

Yorikiyo Nagashima

Elementary Particle Physics

Volume 1: Quantum Field Theory and Particles



WILEY-VCH Verlag GmbH & Co. KGaA

Yorikiyo Nagashima

Elementary Particle Physics

Related Titles

Nagashima, Y.

Elementary Particle Physics

Volume 2: Standard Model and Experiments

Approx. 2013

ISBN 978-3-527-40966-2

Russenschuck, S.

Field Computation for Accelerator Magnets

Analytical and Numerical Methods for Electromagnetic Design and Optimization

2010

ISBN 978-3-527-40769-9

Stock, R. (ed.)

Encyclopedia of Applied High Energy and Particle Physics

2009

ISBN 978-3-527-40691-3

Griffiths, D.

Introduction to Elementary Particles

2008

ISBN 978-3-527-40601-2

Belusevic, R.

Relativity, Astrophysics and Cosmology

2008

ISBN 978-3-527-40764-4

Reiser, M.

Theory and Design of Charged Particle Beams

2008

ISBN 978-3-527-40741-5

Yorikiyo Nagashima

Elementary Particle Physics

Volume 1: Quantum Field Theory and Particles



WILEY-VCH Verlag GmbH & Co. KGaA

The Author

Yorikiyo Nagashima

Osaka University
Japan
nagashimayori@ybb.ne.jp

Cover Image

Japanese symbol that denotes
“void” or “nothing”; also
symbolizes “supreme state of
matter” or “spirit”.

■ All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data:

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

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

© 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Typesetting le-tex publishing services GmbH, Leipzig

Printing and Binding betz-druck GmbH, Darmstadt

Cover Design Adam Design, Weinheim

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

ISBN 978-3-527-40962-4

Foreword

This is a unique book. It covers the entire theoretical and experimental content of particle physics.

Particle physics sits at the forefront of our search for the ultimate structure of matter at the smallest scale, but in the process it has also learned to question the nature of our space and time in which they exist. Going hand in hand with technological advances, particle physics now has extended its reach to studies of the structure and the history of the universe. It seems that both the ultimate small and the ultimate large are linked together. To explore one you must also explore the other.

Thus particle physics covers a vast area. To master it, you usually have to read different books, at increasingly advanced levels, on individual subjects like its historical background, basic experimental data, quantum field theory, mathematical concepts, theoretical models, cosmological concepts, etc. This book covers most of those topics in a single volume in an integrated manner. Not only that, it shows you how to derive each important mathematical formula in minute detail, then asks you to work out many problems, with answers given at the end of the book. Some abstract formulas are immediately followed by their intuitive interpretation and experimental consequences. The same topics are often repeated at different levels of sophistication in different chapters as you read on, which will help deepen your understanding.

All these features are quite unique to this book, and will be most helpful to students as well as laymen or non-experts who want to learn the subject seriously and enjoy it. It can serve both as a text book and as a compendium on particle physics. Even for practicing particle physicists and professors this will be a valuable reference book to keep at hand. Few people like Professor Nagashima, an accomplished experimental physicist who is also conversant with sophisticated theoretical subjects, could have written it.

