}&\lim_{T\to\infty(1-i\epsilon)} \langle 0|\, T\Big\{\phi_I(x)\phi_I(y)\exp\Big[-i\textstyle\int_{-T}^{T} dt\, H_I(t)\Big]\Big\}\,|0\rangle \\ &\qquad = \langle \Omega|\, T\phi(x)\phi(y)\,|\Omega\rangle \cdot \lim_{T\to\infty(1-i\epsilon)} \big(|\langle 0\,|\Omega\rangle\,|^2 e^{-iE_0(2T)}\big).\end{aligned}$$

Looking only at the T-dependent parts of both sides, this implies

$$\exp\Big[\sum_i V_i\Big] \propto \exp\Big[-iE_0(2T)\Big]. \tag{4.55}$$

Since each disconnected diagram V_i contains a factor of $(2\pi)^4\delta^{(4)}(0) = 2T\cdot V$, this gives us a formula for the energy density of the vacuum (relative to the zero of energy set by $H_0\,|0\rangle = 0$):

$$\frac{E_0}{\text{volume}} = i\Big[$$  $$\Big]\Big/\Big[(2\pi)^4\delta^{(4)}(0)\Big]. \tag{4.56}$$

We should emphasize that the right-hand side is independent of T and (volume); in particular it is reassuring to see that E_0 is proportional to the volume of space. In Chapter 11 we will find that this formula is actually useful.

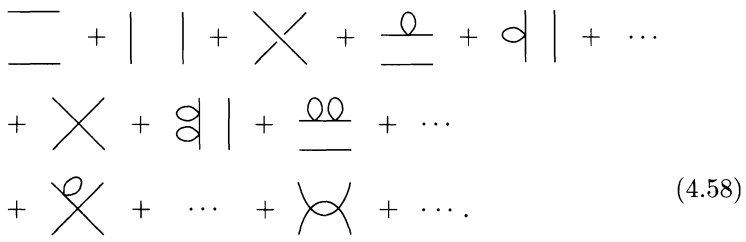
This completes our present analysis of the two-point correlation function. The generalization to higher correlation functions is easy:

$$\langle \Omega|\, T\big[\phi(x_1)\cdots\phi(x_n)\big]\,|\Omega\rangle = \begin{pmatrix}\text{sum of all connected diagrams} \\ \text{with } n \text{ external points}\end{pmatrix}. \tag{4.57}$$

The disconnected diagrams exponentiate, factor, and cancel as before, by the same argument. There is now a potential confusion in terminology, however. By "disconnected" we mean "disconnected from all external points"—exactly the same diagrams as in (4.51). (They are sometimes called "vacuum bubbles"

or "vacuum to vacuum transitions".) In higher correlation functions, diagrams can also be disconnected in another sense. Consider, for example, the four-point function:

$\langle\Omega|\, T\phi_1\phi_2\phi_3\phi_4\,|\Omega\rangle$



(4.58)

In many of these diagrams, external points are disconnected *from each other.* Such diagrams do not exponentiate or factor; they contribute to the amplitude just as do the fully connected diagrams (in which any point can be reached from any other by traveling along the lines).

Note that in ϕ^4 theory, all correlation functions of an odd number of fields vanish, since it is impossible to draw an allowed diagram with an odd number of external points. We can also see this by going back to Wick's theorem: The interaction Hamiltonian H_I contains an even number of fields, so all terms in the perturbation expansion of an odd correlation function will contain an odd number of fields. But it is impossible to fully contract an odd number of fields in pairs, and only fully contracted terms have nonvanishing vacuum expectation value.

4.5 Cross Sections and the S-Matrix

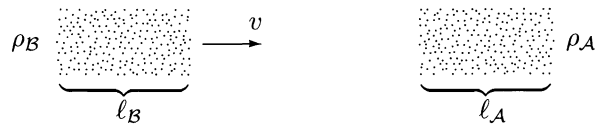
We now have an extremely beautiful formula, Eq. (4.57), for computing an extremely abstract quantity: the n-point correlation function. Our next task is to find equally beautiful ways of computing quantities that can actually be measured: cross sections and decay rates. In this section, after briefly reviewing the definitions of these objects, we will relate them (via a rather technical but fairly careful derivation) to a more primitive quantity, the S-matrix. In the next section we will learn how to compute the matrix elements of the S-matrix using Feynman diagrams.

The Cross Section

The experiments that probe the behavior of elementary particles, especially in the relativistic regime, are scattering experiments. One collides two beams of particles with well-defined momenta, and observes what comes out. The likelihood of any particular final state can be expressed in terms of the *cross section*, a quantity that is intrinsic to the colliding particles, and therefore

allows comparison of two different experiments with different beam sizes and intensities.

The cross section is defined as follows. Consider a target, at rest, of particles of type $\mathcal{A}$, with density $\rho_{\mathcal{A}}$ (particles per unit volume). Aim at this target a bunch of particles of type $\mathcal{B}$, with number density $\rho_{\mathcal{B}}$ and velocity v:



Let $\ell_{\mathcal{A}}$ and $\ell_{\mathcal{B}}$ be the lengths of the bunches of particles. Then we expect the total number of scattering events (or scattering events of any particular desired type) to be proportional to $\rho_{\mathcal{A}}$, $\rho_{\mathcal{B}}$, $\ell_{\mathcal{A}}$, $\ell_{\mathcal{B}}$, and the cross-sectional area A common to the two bunches. The *cross section*, denoted by σ, is just the total number of events (of whatever type desired) divided by all of these quantities:

$$\sigma \equiv \frac{\text{Number of scattering events}}{\rho_{\mathcal{A}}\,\ell_{\mathcal{A}}\,\rho_{\mathcal{B}}\,\ell_{\mathcal{B}}\,A}. \tag{4.59}$$

The definition is symmetric between the $\mathcal{A}$'s and $\mathcal{B}$'s, so of course we could have taken the $\mathcal{B}$'s to be at rest, or worked in any other reference frame.

The cross section has units of area. In fact, it is the effective area of a chunk taken out of one beam, by each particle in the other beam, that subsequently becomes the final state we are interested in.

In real beams, $\rho_{\mathcal{A}}$ and $\rho_{\mathcal{B}}$ are not constant; the particle density is generally larger at the center of the beam than at the edges. We will assume, however, that both the range of the interaction between the particles and the width of the individual particle wavepackets are small compared to the beam diameter. We can then consider $\rho_{\mathcal{A}}$ and $\rho_{\mathcal{B}}$ to be constant in what follows, and remember that, to compute the event rate in an actual accelerator, one must integrate over the beam area:

$$\text{Number of events} = \sigma\,\ell_{\mathcal{A}}\,\ell_{\mathcal{B}} \int d^2x\ \rho_{\mathcal{A}}(x)\,\rho_{\mathcal{B}}(x). \tag{4.60}$$

If the densities are constant, or if we use this formula to compute an effective area A of the beams, then we have simply

$$\text{Number of events} = \frac{\sigma N_{\mathcal{A}} N_{\mathcal{B}}}{A}, \tag{4.61}$$

where $N_{\mathcal{A}}$ and $N_{\mathcal{B}}$ are the total numbers of $\mathcal{A}$ and $\mathcal{B}$ particles.

Cross sections for many different processes may be relevant to a single scattering experiment. In e^+e^- collisions, for example, one can measure the cross sections for production of $\mu^+\mu^-$, $\tau^+\tau^-$, $\mu^+\mu^-\gamma$, $\mu^+\mu^-\gamma\gamma$, etc., and countless processes involving hadron production, not to mention simple e^+e^- scattering. Usually, of course, we wish to measure not only what the final-state particles are, but also the momenta with which they come out. In this case

our definition (4.59) of σ still works, but if we specify the exact momenta desired, σ will be infinitesimal. The solution is to define the *differential cross section*, $d\sigma/(d^3p_1 \cdots d^3p_n)$. It is simply the quantity that, when integrated over any small $d^3p_1 \cdots d^3p_n$, gives the cross section for scattering into that region of final-state momentum space. The various final-state momenta are not all independent: Four components will always be constrained by 4-momentum conservation. In the simplest case, where there are only two final-state particles, this leaves only two unconstrained momentum components, usually taken to be the angles θ and ϕ of the momentum of one of the particles. Integrating $d\sigma/(d^3p_1 d^3p_2)$ over the four constrained momentum components then leaves us with the usual differential cross section $d\sigma/d\Omega$.

A somewhat simpler measurable quantity is the *decay rate* Γ of an unstable particle $\mathcal{A}$ (assumed to be at rest) into a specified final state (of two or more particles). It is defined as

$$\Gamma \equiv \frac{\text{Number of decays per unit time}}{\text{Number of } \mathcal{A} \text{ particles present}}. \tag{4.62}$$

The lifetime τ of the particle is then the reciprocal of the sum of its decay rates into all possible final states. (The particle's half-life is $\tau \cdot \ln 2$.)

In nonrelativistic quantum mechanics, an unstable atomic state shows up in scattering experiments as a *resonance*. Near the resonance energy E_0, the scattering amplitude is given by the Breit-Wigner formula

$$f(E) \propto \frac{1}{E - E_0 + i\Gamma/2}. \tag{4.63}$$

The cross section therefore has a peak of the form

$$\sigma \propto \frac{1}{(E - E_0)^2 + \Gamma^2/4}.$$

The width of the resonance peak is equal to the decay rate of the unstable state.

The Breit-Wigner formula (4.63) also applies in relativistic quantum mechanics. In particular, it gives the scattering amplitude for processes in which initial particles combine to form an unstable particle, which then decays. The unstable particle, viewed as an excited state of the vacuum, is a direct analogue of the unstable nonrelativistic atomic state. If we call the 4-momentum of the unstable particle p and its mass m, we can make a relativistically invariant generalization of (4.63):

$$\frac{1}{p^2 - m^2 + im\Gamma} \approx \frac{1}{2E_{\mathbf{p}}(p^0 - E_{\mathbf{p}} + i(m/E_{\mathbf{p}})\Gamma/2)}. \tag{4.64}$$

The decay rate of the unstable particle in a general frame is $(m/E_{\mathbf{p}})\Gamma$, in accord with relativistic time dilation. Although the two expressions in (4.64) are equal in the vicinity of the resonance, the left-hand side, which is manifestly Lorentz invariant, is much more convenient.

The S-Matrix

How, then, do we calculate a cross section? We must set up wavepackets representing the initial-state particles, evolve this initial state for a very long time with the time-evolution operator $\exp(-iHt)$ of the interacting field theory, and overlap the resulting final state with wavepackets representing some desired set of final-state particles. This gives the probability amplitude for producing that final state, which is simply related to the cross section. We will find that, in the limit where the wavepackets are very narrow in momentum space, the amplitude depends only on the momenta of the wavepackets, not on the details of their shapes.†

A wavepacket representing some desired state $|\phi\rangle$ can be expressed as

$$|\phi\rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{k}}}} \phi(\mathbf{k}) \, |\mathbf{k}\rangle \,, \tag{4.65}$$

where $\phi(\mathbf{k})$ is the Fourier transform of the spatial wavefunction, and $|\mathbf{k}\rangle$ is a one-particle state of momentum $\mathbf{k}$ in the interacting theory. In the free theory, we would have $|\mathbf{k}\rangle = \sqrt{2E_{\mathbf{k}}} a^\dagger_{\mathbf{k}} |0\rangle$. The factor of $\sqrt{2E_{\mathbf{k}}}$ converts our relativistic normalization of $|\mathbf{k}\rangle$ to the conventional normalization in which the sum of all probabilities adds up to 1:

$$\langle\phi|\phi\rangle = 1 \qquad \text{if} \qquad \int \frac{d^3k}{(2\pi)^3} \left|\phi(\mathbf{k})\right|^2 = 1. \tag{4.66}$$

The probability we wish to compute is then

$$\mathcal{P} = \Big| \underbrace{\langle \phi_1 \phi_2 \cdots |}_{\text{future}} \underbrace{| \phi_\mathcal{A} \phi_\mathcal{B} \rangle}_{\text{past}} \Big|^2, \tag{4.67}$$

where $|\phi_\mathcal{A}\phi_\mathcal{B}\rangle$ is a state of two wavepackets constructed in the far past and $\langle\phi_1\phi_2\cdots|$ is a state of several wavepackets (one for each final-state particle) constructed in the far future. The wavepackets are localized in space, so each can be constructed independently of the others. States constructed in this way are called *in* and *out* states. Note that we use the Heisenberg picture: States are time-independent, but the name we give a state depends on the eigenvalues or expectation values of time-dependent operators. Thus states with different names constructed at different times have a nontrivial overlap, which depends on the time dependence of the operators.