Chicago, October 2009

Yoichiro Nambu

Contents

Foreword V

Preface XVII

Acknowledgements XXI

Part One A Field Theoretical Approach 1

1 Introduction 3

1.1 An Overview of the Standard Model 3

1.1.1 What is an Elementary Particle? 3

1.1.2 The Four Fundamental Forces and Their Unification 4

1.1.3 The Standard Model 7

1.2 The Accelerator as a Microscope 11

2 Particles and Fields 13

2.1 What is a Particle? 13

2.2 What is a Field? 21

2.2.1 Force Field 21

2.2.2 Relativistic Wave Equation 25

2.2.3 Matter Field 27

2.2.4 Intuitive Picture of a Field and Its Quantum 28

2.2.5 Mechanical Model of a Classical Field 29

2.3 Summary 32

2.4 Natural Units 33

3 Lorentz Invariance 37

3.1 Rotation Group 37

3.2 Lorentz Transformation 41

3.2.1 General Formalism 41

3.2.2 Lorentz Vectors and Scalars 43

3.3 Space Inversion and Time Reversal 45

3.4 Covariant Formalism 47

3.4.1 Tensors 47

3.4.2 Covariance 48

3.4.3 Supplementing the Time Component 49

3.4.4	Rapidity	51
3.5	Lorentz Operator	53
3.6	Poincaré Group*	56
4	Dirac Equation	59
4.1	Relativistic Schrödinger Equation	59
4.1.1	Dirac Matrix	59
4.1.2	Weyl Spinor	61
4.1.3	Interpretation of the Negative Energy	64
4.1.4	Lorentz-Covariant Dirac Equation	69
4.2	Plane-Wave Solution	71
4.3	Properties of the Dirac Particle	75
4.3.1	Magnetic Moment of the Electron	75
4.3.2	Parity	77
4.3.3	Bilinear Form of the Dirac Spinor	78
4.3.4	Charge Conjugation	79
4.3.5	Chiral Eigenstates	82
4.4	Majorana Particle	84
5	Field Quantization	89
5.1	Action Principle	89
5.1.1	Equations of Motion	89
5.1.2	Hamiltonian Formalism	90
5.1.3	Equation of a Field	91
5.1.4	Noether's Theorem	95
5.2	Quantization Scheme	100
5.2.1	Heisenberg Equation of Motion	100
5.2.2	Quantization of the Harmonic Oscillator	102
5.3	Quantization of Fields	105
5.3.1	Complex Fields	106
5.3.2	Real Field	111
5.3.3	Dirac Field	112
5.3.4	Electromagnetic Field	114
5.4	Spin and Statistics	119
5.5	Vacuum Fluctuation	121
5.5.1	The Casimir Effect*	122
6	Scattering Matrix	127
6.1	Interaction Picture	127
6.2	Asymptotic Field Condition	131
6.3	Explicit Form of the S-Matrix	133
6.3.1	Rutherford Scattering	135
6.4	Relativistic Kinematics	136
6.4.1	Center of Mass Frame and Laboratory Frame	136
6.4.2	Crossing Symmetry	139
6.5	Relativistic Cross Section	141

6.5.1	Transition Rate	141
6.5.2	Relativistic Normalization	142
6.5.3	Incoming Flux and Final State Density	144
6.5.4	Lorentz-Invariant Phase Space	145
6.5.5	Cross Section in the Center of Mass Frame	145
6.6	Vertex Functions and the Feynman Propagator	147
6.6.1	$e e \gamma$ Vertex Function	147
6.6.2	Feynman Propagator	151
6.7	Mott Scattering	157
6.7.1	Cross Section	157
6.7.2	Coulomb Scattering and Magnetic Scattering	161
6.7.3	Helicity Conservation	161
6.7.4	A Method to Rotate Spin	161
6.8	Yukawa Interaction	162
7	Qed: Quantum Electrodynamics	167
7.1	$e-\mu$ Scattering	167
7.1.1	Cross Section	167
7.1.2	Elastic Scattering of Polarized $e-\mu$	171
7.1.3	$e^- e^+ \rightarrow \mu^- \mu^+$ Reaction	174
7.2	Compton Scattering	176
7.3	Bremsstrahlung	181
7.3.1	Soft Bremsstrahlung	183
7.4	Feynman Rules	186
8	Radiative Corrections and Tests of Qed*	191
8.1	Radiative Corrections and Renormalization*	191
8.1.1	Vertex Correction	191
8.1.2	Ultraviolet Divergence	193
8.1.3	Infrared Divergence	197
8.1.4	Infrared Compensation to All Orders*	199
8.1.5	Running Coupling Constant	204
8.1.6	Mass Renormalization	208
8.1.7	Ward–Takahashi Identity	210
8.1.8	Renormalization of the Scattering Amplitude	211
8.2	Tests of QED	213
8.2.1	Lamb Shift	213
8.2.2	$g-2$	214
8.2.3	Limit of QED Applicability	216
8.2.4	E821 BNL Experiment	216
9	Symmetries	221
9.1	Continuous Symmetries	222
9.1.1	Space and Time Translation	223
9.1.2	Rotational Invariance in the Two-Body System	227
9.2	Discrete Symmetries	233

9.2.1	Parity Transformation	233
9.2.2	Time Reversal	240
9.3	Internal Symmetries	251
9.3.1	$U(1)$ Gauge Symmetry	251
9.3.2	Charge Conjugation	252
9.3.3	CPT Theorem	258
9.3.4	$SU(2)$ (Isospin) Symmetry	260
10	Path Integral: Basics	267
10.1	Introduction	267
10.1.1	Bra and Ket	267
10.1.2	Translational Operator	268
10.2	Quantum Mechanical Equations	271
10.2.1	Schrödinger Equation	271
10.2.2	Propagators	272
10.3	Feynman's Path Integral	274
10.3.1	Sum over History	274
10.3.2	Equivalence with the Schrödinger Equation	278
10.3.3	Functional Calculus	279
10.4	Propagators: Simple Examples	282
10.4.1	Free-Particle Propagator	282
10.4.2	Harmonic Oscillator	285
10.5	Scattering Matrix	294
10.5.1	Perturbation Expansion	295
10.5.2	S-Matrix in the Path Integral	297
10.6	Generating Functional	300
10.6.1	Correlation Functions	300
10.6.2	Note on Imaginary Time	302
10.6.3	Correlation Functions as Functional Derivatives	304
10.7	Connection with Statistical Mechanics	306
11	Path Integral Approach to Field Theory	311
11.1	From Particles to Fields	311
11.2	Real Scalar Field	312
11.2.1	Generating Functional	312
11.2.2	Calculation of $\det A$	315
11.2.3	n -Point Functions and the Feynman Propagator	318
11.2.4	Wick's Theorem	319
11.2.5	Generating Functional of Interacting Fields	320
11.3	Electromagnetic Field	321
11.3.1	Gauge Fixing and the Photon Propagator	321
11.3.2	Generating Functional of the Electromagnetic Field	323
11.4	Dirac Field	324
11.4.1	Grassmann Variables	324
11.4.2	Dirac Propagator	331
11.4.3	Generating Functional of the Dirac Field	332

11.5	Reduction Formula	333
11.5.1	Scalar Fields	333
11.5.2	Electromagnetic Field	337
11.5.3	Dirac Field	337
11.6	QED	340
11.6.1	Formalism	340
11.6.2	Perturbative Expansion	342
11.6.3	First-Order Interaction	343
11.6.4	Mott Scattering	345
11.6.5	Second-Order Interaction	346
11.6.6	Scattering Matrix	351
11.6.7	Connected Diagrams	353
11.7	Faddeev–Popov’s Ansatz*	354
11.7.1	A Simple Example*	355
11.7.2	Gauge Fixing Revisited*	356
11.7.3	Faddeev–Popov Ghost*	359
12	Accelerator and Detector Technology	363
12.1	Accelerators	363
12.2	Basic Parameters of Accelerators	364
12.2.1	Particle Species	364
12.2.2	Energy	366
12.2.3	Luminosity	367
12.3	Various Types of Accelerators	369
12.3.1	Low-Energy Accelerators	369
12.3.2	Synchrotron	373
12.3.3	Linear Collider	377
12.4	Particle Interactions with Matter	378
12.4.1	Some Basic Concepts	378
12.4.2	Ionization Loss	381
12.4.3	Multiple Scattering	389
12.4.4	Cherenkov and Transition Radiation	390
12.4.5	Interactions of Electrons and Photons with Matter	394
12.4.6	Hadronic Shower	401
12.5	Particle Detectors	403
12.5.1	Overview of Radioisotope Detectors	403
12.5.2	Detectors that Use Light	404
12.5.3	Detectors that Use Electric Signals	410
12.5.4	Functional Usage of Detectors	415
12.6	Collider Detectors	422
12.7	Statistics and Errors	428
12.7.1	Basics of Statistics	428
12.7.2	Maximum Likelihood and Goodness of Fit	433
12.7.3	Least Squares Method	438