If we set up $|\phi_\mathcal{A}\phi_\mathcal{B}\rangle$ in the remote past, and then take the limit in which the wavepackets $\phi_i(\mathbf{k}_i)$ become concentrated about definite momenta $\mathbf{p}_i$, this defines an *in* state $|\mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle_{\text{in}}$ with definite initial momenta. It is useful to view $|\phi_\mathcal{A}\phi_\mathcal{B}\rangle$ as a linear superposition of such states. It is important, however, to

†Much of this section is based on the treatment of nonrelativistic scattering given in Taylor (1972), Chapters 2, 3, and 17. We concentrate on the additional complications of the relativistic theory, glossing over many subtleties, common to both cases, which Taylor explains carefully.

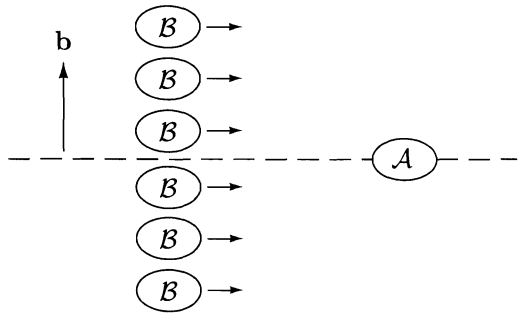


Figure 4.2. Incident wavepackets are uniformly distributed in impact parameter **b**.

take into account the transverse displacement of the wavepacket $\phi_{\mathcal{B}}$ relative to $\phi_{\mathcal{A}}$ in position space (see Fig. 4.2). Although we could leave this implicit in the form of $\phi_{\mathcal{B}}(\mathbf{k}_{\mathcal{B}})$, we instead adopt the convention that our reference momentum-space wavefunctions are collinear (that is, have impact parameter $\mathbf{b} = 0$), and write $\phi_{\mathcal{B}}(\mathbf{k}_{\mathcal{B}})$ with an explicit factor $\exp(-i\mathbf{b}\cdot\mathbf{k}_{\mathcal{B}})$ to account for the spatial translation. Then, since $\phi_{\mathcal{A}}$ and $\phi_{\mathcal{B}}$ are constructed independently at different locations, we can write the initial state as

$$|\phi_{\mathcal{A}}\phi_{\mathcal{B}}\rangle_{\text{in}} = \int \frac{d^3k_{\mathcal{A}}}{(2\pi)^3} \int \frac{d^3k_{\mathcal{B}}}{(2\pi)^3} \frac{\phi_{\mathcal{A}}(\mathbf{k}_{\mathcal{A}})\phi_{\mathcal{B}}(\mathbf{k}_{\mathcal{B}})e^{-i\mathbf{b}\cdot\mathbf{k}_{\mathcal{B}}}}{\sqrt{(2E_{\mathcal{A}})(2E_{\mathcal{B}})}} |\mathbf{k}_{\mathcal{A}}\mathbf{k}_{\mathcal{B}}\rangle_{\text{in}}. \tag{4.68}$$

We could expand $\langle\phi_1\phi_2\cdots|$ in terms of similarly defined *out* states of definite momentum formed in the asymptotic future:‡

$$_{\text{out}}\langle\phi_1\phi_2\cdots| = \left(\prod_f \int \frac{d^3p_f}{(2\pi)^3} \frac{\phi_f(\mathbf{p}_f)}{\sqrt{2E_f}}\right) {}_{\text{out}}\langle\mathbf{p}_1\mathbf{p}_2\cdots|.$$

It is much easier, however, to use the *out* states of definite momentum as the final states in the probability amplitude (4.67), and to multiply by the various normalization factors after squaring the amplitude. This is physically reasonable as long as the detectors of final-state particles mainly measure momentum—that is, they do not resolve positions at the level of de Broglie wavelengths.

We can now relate the probability of scattering in a real experiment to an idealized set of transition amplitudes between the asymptotically defined *in* and *out* states of definite momentum,

$$_{\text{out}}\langle\mathbf{p}_1\mathbf{p}_2\cdots|\mathbf{k}_{\mathcal{A}}\mathbf{k}_{\mathcal{B}}\rangle_{\text{in}}. \tag{4.69}$$

‡Here and below, the product symbol applies (symbolically) to the integral as well as the other factors in parentheses; the integrals apply to what is outside the parentheses as well.

To compute the overlap of *in* states with *out* states, we note that the conventions for defining the two sets of states are related by time translation:

$$\begin{aligned} {}_{\text{out}}\langle \mathbf{p}_1\mathbf{p}_2\cdots|\mathbf{k}_\mathcal{A}\mathbf{k}_\mathcal{B}\rangle_{\text{in}} &= \lim_{T\to\infty} \langle \underbrace{\mathbf{p}_1\mathbf{p}_2\cdots}_{T} | \underbrace{\mathbf{k}_\mathcal{A}\mathbf{k}_\mathcal{B}}_{-T}\rangle \\ &= \lim_{T\to\infty} \langle \mathbf{p}_1\mathbf{p}_2\cdots | \, e^{-iH(2T)} \, |\mathbf{k}_\mathcal{A}\mathbf{k}_\mathcal{B}\rangle \,. \end{aligned} \tag{4.70}$$

In the last line, the states are defined at any common reference time. Thus, the *in* and *out* states are related by the limit of a sequence of unitary operators. This limiting unitary operator is called the *S-matrix*:

$${}_{\text{out}}\langle \mathbf{p}_1\mathbf{p}_2\cdots|\mathbf{k}_\mathcal{A}\mathbf{k}_\mathcal{B}\rangle_{\text{in}} \equiv \langle \mathbf{p}_1\mathbf{p}_2\cdots | \, S \, |\mathbf{k}_\mathcal{A}\mathbf{k}_\mathcal{B}\rangle \,. \tag{4.71}$$

The S-matrix has the following structure: If the particles in question do not interact at all, S is simply the identity operator. Even if the theory contains interactions, the particles have some probability of simply missing one another. To isolate the interesting part of the S-matrix—that is, the part due to interactions—we define the *T-matrix* by

$$S = \mathbf{1} + iT. \tag{4.72}$$

Next we note that the matrix elements of S should reflect 4-momentum conservation. Thus S or T should always contain a factor $\delta^{(4)}(k_\mathcal{A} + k_\mathcal{B} - \sum p_f)$. Extracting this factor, we define the *invariant matrix element* $\mathcal{M}$, by

$$\langle \mathbf{p}_1\mathbf{p}_2\cdots| \, iT \, |\mathbf{k}_\mathcal{A}\mathbf{k}_\mathcal{B}\rangle = (2\pi)^4\delta^{(4)}\big(k_\mathcal{A}+k_\mathcal{B} - \textstyle\sum p_f\big) \cdot i\mathcal{M}(k_\mathcal{A}, k_\mathcal{B} \to p_f). \tag{4.73}$$

We have written this expression in terms of 4-momenta p and k, but of course all 4-momenta are on mass-shell: $p^0 = E_\mathbf{p}$, $k^0 = E_\mathbf{k}$. (Note that our entire treatment is specific to the case where the initial state contains only two particles. For 3→many or many→many interactions, one can invent analogous constructions, but we will not consider such complicated experiments in this book.)

The matrix element $\mathcal{M}$ is analogous to the scattering amplitude f of one-particle quantum mechanics. It is useful because it allows us to separate all the physics that depends on the details of the interaction Hamiltonian ("dynamics") from all the physics that doesn't ("kinematics"). In the next section we will discuss how to compute $\mathcal{M}$ using Feynman diagrams. But first, we must figure out how to reconstruct the cross section σ from $\mathcal{M}$.

To do this, let us calculate, in terms of $\mathcal{M}$, the probability for the initial state $|\phi_\mathcal{A}\phi_\mathcal{B}\rangle$ to scatter and become a final state of n particles whose momenta lie in a small region $d^3p_1 \cdots d^3p_n$. In our normalization, this probability is

$$\mathcal{P}(\mathcal{A}\mathcal{B} \to 1\,2\,\ldots n) = \left(\prod_f \frac{d^3p_f}{(2\pi)^3}\frac{1}{2E_f}\right) \big|{}_{\text{out}}\langle \mathbf{p}_1\cdots\mathbf{p}_n \mid \phi_\mathcal{A}\phi_\mathcal{B}\rangle_{\text{in}}\big|^2. \tag{4.74}$$

For a single target ($\mathcal{A}$) particle and many incident ($\mathcal{B}$) particles with different impact parameters $\mathbf{b}$, the number of scattering events is

$$N = \sum_{\substack{\text{all incident} \\ \text{particles } i}} \mathcal{P}_i = \int d^2b \; n_{\mathcal{B}} \, \mathcal{P}(\mathbf{b}),$$

where $n_{\mathcal{B}}$ is the number density (particles per unit area) of $\mathcal{B}$ particles. Since we are assuming that this number density is constant over the range of the interaction, $n_{\mathcal{B}}$ can be taken outside the integral. The cross section is then

$$\sigma = \frac{N}{n_{\mathcal{B}} N_{\mathcal{A}}} = \frac{N}{n_{\mathcal{B}} \cdot 1} = \int d^2b \; \mathcal{P}(\mathbf{b}). \tag{4.75}$$

Deriving a simple expression for σ in terms of $\mathcal{M}$ is now a fairly straightforward calculation. Combining (4.75), (4.74), and (4.68), we have (writing $d\sigma$ rather than σ since this is an infinitesimal quantity)

$$\begin{aligned} d\sigma = \left(\prod_f \frac{d^3p_f}{(2\pi)^3}\frac{1}{2E_f}\right)\int d^2b \left(\prod_{i=\mathcal{A},\mathcal{B}} \int \frac{d^3k_i}{(2\pi)^3}\frac{\phi_i(\mathbf{k}_i)}{\sqrt{2E_i}}\right) \int \frac{d^3\bar{k}_i}{(2\pi)^3}\frac{\phi_i^*(\bar{\mathbf{k}}_i)}{\sqrt{2\bar{E}_i}}\Bigg) \\ \times e^{i\mathbf{b}\cdot(\bar{\mathbf{k}}_{\mathcal{B}}-\mathbf{k}_{\mathcal{B}})} \big({}_{\text{out}}\langle\{\mathbf{p}_f\}|\{\mathbf{k}_i\}\rangle_{\text{in}}\big)\big({}_{\text{out}}\langle\{\mathbf{p}_f\}|\{\bar{\mathbf{k}}_i\}\rangle_{\text{in}}\big)^*, \end{aligned} \tag{4.76}$$

where we have used $\bar{k}_{\mathcal{A}}$ and $\bar{k}_{\mathcal{B}}$ as dummy integration variables in the second half of the squared amplitude. The d^2b integral can be performed to give a factor of $(2\pi)^2\delta^{(2)}(k_{\mathcal{B}}^{\perp} - \bar{k}_{\mathcal{B}}^{\perp})$. We get more delta functions by writing the final two factors of (4.76) in terms of $\mathcal{M}$. Assuming that we are not interested in the trivial case of forward scattering where no interaction takes place, we can drop the **1** in Eq. (4.72) and write these factors as

$$\begin{aligned} \big({}_{\text{out}}\langle\{\mathbf{p}_f\}|\{\mathbf{k}_i\}\rangle_{\text{in}}\big) &= i\mathcal{M}(\{k_i\} \to \{p_f\})\,(2\pi)^4\delta^{(4)}(\textstyle\sum k_i - \sum p_f); \\ \big({}_{\text{out}}\langle\{\mathbf{p}_f\}|\{\bar{\mathbf{k}}_i\}\rangle_{\text{in}}\big)^* &= -i\mathcal{M}^*(\{\bar{k}_i\} \to \{p_f\})\,(2\pi)^4\delta^{(4)}(\textstyle\sum \bar{k}_i - \sum p_f). \end{aligned}$$