Part Two A Way to the Standard Model 441**13 Spectroscopy 443**

- 13.1 Pre-accelerator Age (1897–1947) 444
- 13.2 Pions 449
- 13.3 πN Interaction 454
 - 13.3.1 Isospin Conservation 454
 - 13.3.2 Partial Wave Analysis 462
 - 13.3.3 Resonance Extraction 466
 - 13.3.4 Argand Diagram: Digging Resonances 472
- 13.4 ρ (770) 475
- 13.5 Final State Interaction 478
 - 13.5.1 Dalitz Plot 478
 - 13.5.2 K Meson 481
 - 13.5.3 Angular Momentum Barrier 484
 - 13.5.4 ω Meson 485
- 13.6 Low-Energy Nuclear Force 487
 - 13.6.1 Spin–Isospin Exchange Force 487
 - 13.6.2 Effective Range 490
- 13.7 High-Energy Scattering 491
 - 13.7.1 Black Sphere Model 491
 - 13.7.2 Regge Trajectory* 494

14 The Quark Model 501

- 14.1 $SU(3)$ Symmetry 501
 - 14.1.1 The Discovery of Strange Particles 502
 - 14.1.2 The Sakata Model 505
 - 14.1.3 Meson Nonets 507
 - 14.1.4 The Quark Model 509
 - 14.1.5 Baryon Multiplets 510
 - 14.1.6 General Rules for Composing Multiplets 511
- 14.2 Predictions of $SU(3)$ 513
 - 14.2.1 Gell-Mann–Okubo Mass Formula 513
 - 14.2.2 Prediction of Ω 514
 - 14.2.3 Meson Mixing 516
- 14.3 Color Degrees of Freedom 519
- 14.4 $SU(6)$ Symmetry 522
 - 14.4.1 Spin and Flavor Combined 522
 - 14.4.2 $SU(6) \times O(3)$ 525
- 14.5 Charm Quark 525
 - 14.5.1 J/ψ 525
 - 14.5.2 Mass and Quantum Number of J/ψ 527
 - 14.5.3 Charmonium 527
 - 14.5.4 Width of J/ψ 533
 - 14.5.5 Lifetime of Charmed Particles 536
 - 14.5.6 Charm Spectroscopy: $SU(4)$ 537

14.5.7	The Fifth Quark b (Bottom)	539
14.6	Color Charge	539
14.6.1	Color Independence	542
14.6.2	Color Exchange Force	544
14.6.3	Spin Exchange Force	545
14.6.4	Mass Formulae of Hadrons	547
15	Weak Interaction	553
15.1	Ingredients of the Weak Force	553
15.2	Fermi Theory	555
15.2.1	Beta Decay	555
15.2.2	Parity Violation	562
15.2.3	π Meson Decay	564
15.3	Chirality of the Leptons	567
15.3.1	Helicity and Angular Correlation	567
15.3.2	Electron Helicity	569
15.4	The Neutrino	571
15.4.1	Detection of the Neutrino	571
15.4.2	Mass of the Neutrino	572
15.4.3	Helicity of the Electron Neutrino	576
15.4.4	The Second Neutrino ν_μ	578
15.5	The Universal V–A Interaction	579
15.5.1	Muon Decay	579
15.5.2	CVC Hypothesis	584
15.6	Strange Particle Decays	589
15.6.1	$\Delta S = \Delta Q$ Rule	589
15.6.2	$\Delta I = 1/2$ Rule	591
15.6.3	$K_{l3} : K^+ \rightarrow \pi^0 + l^+ + \nu$	592
15.6.4	Cabibbo Rotation	596
15.7	Flavor Conservation	598
15.7.1	GIM Mechanism	598
15.7.2	Kobayashi–Maskawa Matrix	600
15.7.3	Tau Lepton	601
15.7.4	The Generation Puzzle	605
15.8	A Step Toward a Unified Theory	608
15.8.1	Organizing the Weak Phenomena	608
15.8.2	Limitations of the Fermi Theory	610
15.8.3	Introduction of $SU(2)$	614
16	Neutral Kaons and CP Violation*	617
16.1	Introduction	618
16.1.1	Strangeness Eigenstates and CP Eigenstates	618
16.1.2	Schrödinger Equation for $K^0 - \bar{K}^0$ States	619
16.1.3	Strangeness Oscillation	622
16.1.4	Regeneration of K_1	626