We can use the second of these delta functions, together with the $\delta^{(2)}(k_{\mathcal{B}}^{\perp} - \bar{k}_{\mathcal{B}}^{\perp})$, to perform all six of the $\bar{k}$ integrals in (4.76). Of the six integrals, only those over $\bar{k}_{\mathcal{A}}^z$ and $\bar{k}_{\mathcal{B}}^z$ require some work:

$$\begin{aligned} &\int d\bar{k}_{\mathcal{A}}^z d\bar{k}_{\mathcal{B}}^z \, \delta(\bar{k}_{\mathcal{A}}^z + \bar{k}_{\mathcal{B}}^z - \textstyle\sum p_f^z)\,\delta(\bar{E}_{\mathcal{A}} + \bar{E}_{\mathcal{B}} - \sum E_f) \\ &\quad = \int d\bar{k}_{\mathcal{A}}^z \, \delta\big(\sqrt{\bar{k}_{\mathcal{A}}^2 + m_{\mathcal{A}}^2} + \sqrt{\bar{k}_{\mathcal{B}}^2 + m_{\mathcal{B}}^2} - \textstyle\sum E_f\big)\Big|_{\bar{k}_{\mathcal{B}}^z = \sum p_f^z - \bar{k}_{\mathcal{A}}^z} \\ &\quad = \frac{1}{\left|\dfrac{\bar{k}_{\mathcal{A}}^z}{\bar{E}_{\mathcal{A}}} - \dfrac{\bar{k}_{\mathcal{B}}^z}{\bar{E}_{\mathcal{B}}}\right|} \equiv \frac{1}{|v_{\mathcal{A}} - v_{\mathcal{B}}|}. \end{aligned} \tag{4.77}$$

In the last line and in the rest of Eq. (4.76) it is understood that the constraints $\bar{k}_{\mathcal{A}}^z + \bar{k}_{\mathcal{B}}^z = \sum p_f^z$ and $\bar{E}_{\mathcal{A}} + \bar{E}_{\mathcal{B}} = \sum E_f$ now apply (in addition to the constraints $\bar{k}_{\mathcal{A}}^{\perp} = k_{\mathcal{A}}^{\perp}$ and $\bar{k}_{\mathcal{B}}^{\perp} = k_{\mathcal{B}}^{\perp}$ coming from the other four integrals).

The difference $|v_{\mathcal{A}} - v_{\mathcal{B}}|$ is the relative velocity of the beams as viewed from the laboratory frame.

Now recall that the initial wavepackets are localized in momentum space, centered on $\mathbf{p}_{\mathcal{A}}$ and $\mathbf{p}_{\mathcal{B}}$. This means that we can evaluate all factors that are smooth functions of $\mathbf{k}_{\mathcal{A}}$ and $\mathbf{k}_{\mathcal{B}}$ at $\mathbf{p}_{\mathcal{A}}$ and $\mathbf{p}_{\mathcal{B}}$, pulling them outside the integrals. These factors include $E_{\mathcal{A}}$, $E_{\mathcal{B}}$, $|v_{\mathcal{A}} - v_{\mathcal{B}}|$, and $\mathcal{M}$—everything except the remaining delta function. After doing this, we arrive at the expression

$$d\sigma = \left(\prod_f \frac{d^3p_f}{(2\pi)^3}\frac{1}{2E_f}\right)\frac{\left|\mathcal{M}(p_{\mathcal{A}}, p_{\mathcal{B}} \to \{p_f\})\right|^2}{2E_{\mathcal{A}}2E_{\mathcal{B}}\,|v_{\mathcal{A}}-v_{\mathcal{B}}|}\int\frac{d^3k_{\mathcal{A}}}{(2\pi)^3}\int\frac{d^3k_{\mathcal{B}}}{(2\pi)^3} \times \left|\phi_{\mathcal{A}}(\mathbf{k}_{\mathcal{A}})\right|^2\left|\phi_{\mathcal{B}}(\mathbf{k}_{\mathcal{B}})\right|^2(2\pi)^4\delta^{(4)}(k_{\mathcal{A}}+k_{\mathcal{B}}-\textstyle\sum p_f). \tag{4.78}$$

To simplify this formula further, we should think a bit more about the properties of real particle detectors. We have already noted that real detectors project mainly onto eigenstates of momentum. But real detectors have finite resolution; that is, they sum incoherently over momentum bites of finite size. Normally, the measurement of the final-state momentum is not of such high quality that it can resolve the small variation of this momentum that results from the momentum spread of the initial wavepackets $\phi_{\mathcal{A}}$, $\phi_{\mathcal{B}}$. In that case, we may treat even the momentum vector $k_{\mathcal{A}} + k_{\mathcal{B}}$ inside the delta function as being well approximated by its central value $p_{\mathcal{A}} + p_{\mathcal{B}}$. With this further approximation, we can perform the integrals over $k_{\mathcal{A}}$ and $k_{\mathcal{B}}$ using the normalization condition (4.66). This produces the final form of the relation between S-matrix elements and cross sections,

$$d\sigma = \frac{1}{2E_{\mathcal{A}}2E_{\mathcal{B}}\,|v_{\mathcal{A}}-v_{\mathcal{B}}|}\left(\prod_f \frac{d^3p_f}{(2\pi)^3}\frac{1}{2E_f}\right) \times \left|\mathcal{M}(p_{\mathcal{A}}, p_{\mathcal{B}} \to \{p_f\})\right|^2(2\pi)^4\delta^{(4)}(p_{\mathcal{A}}+p_{\mathcal{B}}-\textstyle\sum p_f). \tag{4.79}$$

All dependence on the shapes of the wavepackets has disappeared.

The integral over final-state momenta in (4.79) has the structure

$$\int d\Pi_n = \left(\prod_f \int \frac{d^3p_f}{(2\pi)^3}\frac{1}{2E_f}\right)(2\pi)^4\delta^{(4)}(P - \textstyle\sum p_f), \tag{4.80}$$

with P the total initial 4-momentum. This integral is manifestly Lorentz invariant, since it is built up from invariant 3-momentum integrals constrained by a 4-momentum delta function. This integral is known as *relativistically invariant n-body phase space.* Of the other ingredients in (4.79), the matrix element $\mathcal{M}$ is also Lorentz invariant. The Lorentz transformation property of (4.79) therefore comes entirely from the prefactor

$$\frac{1}{E_{\mathcal{A}}E_{\mathcal{B}}|v_{\mathcal{A}} - v_{\mathcal{B}}|} = \frac{1}{|E_{\mathcal{B}}p^z_{\mathcal{A}} - E_{\mathcal{A}}p^z_{\mathcal{B}}|} = \frac{1}{|\epsilon_{\mu x y \nu}p^\mu_{\mathcal{A}}p^\nu_{\mathcal{B}}|}.$$

This is not Lorentz invariant, but it is invariant to boosts along the z-direction. In fact, this expression has exactly the transformation properties of a cross-sectional area.

For the special case of two particles in the final state, we can simplify the general expression (4.79) by partially evaluating the phase-space integrals in the center-of-mass frame. Label the momenta of the two final particles p_1 and p_2. We first choose to integrate all three components of $\mathbf{p}_2$ over the delta functions enforcing 3-momentum conservation. This sets $\mathbf{p}_2 = -\mathbf{p}_1$ and converts the integral over two-body phase space to the form

$$\int d\Pi_2 = \int \frac{dp_1\, p_1^2\, d\Omega}{(2\pi)^3\, 2E_1\, 2E_2}\, (2\pi)\delta(E_{\rm cm} - E_1 - E_2), \tag{4.81}$$

where $E_1 = \sqrt{p_1^2 + m_1^2}$, $E_2 = \sqrt{p_1^2 + m_2^2}$, and $E_{\rm cm}$ is the total initial energy. Integrating over the final delta function gives

$$\begin{aligned} \int d\Pi_2 &= \int d\Omega\, \frac{p_1^2}{16\pi^2 E_1 E_2} \left(\frac{p_1}{E_1} + \frac{p_1}{E_2}\right)^{-1} \\ &= \int d\Omega\, \frac{1}{16\pi^2}\, \frac{|\mathbf{p}_1|}{E_{\rm cm}}. \end{aligned} \tag{4.82}$$

For reactions symmetric about the collision axis, two-body phase space can be written simply as an integral over the polar angle in the center-of-mass frame:

$$\int d\Pi_2 = \int d\cos\theta\, \frac{1}{16\pi}\, \frac{2|\mathbf{p}_1|}{E_{\rm cm}}. \tag{4.83}$$

The last factor tends to 1 at high energy.

Applying this simplification to (4.79), we find the following form of the cross section for two final-state particles:

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm CM} = \frac{1}{2E_{\mathcal{A}} 2E_{\mathcal{B}}\, |v_{\mathcal{A}} - v_{\mathcal{B}}|}\, \frac{|\mathbf{p}_1|}{(2\pi)^2\, 4E_{\rm cm}} \left|\mathcal{M}(p_{\mathcal{A}}, p_{\mathcal{B}} \to p_1, p_2)\right|^2. \tag{4.84}$$

In the special case where all four particles have identical masses (including the commonly seen limit $m \to 0$), this reduces to the formula quoted in Chapter 1,

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm CM} = \frac{|\mathcal{M}|^2}{64\pi^2 E_{\rm cm}^2} \qquad \text{(all four masses identical)}. \tag{4.85}$$

To conclude this section, we should derive a formula for the differential decay rate, $d\Gamma$, in terms of $\mathcal{M}$. The correct expression is only a slight modification of (4.79), and is quite easy to guess: Just remove from (4.79) the factors that do not make sense when the initial state consists of a single particle. The definition of Γ assumes that the decaying particle is at rest, so the normalization factor $(2E_{\mathcal{A}})^{-1}$ becomes $(2m_{\mathcal{A}})^{-1}$. (In any other frame, this factor would give the usual time dilation.) Thus the decay rate formula is

$$d\Gamma = \frac{1}{2m_{\mathcal{A}}} \left(\prod_f \frac{d^3 p_f}{(2\pi)^3}\, \frac{1}{2E_f}\right) \left|\mathcal{M}(m_{\mathcal{A}} \to \{p_f\})\right|^2 (2\pi)^4 \delta^{(4)}(p_{\mathcal{A}} - \textstyle\sum p_f). \tag{4.86}$$

Unfortunately, the meaning of this formula is far from clear. Since an unstable particle cannot be sent into the infinitely distant past, our definition (4.73)

of $\mathcal{M}(m_{\mathcal{A}} \to \{p_f\})$ in terms of the S-matrix makes no sense in this context. Nevertheless formula (4.86) is correct, when $\mathcal{M}$ is computed according to the Feynman rules for S-matrix elements that we will present in the following section. We postpone the further discussion of these matters, and the proof of Eq. (4.86), until Section 7.3. Until then, an intuitive notion of $\mathcal{M}$ as a transition amplitude should suffice.

Equations (4.79) and (4.86) are completely general, whether or not the final state contains several identical particles. (The computation of $\mathcal{M}$, of course, will be quite different when identical particles are present, but that is another matter.) When integrating either of these formulae to obtain a *total* cross section or decay rate, however, we must be careful to avoid counting the same final state several times. If there are n identical particles in the final state, we must either restrict the integration to inequivalent configurations, or divide by $n!$ after integrating over all sets of momenta.

4.6 Computing S-Matrix Elements from Feynman Diagrams

Now that we have formulae for cross sections and decay rates in terms of the invariant matrix element $\mathcal{M}$, the only remaining task is to find a way of computing $\mathcal{M}$ for various processes in various interacting field theories. In this section we will write down (and try to motivate) a formula for $\mathcal{M}$ in terms of Feynman diagrams. We postpone the actual proof of this formula until Section 7.2, since the proof is somewhat technical and will be much easier to understand after we have seen how the formula is used.

Recall from its definition, Eq. (4.71), that the S-matrix is simply the time-evolution operator, $\exp(-iHt)$, in the limit of very large t:

$$\langle \mathbf{p}_1 \mathbf{p}_2 \cdots | \, S \, | \mathbf{k}_{\mathcal{A}} \mathbf{k}_{\mathcal{B}} \rangle = \lim_{T \to \infty} \langle \mathbf{p}_1 \mathbf{p}_2 \cdots | \, e^{-iH(2T)} \, | \mathbf{k}_{\mathcal{A}} \mathbf{k}_{\mathcal{B}} \rangle \, . \tag{4.87}$$

To compute this quantity we would like to replace the external plane-wave states in (4.87), which are eigenstates of H, with their counterparts in the unperturbed theory, which are eigenstates of H_0. We successfully made such a replacement for the vacuum state $|\Omega\rangle$ in Eq. (4.27):

$$|\Omega\rangle = \lim_{T \to \infty(1-i\epsilon)} \left(e^{-iE_0 T} \langle \Omega | \, 0 \rangle \right)^{-1} e^{-iHT} \, |0\rangle \, .$$

This time we would like to find a relation of the form

$$|\mathbf{k}_{\mathcal{A}} \mathbf{k}_{\mathcal{B}}\rangle \propto \lim_{T \to \infty(1-i\epsilon)} e^{-iHT} |\mathbf{k}_{\mathcal{A}} \mathbf{k}_{\mathcal{B}}\rangle_0, \tag{4.88}$$

where we have omitted some unknown phases and overlap factors like those in (4.27). To find such a relation would not be easy. In (4.27), we used the fact that the vacuum was the state of absolute lowest energy. Here we can use only the much weaker statement that the external states with well-separated initial and final particles have the lowest energy consistent with the predetermined

nonzero values of momentum. The problem is a deep one, and it is associated with one of the most fundamental difficulties of field theory, that interactions affect not only the scattering of distinct particles but also the form of the single-particle states themselves.

If the formula (4.88) could somehow be justified, we could use it to rewrite the right-hand side of (4.87) as

$$\begin{aligned} &\lim_{T\to\infty(1-i\epsilon)} {}_0\langle \mathbf{p}_1 \cdots \mathbf{p}_n | e^{-iH(2T)} | \mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle_0 \\ &\qquad \propto \lim_{T\to\infty(1-i\epsilon)} {}_0\langle \mathbf{p}_1 \cdots \mathbf{p}_n | T\left(\exp\left[-i\int_{-T}^{T} dt\, H_I(t)\right]\right) | \mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle_0 . \end{aligned} \tag{4.89}$$

In the evaluation of vacuum expectation values, the awkward proportionality factors between free and interacting vacuum states cancelled out of the final formula, Eq. (4.31). In the present case those factors are so horrible that we have not even attempted to write them down; we only hope that a similar dramatic cancellation will take place here. In fact such a cancellation does take place, although it is not easy to derive this conclusion from our present approach. Up to one small modification (which is unimportant for our present purposes), the formula for the nontrivial part of the S-matrix can be simplified to the following form:

$$\begin{aligned} &\langle \mathbf{p}_1 \cdots \mathbf{p}_n |\, iT \,| \mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle \\ &\quad = \lim_{T\to\infty(1-i\epsilon)} \left({}_0\langle \mathbf{p}_1 \cdots \mathbf{p}_n | T\left(\exp\left[-i\int_{-T}^{T} dt\, H_I(t)\right]\right) | \mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle_0 \right)_{\substack{\text{connected,}\\ \text{amputated}}} \end{aligned} \tag{4.90}$$

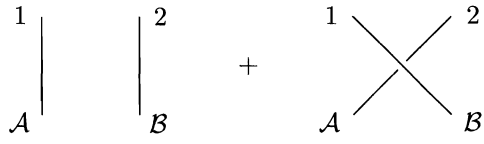
The attributes "connected" and "amputated" refer to restrictions on the class of possible Feynman diagrams; these terms will be defined in a moment. We will prove Eq. (4.90) in Section 7.2. In the remainder of this section, we will explain this formula and motivate the new restrictions that we have added.

First we must learn how to represent the matrix element in (4.90) as a sum of Feynman diagrams. Let us evaluate the first few terms explicitly, in ϕ^4 theory, for the case of two particles in the final state. The first term is

$$\begin{aligned} {}_0\langle \mathbf{p}_1\mathbf{p}_2 | \mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle_0 &= \sqrt{2E_1 2E_2 2E_\mathcal{A} 2E_\mathcal{B}}\, \langle 0|\, a_1 a_2 a_\mathcal{A}^\dagger a_\mathcal{B}^\dagger \,|0\rangle \\ &= 2E_\mathcal{A} 2E_\mathcal{B} (2\pi)^6 \Big(\delta^{(3)}(\mathbf{p}_\mathcal{A} - \mathbf{p}_1)\delta^{(3)}(\mathbf{p}_\mathcal{B} - \mathbf{p}_2) \\ &\qquad + \delta^{(3)}(\mathbf{p}_\mathcal{A} - \mathbf{p}_2)\delta^{(3)}(\mathbf{p}_\mathcal{B} - \mathbf{p}_1) \Big). \end{aligned} \tag{4.91}$$

The delta functions force the final state to be identical to the initial state, so this term is part of the '$\mathbf{1}$' in $S = \mathbf{1} + iT$, and does not contribute to the

scattering matrix element $\mathcal{M}$. We can represent it diagrammatically as



The next term in $\langle \mathbf{p}_1\mathbf{p}_2|\, S\, |\mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle$ is

$$\begin{aligned} &{}_0\langle \mathbf{p}_1\mathbf{p}_2|T\Big(-i\frac{\lambda}{4!}\int d^4x\,\phi_I^4(x)\Big)|\mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle_0 \\ &\quad = {}_0\langle \mathbf{p}_1\mathbf{p}_2|N\Big(-i\frac{\lambda}{4!}\int d^4x\,\phi_I^4(x) + \text{contractions}\Big)|\mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle_0, \end{aligned} \tag{4.92}$$

using Wick's theorem. Since the external states are not $|0\rangle$, terms that are not fully contracted do not necessarily vanish; we can use an annihilation operator from $\phi_I(x)$ to annihilate an initial-state particle, or a creation operator from $\phi_I(x)$ to produce a final-state particle. For example,

$$\begin{aligned} \phi_I^+(x)|\mathbf{p}\rangle_0 &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2E_\mathbf{k}}}\,a_\mathbf{k}e^{-ik\cdot x}\,\sqrt{2E_\mathbf{p}}a_\mathbf{p}^\dagger\,|0\rangle \\ &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2E_\mathbf{k}}}\,e^{-ik\cdot x}\sqrt{2E_\mathbf{p}}(2\pi)^3\delta^{(3)}(\mathbf{k}-\mathbf{p})\,|0\rangle \\ &= e^{-ip\cdot x}\,|0\rangle\,. \end{aligned} \tag{4.93}$$

An uncontracted ϕ_I operator inside the N-product of (4.92) has two terms: ϕ_I^+ on the far right and ϕ_I^- on the far left. We get one contribution to the S-matrix element for each way of commuting the a of ϕ_I^+ past an initial-state $a^\dagger$, and one contribution for each way of commuting the $a^\dagger$ of ϕ_I^- past a final-state a. It is natural, then, to define the contractions of field operators with external states as follows:

$$\overbracket{\phi_I(x)|\mathbf{p}\rangle} = e^{-ip\cdot x}\,|0\rangle\,; \qquad \overbracket{\langle\mathbf{p}|\phi_I(x)} = \langle 0|\,e^{+ip\cdot x}. \tag{4.94}$$

To evaluate an S-matrix element such as (4.92), we simply write down all possible full contractions of the ϕ_I operators *and* the external-state momenta.

To see that this prescription is correct, let us evaluate (4.92) in detail. The N-product contains terms of the form

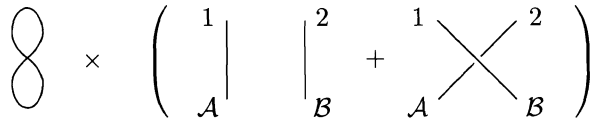
$$\phi\phi\phi\phi; \qquad \overbracket{\phi\phi\phi\phi}; \qquad \overbracket{\phi\phi}\,\overbracket{\phi\phi}. \tag{4.95}$$

The last term, in which the ϕ operators are fully contracted with each other, is

equal to a vacuum bubble diagram times the value of (4.91) calculated above:

$$-i\frac{\lambda}{4!}\int d^4x\,{}_0\langle \mathbf{p}_1\mathbf{p}_2|\overbrace{\phi\phi\phi\phi}|\mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle_0$$

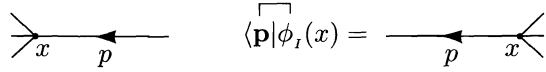
$$= \quad \times \quad \left(\quad + \quad \right) \tag{4.96}$$



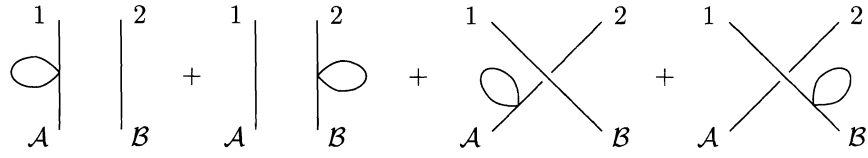
This is just another contribution to the trivial part of the S-matrix, so we ignore it.