16.1.5	Discovery of CP Violation	630
16.2	Formalism of CP and CPT Violation	632
16.2.1	CP, T, CPT Transformation Properties	632
16.2.2	Definition of CP Parameters	635
16.3	CP Violation Parameters	640
16.3.1	Observed Parameters	640
16.3.2	ϵ and ϵ'	644
16.4	Test of T and CPT Invariance	653
16.4.1	Definition of T- and CPT-Violating Amplitudes	654
16.4.2	T Violation	654
16.4.3	CPT violation	656
16.4.4	Possible Violation of Quantum Mechanics	662
16.5	Experiments on CP Parameters	664
16.5.1	CLEAR	664
16.5.2	NA48/KTeV	666
16.6	Models of CP Violation	673
17	Hadron Structure	679
17.1	Historical Overview	679
17.2	Form Factor	680
17.3	e - p Elastic Scattering	683
17.4	Electron Proton Deep Inelastic Scattering	687
17.4.1	Cross-Section Formula for Inelastic Scattering	687
17.4.2	Bjorken Scaling	690
17.5	Parton Model	693
17.5.1	Impulse Approximation	693
17.5.2	Electron-Parton Scattering	696
17.6	Scattering with Equivalent Photons	699
17.6.1	Transverse and Longitudinal Photons	699
17.6.2	Spin of the Target	702
17.6.3	Photon Flux	703
17.7	How to Do Neutrino Experiments	705
17.7.1	Neutrino Beams	705
17.7.2	Neutrino Detectors	709
17.8	ν - p Deep Inelastic Scattering	712
17.8.1	Cross Sections and Structure Functions	712
17.8.2	ν , $\bar{\nu}$ - q Scattering	715
17.8.3	Valence Quarks and Sea Quarks	716
17.8.4	Comparisons with Experimental Data	717
17.8.5	Sum Rules	719
17.9	Parton Model in Hadron-Hadron Collisions	721
17.9.1	Drell-Yan Process	721
17.9.2	Other Hadronic Processes	724
17.10	A Glimpse of QCD's Power	725

18	Gauge Theories	729
18.1	Historical Prelude	729
18.2	Gauge Principle	731
18.2.1	Formal Definition	731
18.2.2	Gravity as a Geometry	733
18.2.3	Parallel Transport and Connection	734
18.2.4	Rotation in Internal Space	737
18.2.5	Curvature of a Space	739
18.2.6	Covariant Derivative	741
18.2.7	Principle of Equivalence	743
18.2.8	General Relativity and Gauge Theory	745
18.3	Aharonov–Bohm Effect	748
18.4	Nonabelian Gauge Theories	754
18.4.1	Isospin Operator	754
18.4.2	Gauge Potential	755
18.4.3	Isospin Force Field and Equation of Motion	757
18.5	QCD	760
18.5.1	Asymptotic Freedom	762
18.5.2	Confinement	767
18.6	Unified Theory of the Electroweak Interaction	770
18.6.1	$SU(2) \times U(1)$ Gauge Theory	770
18.6.2	Spontaneous Symmetry Breaking	774
18.6.3	Higgs Mechanism	778
18.6.4	Glashow–Weinberg–Salam Electroweak Theory	782
18.6.5	Summary of GWS Theory	784
19	Epilogue	787
19.1	Completing the Picture	788
19.2	Beyond the Standard Model	789
19.2.1	Neutrino Oscillation	789
19.2.2	GUTs: Grand Unified Theories	791
19.2.3	Supersymmetry	792
19.2.4	Superstring Model	795
19.2.5	Extra Dimensions	796
19.2.6	Dark Matter	797
19.2.7	Dark Energy	798
Appendix A Spinor Representation		803
A.1	Definition of a Group	803
A.1.1	Lie Group	804
A.2	$SU(2)$	805
A.3	Lorentz Operator for Spin 1/2 Particle	809
A.3.1	$SL(2, C)$ Group	809
A.3.2	Dirac Equation: Another Derivation	811

Appendix B Coulomb Gauge 813

- B.1 Quantization of the Electromagnetic Field in the Coulomb Gauge 814

Appendix C Dirac Matrix and Gamma Matrix Traces 817

- C.1 Dirac Plane Wave Solutions 817
C.2 Dirac γ Matrices 817
C.2.1 Traces of the γ Matrices 818
C.2.2 Levi-Civita Antisymmetric Tensor 819
C.3 Spin Sum of $|\mathcal{M}_{fi}|^2$ 819
C.3.1 A Frequently Used Example 820
C.3.2 Polarization Sum of the Vector Particle 822
C.4 Other Useful Formulae 823

Appendix D Dimensional Regularization 825

- D.1 Photon Self-Energy 825
D.2 Electron Self-Energy 830

Appendix E Rotation Matrix 833

- E.1 Angular Momentum Operators 833
E.2 Addition of the Angular Momentum 835
E.3 Rotational Matrix 835

Appendix F C, P, T Transformation 839

Appendix G $SU(3)$, $SU(n)$ and the Quark Model 841

- G.1 Generators of the Group 841
G.1.1 Adjoint Representation 842
G.1.2 Direct Product 843
G.2 $SU(3)$ 844
G.2.1 Structure Constants 844
G.2.2 Irreducible Representation of a Direct Product 846
G.2.3 Tensor Analysis 851
G.2.4 Young Diagram 854

Appendix H Mass Matrix and Decaying States 859

- H.1 The Decay Formalism 859

Appendix I Answers to the Problems 865

Appendix J Particle Data 915

Appendix K Constants 917

References 919

Index 929

Preface

The purpose of this book is to present introductory explanations to junior graduate students of major advances in particle physics made in the last century. It is also aimed at laypeople who have not received a standard education in physics but are willing to make an extra effort to follow mathematical logic. Therefore it is organized to be suitable for self-study and is as self-contained as possible. This is why it starts with an introduction to field theory while the main purpose is to talk about particle physics. Quantum field theory is an essential mathematical tool for understanding particle physics, which has fused quantum mechanics with special relativity. However, phenomena that need the theory of relativity are almost completely limited to particle physics, cosmology and parts of astrophysics, which are not topics of everyday life. Few students dare to take a course in special or general relativity. This fact, together with the mathematical complexity of field theory, makes particle physics hard to approach. If one tries to make a self-contained explanation of mathematical aspects of the field, hardly any space is left for the details of particle-physics phenomena. While good textbooks on field theories are abundant, very few talk about the rich phenomena in the world of elementary particle physics. One reason may be that the experimental progress is so rapid that any textbook that tries to take an in-depth approach to the experimental aspect of the field becomes obsolete by the time it is published. In addition, when one makes a great effort to find a good book, basic formulae are usually given without explanation. The author had long felt the need for a good textbook suitable for a graduate course, well balanced between theory and experiment.

Without doubt many people have been fascinated by the world that quantum mechanics opens up. After a student has learned the secrets of its mystery, he (or she) must realize that his world view has changed for life. But at what depth would it have been if he thought he had grasped the idea without knowledge of the Schrödinger equation? This is why the author feels uneasy with presentation of the basic formulae as given from heaven, as is often the case in many textbooks on experimental particle physics.

This book is the first of two volumes. Volume I paves the way to the Standard Model and Volume II develops applications and discusses recent progress.

Volume I, i.e. this book, is further divided into two parts, a field theoretical approach and the way to the Standard Model. The aim of Part I is to equip students

with tools so they can calculate basic formulae to analyze the experimental data using quantized field theories at least at the tree level. Tailored to the above purpose, the author's intention is to present a readable introduction and, hopefully, an intermediate step to the study of advanced field theories.

Chapters 1 and 2 are an introduction and outline of what the Standard Model is. Chapter 3 picks out essential ingredients of special relativity relevant for particle physics and prepares for easy understanding of relativistic formulae that follow. Chapters 4–7, starting from the Dirac equation, field quantization, the scattering matrix and leading to QED (quantum electrodynamics), should be treated as one, closed subject to provide basic knowledge of relativistic field theories and step-by-step acquisition of the necessary tools for calculating various dynamic processes. For the treatment of higher order radiative corrections, only a brief explanation is given in Chapter 8, to discuss how mathematical consistency of the whole theory is achieved by renormalization.