Next consider the second term of (4.95), in which two of the four ϕ operators are contracted. The normal-ordered product of the remaining two fields looks like $(a^\dagger a^\dagger + 2a^\dagger a + aa)$. As we commute these operators past the a's and $a^\dagger$'s of the initial and final states, we find that only a term with an equal number of a's and $a^\dagger$'s can survive. In the language of contractions, this says that one of the ϕ's must be contracted with an initial-state $|\mathbf{p}\rangle$, the other with a final-state $\langle\mathbf{p}|$. The uncontracted $|\mathbf{p}\rangle$ and $\langle\mathbf{p}|$ give a delta function as in (4.91). To represent these quantities diagrammatically, we introduce *external lines* to our Feynman rules:

$$\overbrace{\phi_I(x)|\mathbf{p}\rangle} = \qquad \overbrace{\langle\mathbf{p}|\phi_I(x)} = \tag{4.97}$$



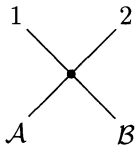
Feynman diagrams for S-matrix elements will always contain external lines, rather than the external points of diagrams for correlation functions. The second term of (4.95) thus yields four diagrams:



The integration $\int d^4x$ produces a momentum-conserving delta function at each vertex (including the external momenta), so these diagrams again describe trivial processes in which the initial and final states are identical. This illustrates a general principle: Only *fully connected* diagrams, in which all external lines are connected to each other, contribute to the T-matrix.

Finally, consider the term of (4.95) in which none of the ϕ operators are contracted with each other. Our prescription tells us to contract two of the ϕ's with $|\mathbf{p}_\mathcal{A}\mathbf{p}_\mathcal{B}\rangle$ and the other two with $\langle\mathbf{p}_1\mathbf{p}_2|$. There are 4! ways to do this.

Thus we obtain the diagram



$$= (4!) \cdot \left(-i\frac{\lambda}{4!}\right) \int d^4x\, e^{-i(p_\mathcal{A}+p_\mathcal{B}-p_1-p_2)\cdot x} \tag{4.98}$$

$$= -i\lambda\,(2\pi)^4 \delta^{(4)}(p_\mathcal{A} + p_\mathcal{B} - p_1 - p_2).$$

This is exactly of the form $i\mathcal{M}(2\pi)^4\delta^{(4)}(p_\mathcal{A} + p_\mathcal{B} - p_1 - p_2)$, with $\mathcal{M} = -\lambda$.

Before continuing our discussion of Feynman diagrams for S-matrix elements, we should certainly pause to turn this result into a cross section. For scattering in the center-of-mass frame, we can simply plug $|\mathcal{M}|^2 = \lambda^2$ into Eq. (4.85) to obtain

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{\lambda^2}{64\pi^2 E_{\text{cm}}^2}. \tag{4.99}$$

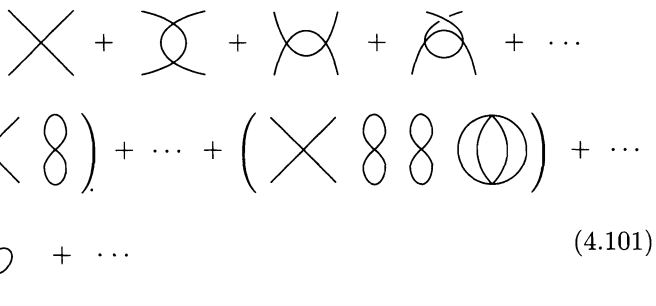
We have just computed our first quantum field theory cross section. It is a rather dull result, having no angular dependence at all. (This situation will be remedied when we consider fermions in the next section.) Integrating over $d\Omega$, and dividing by 2 since there are two identical particles in the final state, we find the total cross section,

$$\sigma_{\text{total}} = \frac{\lambda^2}{32\pi E_{\text{cm}}^2}. \tag{4.100}$$

In practice, one would probably use this result to measure the value of λ.

Returning to our general discussion, let us consider some higher-order contributions to the T-matrix for the process $\mathcal{A}, \mathcal{B} \to 1, 2$. If we ignore, for the moment, the "connected and amputated" prescription, we have the formula

$\langle \mathbf{p}_1 \mathbf{p}_2 | \, iT \, | \mathbf{p}_\mathcal{A} \mathbf{p}_\mathcal{B} \rangle \stackrel{?}{=}$



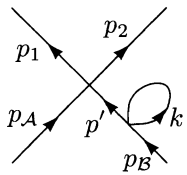
(4.101)

plus diagrams in which the four external lines are not all connected to each other. We have already seen that this last class of diagrams gives no contribution to the T-matrix. The first diagram shown in (4.101) gives the lowest-order contribution to T, which we calculated above. The next three diagrams give

expected corrections to this amplitude, involving creation and annihilation of additional "virtual" particles.

The diagrams in the second line of (4.101) contain disconnected "vacuum bubbles". By the same argument as at the end of Section 4.4, the disconnected pieces exponentiate to an overall phase factor giving the shift of the energy of the interacting vacuum state upon which the scattering takes place. Thus they are irrelevant to S. We have now seen that only fully connected diagrams give sensible contributions to S-matrix elements.

The last diagram is more problematical; let us evaluate it. After integrating over the two vertex positions, we obtain

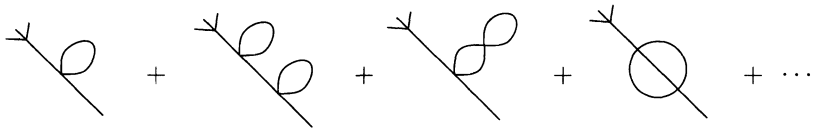


$$= \frac{1}{2}\int \frac{d^4p'}{(2\pi)^4}\frac{i}{p'^2-m^2}\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2} \times (-i\lambda)(2\pi)^4\delta^{(4)}(p_{\mathcal{A}}+p'-p_1-p_2) \times (-i\lambda)(2\pi)^4\delta^{(4)}(p_{\mathcal{B}}-p'). \tag{4.102}$$

We can integrate over p' using the second delta function. It tells us to evaluate

$$\frac{1}{p'^2-m^2}\Big|_{p'=p_{\mathcal{B}}} = \frac{1}{p_{\mathcal{B}}^2-m^2} = \frac{1}{0}.$$

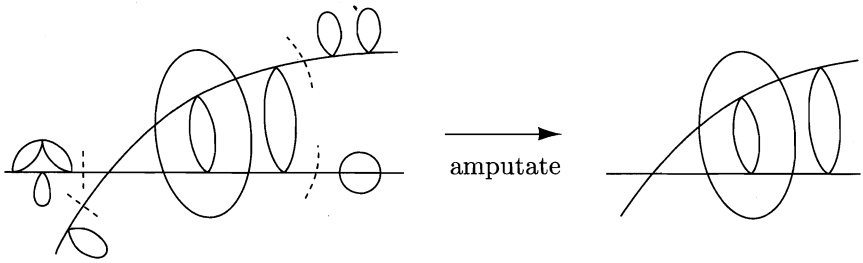
We get infinity, since $p_{\mathcal{B}}$, being the momentum of an external particle, is on-shell: $p_{\mathcal{B}}^2 = m^2$. This is a disaster. Clearly, our formula for S makes sense only if we exclude diagrams of this form, that is, diagrams with loops connected to only one external leg. Fortunately, this is physically reasonable: In the same way that the vacuum bubble diagrams represent the evolution of $|0\rangle$ into $|\Omega\rangle$, these external leg corrections,



represent the evolution of $|\mathbf{p}\rangle_0$ into $|\mathbf{p}\rangle$, the single-particle state of the interacting theory. Since these corrections have nothing to do with the scattering process, we should exclude them from the computation of S.

For a general diagram with external legs, we define *amputation* in the following way. Starting from the tip of each external leg, find the last point at which the diagram can be cut by removing a single propagator, such that this operation separates the leg from the rest of the diagram. Cut there. For

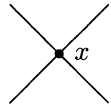
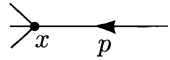
example:



Let us summarize our prescription for calculating scattering amplitudes. Our formula for S-matrix elements, Eq. (4.90), can be rewritten

$$i\mathcal{M} \cdot (2\pi)^4 \delta^{(4)}(p_A + p_B - \sum p_f) = \left(\begin{array}{c} \text{sum of all connected, amputated Feynman} \\ \text{diagrams with } p_A,\ p_B \text{ incoming, } p_f \text{ outgoing} \end{array} \right). \tag{4.103}$$

By 'connected', we now mean fully connected, that is, with no vacuum bubbles, and all external legs connected to each other. The Feynman rules for scattering amplitudes in ϕ^4 theory are, in position space,

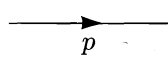
1. For each propagator,  $= D_F(x-y)$;
2. For each vertex,  $= (-i\lambda)\int d^4x$;
3. For each external line,  $= e^{-ip\cdot x}$;
4. Divide by the symmetry factor.

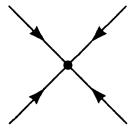
Notice that the factor for an ingoing line is just the amplitude for that particle to be found at the vertex it connects to, i.e., the particle's wavefunction. Similarly, the factor for an outgoing line is the amplitude for a particle produced at the vertex to have the desired final momentum.

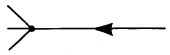
Just as with the Feynman rules for correlation functions, it is usually simpler to introduce the momentum-space representation of the propagators, carry out the vertex integrals to obtain momentum-conserving delta functions, and use these delta functions to evaluate as many momentum integrals as possible. In a scattering amplitude, however, there will always be an overall delta function, which can be used to cancel the one on the left-hand side of Eq. (4.103). We are then left with

$$i\mathcal{M} = \text{sum of all connected, amputated diagrams,} \tag{4.104}$$

where the diagrams are evaluated according to the following rules:

1. For each propagator,  $= \dfrac{i}{p^2 - m^2 + i\epsilon}$;

2. For each vertex,  $= -i\lambda$;

3. For each external line,  $= 1$;

4. Impose momentum conservation at each vertex;

5. Integrate over each undetermined loop momentum: $\displaystyle\int \frac{d^4p}{(2\pi)^4}$;

6. Divide by the symmetry factor.

This is our final version of the Feynman rules for ϕ^4 theory; these rules are also listed in the Appendix, for reference.

Actually, Eq. (4.103) still isn't quite correct. One more modification is necessary, involving the proportionality factors that were omitted from Eq. (4.89). But the modification affects only diagrams containing loops, so we postpone its discussion until Chapters 6 and 7, where we first evaluate such diagrams. We will prove the corrected formula (4.103) in Section 7.2, by relating S-matrix elements to correlation functions, for which we have actually derived a formula in terms of Feynman diagrams.

4.7 Feynman Rules for Fermions

So far in this chapter we have discussed only ϕ^4 theory, in order to avoid unnecessary complication. We are now ready to generalize our results to theories containing fermions.

Our treatment of correlation functions in Section 4.2 generalizes without difficulty. Lorentz invariance requires that the interaction Hamiltionian H_I be a product of an even number of spinor fields, so no difficulties arise in defining the time-ordered exponential of H_I.

To apply Wick's theorem, however, we must generalize the definitions of the time-ordering and normal-ordering symbols to include fermions. We saw at the end of Section 3.5 that the time-ordering operator T acting on two spinor fields is most conveniently defined with an additional minus sign:

$$T\big(\psi(x)\bar{\psi}(y)\big) \equiv \begin{cases} \psi(x)\bar{\psi}(y) & \text{for } x^0 > y^0; \\ -\bar{\psi}(y)\psi(x) & \text{for } x^0 < y^0. \end{cases} \tag{4.105}$$

With this definition, the Feynman propagator for the Dirac field is

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)} = \langle 0|\, T\psi(x)\bar{\psi}(y)\, |0\rangle\,. \tag{4.106}$$

For products of more than two spinor fields, we generalize this definition in the natural way: The time-ordered product picks up one minus sign for each interchange of operators that is necessary to put the fields in time order. For example,

$$T(\psi_1\psi_2\psi_3\psi_4) = (-1)^3\psi_3\psi_1\psi_4\psi_2 \qquad \text{if } x_3^0 > x_1^0 > x_4^0 > x_2^0.$$

The definition of the normal-ordered product of spinor fields is analogous: Put in an extra minus sign for each fermion interchange. The anticommutation properties make it possible to write a normal-ordered product in several ways, but with our conventions these are completely equivalent:

$$N(a_{\mathbf{p}}a_{\mathbf{q}}a_{\mathbf{r}}^\dagger) = (-1)^2 a_{\mathbf{r}}^\dagger a_{\mathbf{p}}a_{\mathbf{q}} = (-1)^3 a_{\mathbf{r}}^\dagger a_{\mathbf{q}}a_{\mathbf{p}}.$$

Using these definitions, it is not hard to generalize Wick's theorem. Consider first the case of two Dirac fields, say $T[\psi(x)\bar{\psi}(y)]$. In analogy with (4.37), define the contraction of two fields by

$$T[\psi(x)\bar{\psi}(y)] = N[\psi(x)\bar{\psi}(y)] + \overbracket{\psi(x)\bar{\psi}(y)}. \tag{4.107}$$

Explicitly, for the Dirac field,

$$\overbracket{\psi(x)\bar{\psi}(y)} \equiv \left\{ \begin{array}{ll} \{\psi^+(x), \bar{\psi}^-(y)\} & \text{for } x^0 > y^0 \\ -\{\bar{\psi}^+(y), \psi^-(x)\} & \text{for } x^0 < y^0 \end{array} \right\} = S_F(x-y); \tag{4.108}$$

$$\overbracket{\psi(x)\psi(y)} = \overbracket{\bar{\psi}(x)\bar{\psi}(y)} = 0. \tag{4.109}$$

Define contractions under the normal-ordering symbol to include minus signs for operator interchanges:

$$N(\overbracket{\psi_1\psi_2}\bar{\psi}_3\bar{\psi}_4) = -\overbracket{\psi_1\bar{\psi}_3}\, N(\psi_2\bar{\psi}_4) = -S_F(x_1 - x_3)\, N(\psi_2\bar{\psi}_4). \tag{4.110}$$

With these conventions, Wick's theorem takes the same form as before:

$$T[\psi_1\bar{\psi}_2\psi_3\cdots] = N[\psi_1\bar{\psi}_2\psi_3\cdots + \text{all possible contractions}]. \tag{4.111}$$

The proof is essentially unchanged from the bosonic case, since all extra minus signs are accounted for by the above definitions.

Yukawa Theory

Writing down the Feynman rules for fermion correlation functions would now be easy, but instead let's press on and discuss scattering processes. For definiteness, we begin by analyzing the Yukawa theory:

$$H = H_{\text{Dirac}} + H_{\text{Klein-Gordon}} + \int d^3x\, g\bar{\psi}\psi\phi. \tag{4.112}$$

This is a simplified model of Quantum Electrodynamics. In this section we will carefully work out the rules of calculation for Yukawa theory, so that in the next section we can guess the rules for QED without too much difficulty.

To be even more specific, consider the two-particle scattering reaction

$$\text{fermion}(p) + \text{fermion}(k) \longrightarrow \text{fermion}(p') + \text{fermion}(k').$$

The leading contribution comes from the H_I^2 term of the S-matrix:

$${}_0\langle \mathbf{p}', \mathbf{k}' | \, T\Big(\frac{1}{2!}(-ig)\int d^4x\, \bar{\psi}_I\psi_I\phi_I \;\; (-ig)\int d^4y\, \bar{\psi}_I\psi_I\phi_I\Big)|\mathbf{p}, \mathbf{k}\rangle_0. \tag{4.113}$$

To evaluate this expression, use Wick's theorem to reduce the T-product to an N-product of contractions, then act the uncontracted fields on the initial- and final-state particles. Represent this latter process as the contraction

$$\begin{aligned}\overbrace{\psi_I(x)|\mathbf{p}, s\rangle} &= \int \frac{d^3p'}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}'}}}\sum_{s'} a^{s'}_{\mathbf{p}'}u^{s'}(p')e^{-ip'\cdot x}\sqrt{2E_{\mathbf{p}}}a^{s\dagger}_{\mathbf{p}}\,|0\rangle \\ &= e^{-ip\cdot x}u^s(p)\,|0\rangle\,.\end{aligned} \tag{4.114}$$

Similar expressions hold for the contraction of $\bar{\psi}_I$ with a final-state fermion, and for contractions of ψ_I and $\bar{\psi}_I$ with antifermion states. Note that ψ_I can be contracted with a fermion on the right or an antifermion on the left; the opposite is true for $\bar{\psi}_I$.

We can write a typical contribution to the matrix element (4.113) as the contraction

$$\langle \mathbf{p}', \mathbf{k}' | \tfrac{1}{2!}(-ig)\textstyle\int d^4x\, \bar{\psi}\psi\phi \;\; (-ig)\int d^4y\, \bar{\psi}\psi\phi\, |\mathbf{p},\mathbf{k}\rangle. \tag{4.115}$$

Up to a possible minus sign, the value of this quantity is

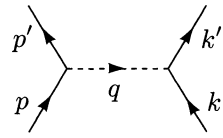
$$\begin{aligned}(-ig)^2\int \frac{d^4q}{(2\pi)^4}\frac{i}{q^2-m_\phi^2}&(2\pi)^4\delta^{(4)}(p'-p+q) \\ &\times (2\pi)^4\delta^{(4)}(k'-k-q)\bar{u}(p')u(p)\bar{u}(k')u(k).\end{aligned}$$

(We have dropped the factor 1/2! because there is a second, identical term that comes from interchanging x and y in (4.115).) Using either delta function to perform the integral, we find that this expression takes the form $i\mathcal{M}(2\pi)^4\delta^{(4)}(\Sigma p)$, with

$$i\mathcal{M} = \frac{-ig^2}{q^2-m_\phi^2}\bar{u}(p')u(p)\bar{u}(k')u(k). \tag{4.116}$$

When writing it in this way, we must remember to impose the constraints $p - p' = q = k' - k$.

Instead of working from (4.115), we could draw a Feynman diagram:



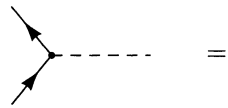
We denote scalar particles by dashed lines, and fermions by solid lines. The S-matrix element could then be obtained directly from the following momentum-space Feynman rules.

1. Propagators:

$$\overbrace{\phi(x)\phi(y)} = \quad \underset{q}{\text{- - - - ▸ - - - -}} \quad = \frac{i}{q^2 - m_\phi^2 + i\epsilon}$$

$$\overbrace{\psi(x)\bar{\psi}(y)} = \quad \underset{p}{\longrightarrow} \quad = \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}$$

2. Vertices:



3. External leg contractions:

$$\overbrace{\phi\,|\mathbf{q}\rangle} = \underset{q}{\text{- - ◂ - -}} = 1 \qquad \overbrace{\langle\mathbf{q}|\,\phi} = \underset{q}{\text{- - ◂ - -}} = 1$$

$$\overbrace{\psi\,\underbrace{|\mathbf{p},s\rangle}_{\text{fermion}}} = \underset{p}{\longleftarrow} = u^s(p) \qquad \overbrace{\underbrace{\langle\mathbf{p},s|}_{\text{fermion}}\,\bar{\psi}} = \underset{p}{\longleftarrow} = \bar{u}^s(p)$$

$$\overbrace{\bar{\psi}\,\underbrace{|\mathbf{k},s\rangle}_{\text{antifermion}}} = \underset{k}{\longrightarrow} = \bar{v}^s(k) \qquad \overbrace{\underbrace{\langle\mathbf{k},s|}_{\text{antifermion}}\,\psi} = \underset{k}{\longrightarrow} = v^s(k)$$

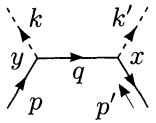
4. Impose momentum conservation at each vertex.
5. Integrate over each undetermined loop momentum.
6. Figure out the overall sign of the diagram.

Several comments are in order regarding these rules.

First, note that the $1/n!$ from the Taylor series of the time-ordered exponential is always canceled by the $n!$ ways of interchanging vertices to obtain the same contraction. The diagrams of Yukawa theory never have symmetry factors, since the three fields $(\bar{\psi}\psi\phi)$ in H_I cannot substitute for one another in contractions.

Second, the direction of the momentum on a fermion line is always significant. On external lines, as for bosons, the direction of the momentum is always

ingoing for initial-state particles and outgoing for final-state particles. This follows immediately from the expansions of ψ and $\bar{\psi}$, where the annihilation operators $a_{\mathbf{p}}$ and $b_{\mathbf{p}}$ both multiply $e^{-ip\cdot x}$ and the creation operators $a^\dagger_{\mathbf{p}}$ and $b^\dagger_{\mathbf{p}}$ both multiply $e^{+ip\cdot x}$. On internal fermion lines (propagators), the momentum must be assigned in the direction of particle-number flow (for electrons, this is the direction of negative charge flow). This requirement is most easily seen by working out an example from first principles. Consider the annihilation of a fermion and an antifermion into two bosons:



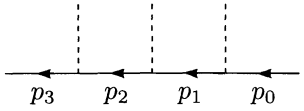
$$= \langle \mathbf{k},\mathbf{k}'| \int d^4x\, \phi\bar{\psi}\psi \int d^4y\, \phi\bar{\psi}\psi\, |\mathbf{p},\mathbf{p}'\rangle$$

$$\sim \int d^4x \int d^4y\, e^{ik'\cdot x}\, \bar{v}(p') e^{-ip'\cdot x} \int \frac{d^4q}{(2\pi)^4} \frac{i(\not{q}+m)}{q^2-m^2} e^{-iq\cdot(x-y)}\, u(p) e^{-ip\cdot y}\, e^{ik\cdot y}.$$

The integrals over x and y give delta functions that force q to flow from y to x, as shown. On internal boson lines the direction of the momentum is irrelevant and may be chosen for convenience, since $D_F(x-y) = D_F(y-x)$.

It is conventional to draw arrows on fermion lines, as shown, to represent the direction of particle-number flow. The momentum assigned to a fermion propagator then flows in the direction of this arrow. For external antiparticles, however, the momentum flows opposite to the arrow; it helps to show this explicitly by drawing a second arrow next to the line.

Third, note that in our examples the Dirac indices contract together along the fermion lines. This will also happen in more complicated diagrams:



$$\sim \bar{u}(p_3) \cdot \frac{i(\not{p}_2 + m)}{p_2^2 - m^2} \cdot \frac{i(\not{p}_1 + m)}{p_1^2 - m^2} \cdot u(p_0). \qquad (4.117)$$

Finally, let's take a moment to worry about fermion minus signs. Return to the example of the fermion-fermion scattering process. We adopt a sign convention for the initial and final states:

$$|\mathbf{p},\mathbf{k}\rangle \sim a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{k}} |0\rangle\,, \qquad \langle \mathbf{p}',\mathbf{k}'| \sim \langle 0|\, a_{\mathbf{k}'} a_{\mathbf{p}'}\,, \qquad (4.118)$$

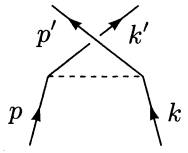
so that $(|p,k\rangle)^\dagger = \langle p,k|$. Then the contraction

$$\langle \mathbf{p}',\mathbf{k}'|(\bar{\psi}\psi)_x\,(\bar{\psi}\psi)_y|\mathbf{p},\mathbf{k}\rangle \sim \langle 0|\, a_{\mathbf{k}'} a_{\mathbf{p}'}\, \bar{\psi}_x\psi_x\, \bar{\psi}_y\psi_y\, a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{k}}\, |0\rangle$$

can be untangled by moving $\bar{\psi}_y$ two spaces to the left, and so picks up a factor of $(-1)^2 = +1$. But note that in the contraction

$$\langle \mathbf{p}',\mathbf{k}'|(\bar{\psi}\psi)_x\,(\bar{\psi}\psi)_y|\mathbf{p},\mathbf{k}\rangle \sim \langle 0|\, a_{\mathbf{k}'} a_{\mathbf{p}'}\, \bar{\psi}_x\psi_x\, \bar{\psi}_y\psi_y\, a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{k}}\, |0\rangle\,,$$

it is sufficient to move the $\bar{\psi}_y$ one space to the left, giving a factor of -1. This contraction corresponds to the diagram



The full result, to lowest order, for the S-matrix element for this process is therefore

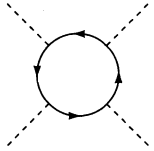
$$
\begin{aligned}
i\mathcal{M} = {} & \\
& = (-ig^2)\left(\bar{u}(p')u(p)\frac{1}{(p'-p)^2 - m_\phi^2}\bar{u}(k')u(k)\right. \\
& \qquad \left. - \bar{u}(p')u(k)\frac{1}{(p'-k)^2 - m_\phi^2}\bar{u}(k')u(p)\right).
\end{aligned}
\tag{4.119}
$$

The minus sign difference between these diagrams is a reflection of Fermi statistics. Turning this expression into an explicit cross section would require some additional work; we postpone such calculations until Chapter 5, when we can work with QED instead of the less interesting Yukawa theory.

In complicated diagrams, one can often simplify the determination of the minus signs by noting that the product $(\bar{\psi}\psi)$, or any other pair of fermions, *commutes* with any operator. Thus,

$$
\begin{aligned}
\cdots(\bar{\psi}\psi)_x(\bar{\psi}\psi)_y(\bar{\psi}\psi)_z(\bar{\psi}\psi)_w\cdots &= \cdots(+1)(\bar{\psi}\psi)_x(\bar{\psi}\psi)_z(\bar{\psi}\psi)_y(\bar{\psi}\psi)_w\cdots \\
&= \cdots S_F(x-z)S_F(z-y)S_F(y-w)\cdots .
\end{aligned}
$$

But note that in a closed loop of n fermion propagators we have



$$
\begin{aligned}
&= \bar{\psi}\psi\ \bar{\psi}\psi\ \bar{\psi}\psi\ \bar{\psi}\psi \\
&= (-1)\ \mathrm{tr}\big[\psi\ \bar{\psi}\psi\ \bar{\psi}\psi\ \bar{\psi}\psi\ \bar{\psi}\big] \\
&= (-1)\ \mathrm{tr}\big[S_F\, S_F\, S_F\, S_F\big].
\end{aligned}
\tag{4.120}
$$

A closed fermion loop always gives a factor of -1 and the *trace* of a product of Dirac matrices.

The Yukawa Potential

We now have all the formal rules we need to compute scattering amplitudes in Yukawa theory. Before going on to discuss QED, let us briefly descend from abstraction to concrete physics, and consider one very simple application of these rules: the scattering of *distinguishable* fermions, in the nonrelativistic limit. By comparing the amplitude for this process to the Born approximation formula from nonrelativistic quantum mechanics, we can determine the potential $V(r)$ created by the Yukawa interaction.

If the two interacting particles are distinguishable, only the first diagram in (4.119) contributes. To evaluate the amplitude in the nonrelativistic limit, we keep terms only to lowest order in the 3-momenta. Thus, up to $\mathcal{O}(\mathbf{p}^2, \mathbf{p}'^2, \ldots)$,

$$\begin{aligned} p &= (m, \mathbf{p}), & k &= (m, \mathbf{k}), \\ p' &= (m, \mathbf{p}'), & k' &= (m, \mathbf{k}'). \end{aligned} \tag{4.121}$$

Using these expressions, we have

$$(p' - p)^2 = -|\mathbf{p}' - \mathbf{p}|^2 + \mathcal{O}(\mathbf{p}^4),$$

$$u^s(p) = \sqrt{m}\begin{pmatrix} \xi^s \\ \xi^s \end{pmatrix}, \quad \text{etc.},$$

where ξ^s is a two-component constant spinor normalized to $\xi^{s'\dagger}\xi^s = \delta^{ss'}$. The spinor products in (4.119) are then

$$\begin{aligned} \bar{u}^{s'}(p')u^s(p) &= 2m\xi^{s'\dagger}\xi^s = 2m\delta^{ss'}; \\ \bar{u}^{r'}(k')u^r(k) &= 2m\xi^{r'\dagger}\xi^r = 2m\delta^{rr'}. \end{aligned} \tag{4.122}$$

So our first physical conclusion is that the spin of each particle is separately conserved in this nonrelativistic scattering interaction—a pleasing result.

Putting together the pieces of the scattering amplitude (4.119), we find

$$i\mathcal{M} = \frac{ig^2}{|\mathbf{p}' - \mathbf{p}|^2 + m_\phi^2} 2m\delta^{ss'} 2m\delta^{rr'}. \tag{4.123}$$

This should be compared with the Born approximation to the scattering amplitude in nonrelativistic quantum mechanics, written in terms of the potential function $V(\mathbf{x})$:

$$\langle p'|iT|p\rangle = -i\widetilde{V}(\mathbf{q})\,(2\pi)\delta(E_{\mathbf{p}'} - E_{\mathbf{p}}), \qquad (\mathbf{q} = \mathbf{p}' - \mathbf{p}). \tag{4.124}$$

So apparently, for the Yukawa interaction,

$$\widetilde{V}(\mathbf{q}) = \frac{-g^2}{|\mathbf{q}|^2 + m_\phi^2}. \tag{4.125}$$

(The factors of $2m$ in (4.123) arise from our relativistic normalization conventions, and must be dropped when comparing to (4.124), which assumes conventional nonrelativistic normalization of states. The additional $\delta^{(3)}(\mathbf{p}' - \mathbf{p})$ goes away when we integrate over the momentum of the target.)

Inverting the Fourier transform to find $V(\mathbf{x})$ requires a short calculation:

$$\begin{aligned} V(\mathbf{x}) &= \int \frac{d^3q}{(2\pi)^3} \frac{-g^2}{|\mathbf{q}|^2 + m_\phi^2} e^{i\mathbf{q}\cdot\mathbf{x}} \\ &= \frac{-g^2}{4\pi^2} \int_0^\infty dq\, q^2 \frac{e^{iqr} - e^{-iqr}}{iqr} \frac{1}{q^2 + m_\phi^2} \\ &= \frac{-g^2}{4\pi^2 ir} \int_{-\infty}^{\infty} dq \frac{q\, e^{iqr}}{q^2 + m_\phi^2}. \end{aligned} \tag{4.126}$$

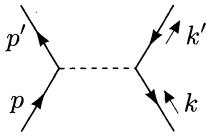
The contour of this integral can be closed above in the complex plane, and we pick up the residue of the simple pole at $q = +im_\phi$. Thus we find

$$V(r) = -\frac{g^2}{4\pi} \frac{1}{r} e^{-m_\phi r}, \tag{4.127}$$

an *attractive* "Yukawa potential", with range $1/m_\phi = \hbar/m_\phi c$, the Compton wavelength of the exchanged boson. Yukawa made this potential the basis for his theory of the nuclear force, and worked backwards from the range of the force (about 1 fm) to predict the mass (about 200 MeV) of the required boson, the pion.

What happens if instead we scatter particles off of *anti*particles? For the process

$$f_1(p)\bar{f}_2(k) \longrightarrow f_1(p')\bar{f}_2(k'),$$



we need to evaluate (nonrelativistically)

$$\bar{v}^s(k)v^{s'}(k') \approx m(\xi^{s\dagger}, -\xi^{s\dagger}) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \xi^{s'} \\ -\xi^{s'} \end{pmatrix} = -2m\delta^{ss'}. \tag{4.128}$$

We must also work out the fermion minus sign. Using $|\mathbf{p}, \mathbf{k}\rangle = a_\mathbf{p}^\dagger b_\mathbf{k}^\dagger |0\rangle$ and $\langle \mathbf{p}', \mathbf{k}'| = \langle 0| b_{\mathbf{k}'} a_{\mathbf{p}'}$, we can write the contracted matrix element as

$$\langle \mathbf{p}', \mathbf{k}'| \bar{\psi}\psi\ \bar{\psi}\psi\ |\mathbf{p}, \mathbf{k}\rangle = \langle 0| b_{\mathbf{k}'} a_{\mathbf{p}'}\ \bar{\psi}\psi\ \bar{\psi}\psi\ a_\mathbf{p}^\dagger b_\mathbf{k}^\dagger |0\rangle .$$

To untangle the contractions requires three operator interchanges, so there is an overall factor of -1. This cancels the extra minus sign in (4.128), and therefore we see that the Yukawa potential between a fermion and an antifermion is also attractive, and identical in strength to that between two fermions.

The remaining case to consider is scattering of two antifermions. It should not be surprising that the potential is again attractive; there is an additional minus sign from changing the other $\bar{u}u$ into $\bar{v}v$, and the number of interchanges necessary to untangle the contractions is even. Thus we conclude that the Yukawa potential is *universally attractive*, whether it is between a pair of fermions, a pair of antifermions, or one of each.

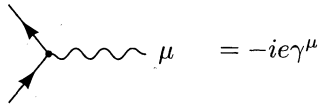
4.8 Feynman Rules for Quantum Electrodynamics

Now we are ready to step from Yukawa theory to Quantum Electrodynamics. To do this, we replace the scalar particle ϕ with a vector particle A_μ, and replace the Yukawa interaction Hamiltonian with

$$H_{\text{int}} = \int d^3x \, e\bar{\psi}\gamma^\mu\psi A_\mu. \tag{4.129}$$

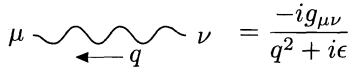
How do the Feynman rules change? The answer, though difficult to prove, is easy to guess. In addition to the fermion rules from the previous section, we have

New vertex:



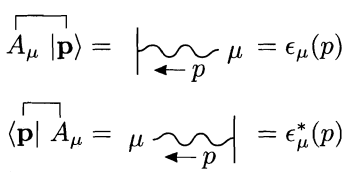
$= -ie\gamma^\mu$

Photon propagator:



$= \dfrac{-ig_{\mu\nu}}{q^2 + i\epsilon}$

External photon lines:



$= \epsilon_\mu(p)$

$= \epsilon^*_\mu(p)$

Photons are conventionally drawn as wavy lines. The symbol $\epsilon_\mu(p)$ stands for the *polarization vector* of the initial- or final-state photon.

To justify these rules, recall that in Lorentz gauge (which we employ to retain explicit relativistic invariance) the field equation for A_μ is

$$\partial^2 A_\mu = 0. \tag{4.130}$$

Thus each component of A separately obeys the Klein-Gordon equation (with $m = 0$). The momentum-space solutions of this equation are $\epsilon_\mu(p)e^{-ip\cdot x}$, where $p^2 = 0$ and $\epsilon_\mu(p)$ is any 4-vector. The interpretation of ϵ as the polarization vector of the field should be familiar from classical electromagnetism. If we expand the quantized electromagnetic field in terms of classical solutions of the wave equation, as we did for the Klein-Gordon field, we find

$$A_\mu(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_{r=0}^{3} \left(a^r_{\mathbf{p}} \epsilon^r_\mu(p) e^{-ip\cdot x} + a^{r\dagger}_{\mathbf{p}} \epsilon^{r*}_\mu(p) e^{ip\cdot x} \right), \tag{4.131}$$

where $r = 0, 1, 2, 3$ labels a basis of polarization vectors. The external line factors in the Feynman rules above follow immediately from this expansion, just as we obtained u's and v's as the external line factors for Dirac particles. The only subtlety is that we must restrict initial- and final-state photons to be transversely polarized: Their polarization vectors are always of the form $\epsilon^\mu = (0, \boldsymbol{\epsilon})$, where $\mathbf{p} \cdot \boldsymbol{\epsilon} = 0$. For $\mathbf{p}$ along the z-axis, the right- and left-handed polarization vectors are $\epsilon^\mu = (0, 1, \pm i, 0)/\sqrt{2}$.

The form of the QED vertex factor is also easy to justify, by simply looking at the interaction Hamiltonian (4.129). Note that the γ matrix in a QED amplitude will sit between spinors or other γ matrices, with the Dirac indices contracted along the fermion line. Note also that this interaction term is specific to the case of an electron (and its antiparticle, the positron). In general, for a Dirac particle with electric charge $Q|e|$,

$$= -iQ|e|\gamma^\mu.$$

For example, an electron has $Q = -1$, an up quark has $Q = +2/3$, and a down quark has $Q_d = -1/3$.

There is no easy way to derive the form of the photon propagator, so for now we will settle for a plausibility argument. Since the electromagnetic field in Lorentz gauge obeys the massless Klein-Gordon equation, it should come as no surprise that the photon propagator is nearly identical to the massless Klein-Gordon propagator. The factor of $-g_{\mu\nu}$, however, requires explanation. Lorentz invariance dictates that the photon propagator be an isotropic second-rank tensor that can dot together the γ^μ and γ^ν from the vertices at each end. The simplest candidate is $g^{\mu\nu}$. To understand the overall sign of the propagator, evaluate its Fourier transform:

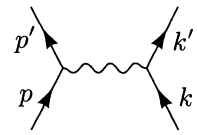
$$\int \frac{d^4q}{(2\pi)^4} \frac{-ig_{\mu\nu}}{q^2 + i\epsilon} e^{-iq\cdot(x-y)} = \int \frac{d^3q}{(2\pi)^3} \frac{1}{2|\mathbf{q}|} e^{-iq\cdot(x-y)} \cdot (-g_{\mu\nu}). \tag{4.132}$$

Presumably this is equal to $\langle 0|\, T\big[A_\mu(x) A_\nu(y)\big]\, |0\rangle$. Now set $\mu = \nu$, and take the limit $x^0 \to y^0$ from the positive direction. Then this quantity becomes the norm of the state $A_\mu(x)\,|0\rangle$, which should be positive. We see that our choice of signs in the propagator implies that the three states created by A_i, with $i = 1, 2, 3$, indeed have positive norm. These states include all real (non-virtual) photons, which always have spacelike polarizations. Unfortunately, because $g_{\mu\nu}$ is not positive definite, the states created by A_0 inevitably have negative norm. This is potentially a serious problem for any theory with vector particles. For Quantum Electrodynamics, we will show in Section 5.5 that the negative-norm states created by A_0 are never produced in physical processes. In Section 9.4 we will give a careful derivation of the photon propagator.

The Coulomb Potential

As a simple application of these Feynman rules, and to better understand the sign of the propagator, let us repeat the nonrelativistic scattering calculation of the previous section, this time for QED. The leading-order contribution is

$$i\mathcal{M} = \quad = (-ie)^2 \bar{u}(p')\gamma^\mu u(p) \frac{-ig_{\mu\nu}}{(p'-p)^2} \bar{u}(k')\gamma^\nu u(k). \quad (4.133)$$



In the nonrelativistic limit,

$$\bar{u}(p')\gamma^0 u(p) = u^\dagger(p')u(p) \approx +2m\xi'^\dagger\xi.$$

You can easily verify that the other terms, $\bar{u}(p')\gamma^i u(p)$, vanish if $p = p' = 0$; they can therefore be neglected compared to $\bar{u}(p')\gamma^0 u(p)$ in the nonrelativistic limit. Thus we have

$$\begin{aligned} i\mathcal{M} &\approx \frac{+ie^2}{-|\mathbf{p}'-\mathbf{p}|^2}(2m\xi'^\dagger\xi)_p(2m\xi'^\dagger\xi)_k \cdot g_{00} \\ &= \frac{-ie^2}{|\mathbf{p}'-\mathbf{p}|^2}(2m\xi'^\dagger\xi)_p(2m\xi'^\dagger\xi)_k. \end{aligned} \quad (4.134)$$

Comparing this to the Yukawa case (4.123), we see that there is an extra factor of -1; the potential is a *repulsive* Yukawa potential with $m = 0$, that is, a repulsive Coulomb potential:

$$V(r) = \frac{e^2}{4\pi r} = \frac{\alpha}{r}, \quad (4.135)$$

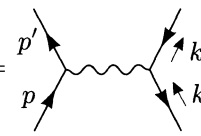
where $\alpha = e^2/4\pi \approx 1/137$ is the fine-structure constant.

For particle-antiparticle scattering, note first that

$$\bar{v}(k)\gamma^0 v(k') = v^\dagger(k)v(k') \approx +2m\xi^\dagger\xi'.$$

The presence of the γ^0 eliminates the minus sign that we found in the Yukawa case. The nonrelativistic scattering amplitude is therefore

$$i\mathcal{M} = \quad = (-1) \cdot \frac{-ie^2}{|\mathbf{p}'-\mathbf{p}|^2}(+2m\xi'^\dagger\xi)_p(+2m\xi^\dagger\xi')_k, \quad (4.136)$$



where the (-1) is the same fermion minus sign we saw in the Yukawa case. This is an *attractive* potential. Similarly, for antifermion-antifermion scattering one finds a repulsive potential. We have just verified that in quantum field theory, when a vector particle is exchanged, like charges repel while unlike charges attract.

Note that the repulsion in fermion-fermion scattering came entirely from the extra factor $-g_{00} = -1$ in the vector boson propagator. A tensor boson, such as the graviton, would have a propagator

$$\mu\,\nu \sim\sim\sim\sim \rho\,\sigma = \frac{1}{2}\Big((-g_{\mu\rho})(-g_{\nu\sigma}) + (-g_{\mu\sigma})(-g_{\nu\rho})\Big)\left(\frac{i}{q^2 + i\epsilon}\right),$$

which in nonrelativistic collisions gives a factor $(-g_{00})^2 = +1$; this will result in a universally attractive potential. It is reassuring to see that quantum field theory does indeed reproduce the obvious features of the electric and gravitational forces:

Exchanged particle	ff and $\bar{f}\bar{f}$	$f\bar{f}$
scalar (Yukawa)	attractive	attractive
vector (electricity)	repulsive	attractive
tensor (gravity)	attractive	attractive

Problems

4.1 Let us return to the problem of the creation of Klein-Gordon particles by a classical source. Recall from Chapter 2 that this process can be described by the Hamiltonian

$$H = H_0 + \int d^3x\,(-j(t,\mathbf{x})\phi(x)),$$

where H_0 is the free Klein-Gordon Hamiltonian, $\phi(x)$ is the Klein-Gordon field, and $j(x)$ is a c-number scalar function. We found that, if the system is in the vacuum state before the source is turned on, the source will create a mean number of particles

$$\langle N\rangle = \int \frac{d^3p}{(2\pi)^3}\,\frac{1}{2E_{\mathbf{p}}}\,|\tilde{j}(p)|^2.$$

In this problem we will verify that statement, and extract more detailed information, by using a perturbation expansion in the strength of the source.

(a) Show that the probability that the source creates *no* particles is given by

$$P(0) = \left|\langle 0|\,T\Big\{\exp[i\int d^4x\,j(x)\phi_I(x)]\Big\}\,|0\rangle\right|^2.$$

(b) Evaluate the term in P(0) of order j^2, and show that $P(0) = 1 - \lambda + \mathcal{O}(j^4)$, where λ equals the expression given above for $\langle N\rangle$.

(c) Represent the term computed in part (b) as a Feynman diagram. Now represent the whole pertubation series for $P(0)$ in terms of Feynman diagrams. Show that this series exponentiates, so that it can be summed exactly: $P(0) = \exp(-\lambda)$.

(d) Compute the probability that the source creates one particle of momentum k. Perform this computation first to $\mathcal{O}(j)$ and then to all orders, using the trick of part (c) to sum the series.

(e) Show that the probability of producing n particles is given by

$$P(n) = (1/n!)\lambda^n \exp(-\lambda).$$

This is a *Poisson distribution.*

(f) Prove the following facts about the Poisson distribution:

$$\sum_{n=0}^{\infty} P(n) = 1; \qquad \langle N \rangle = \sum_{n=0}^{\infty} n\, P(n) = \lambda.$$

The first identity says that the $P(n)$'s are properly normalized probabilities, while the second confirms our proposal for $\langle N \rangle$. Compute the mean square fluctuation $\left\langle (N - \langle N \rangle)^2 \right\rangle$.

4.2 Decay of a scalar particle. Consider the following Lagrangian, involving two real scalar fields Φ and ϕ:

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu \Phi)^2 - \tfrac{1}{2}M^2\Phi^2 + \tfrac{1}{2}(\partial_\mu \phi)^2 - \tfrac{1}{2}m^2\phi^2 - \mu\Phi\phi\phi.$$

The last term is an interaction that allows a Φ particle to decay into two ϕ's, provided that $M > 2m$. Assuming that this condition is met, calculate the lifetime of the Φ to lowest order in μ.

4.3 Linear sigma model. The interactions of pions at low energy can be described by a phenomenological model called the *linear sigma model.* Essentially, this model consists of N real scalar fields coupled by a ϕ^4 interaction that is symmetric under rotations of the N fields. More specifically, let $\Phi^i(x)$, $i = 1, \ldots, N$ be a set of N fields, governed by the Hamiltonian

$$H = \int d^3x \left(\tfrac{1}{2}(\Pi^i)^2 + \tfrac{1}{2}(\nabla\Phi^i)^2 + V(\Phi^2) \right),$$

where $(\Phi^i)^2 = \mathbf{\Phi} \cdot \mathbf{\Phi}$, and

$$V(\Phi^2) = \tfrac{1}{2}m^2(\Phi^i)^2 + \tfrac{\lambda}{4}((\Phi^i)^2)^2$$

is a function symmetric under rotations of $\mathbf{\Phi}$. For (classical) field configurations of $\Phi^i(x)$ that are constant in space and time, this term gives the only contribution to H; hence, V is the field potential energy.

(What does this Hamiltonian have to do with the strong interactions? There are two types of light quarks, u and d. These quarks have identical strong interactions, but different masses. If these quarks are massless, the Hamiltonian of the strong interactions is invariant to unitary transformations of the 2-component object (u, d):

$$\begin{pmatrix} u \\ d \end{pmatrix} \rightarrow \exp(i\boldsymbol{\alpha} \cdot \boldsymbol{\sigma}/2) \begin{pmatrix} u \\ d \end{pmatrix}.$$

This transformation is called an *isospin* rotation. If, in addition, the strong interactions are described by a vector "gluon" field (as is true in QCD), the strong interaction Hamiltonian is invariant to the isospin rotations done separately on the left-handed and right-handed components of the quark fields. Thus, the complete symmetry of QCD with two massless quarks is $SU(2) \times SU(2)$. It happens that $SO(4)$, the group of rotations in 4 dimensions, is isomorphic to $SU(2) \times SU(2)$, so for $N = 4$, the linear sigma model has the same symmetry group as the strong interactions.)