Chapter 9 deals with space-time and internal symmetry, somewhat independent topics that play a double role as an introduction to the subject and appendices to particle physics, which is discussed in Part II.

Two chapters, 10 and 11, on the path integral are included as an addition. The method is very powerful if one wants to go into the formal aspects of field theory, including non-Abelian gauge theories, in any detail. Besides it is the modern way of interpreting quantum mechanics that has many applications in various fields, and the author felt acquaintance with the path integral is indispensable. However, from a practical point of view of calculating the tree-level formulae, the traditional canonical quantization method is all one needs. With all the new concepts and mathematical preparation, the chapter stops at a point where it has rederived what had already been calculated before. So readers have the choice of skipping these chapters and coming back later when their interest has been aroused.

By jumping to Chapter 18 which describes axioms of the Standard Model, the field theoretical approach can be concluded in a self-contained way, at least formally. But it was placed at the end of Part II because the phenomenological approach also culminates in the formulation of the unified gauge theories. Although two alternative approaches are prepared to reach the goal, the purpose of the book is to present particle physics, and field theory is to be considered as a tool and treated as such.

As part of the minimum necessary knowledge, a chapter was added at the end of Part I on basic techniques of experimental measurements, the other necessary tool to understand particle physics phenomena. Readers who are perplexed with the combination may skip this chapter, but it is the author's wish that even students who aim at a theoretical career should understand at least this level of experimental techniques. After all, the essence of natural science is to explain facts and theorists should understand what the data really mean. In addition, a few representative experiments are picked out and scattered throughout the book with appropriate descriptions. They were chosen for their importance in physics but also in order to demonstrate how modern experiments are carried out.

In Part II, various particle phenomena are presented, the $SU(3)$ group is introduced, basic formulae are calculated and explanations are given. The reader should be able to understand the quark structure of matter and basic rules on which the modern unified theory is constructed. Those who do not care how the formulae are derived may skip most of Part I and start from Part II, treating Part I as an appendix, although that is not what the author intended. Part II ends when the primary equations of the Standard Model of particle physics have been derived.

With this goal in mind, after presenting an overview of hadron spectroscopy in Chapter 13 and the quark model in Chapter 14, we concentrate mainly on the electro-weak force phenomena in Chapter 15 to probe their dynamical structure and explain processes that lead to a unified view of electromagnetic and weak interactions. Chapter 17 is on the hadron structure, a dynamical aspect of the quark model that has led to QCD. Chapter 18 presents the principles of gauge theory; its structure and physical interpretation are given. Understanding the axioms of the Standard Model is the goal of the whole book.

Chapter 16 on CP violation in K mesons is an exception to the outline described above. It was included for two reasons. First, because it is an old topic, dating back to 1964, and second, because it is also a new topic, which in author's personal perception may play a decisive role in the future course of particle physics. However, discussions from the viewpoint of unified theories is deferred until we discuss the role of B-physics in Volume II. Important though it is, its study is a side road from the main route of reaching the Standard Model, and this chapter should be considered as a topic independent of the rest; one can study it when one's interest is aroused.

Although the experimental data are all up-to-date, new phenomena that the Standard Model predicted and were discovered after its establishment in the early 1970s are deferred to Volume II. The last chapter concludes what remains to be done on the Standard Model and introduces some intriguing ideas about the future of particle physics.

Volume II, which has yet to come, starts with the axioms of gauge field theories and presents major experimental facts that the Standard Model predicted and whose foundations were subsequently established. These include W, Z, QCD jet phenomena and CP violation in B physics in both electron and hadron colliders. The second half of Volume II deals issues beyond the Standard Model and recent developments of particle physics, the Higgs, the neutrino, grand unified theories, unsolved problems in the Standard Model, such as axions, and the connection with cosmology.

A knowledge of quantum mechanics and the basics of special relativity is required to read this book, but that is all that is required. One thing the author aimed at in writing this series is to make a reference book so that it can serve as a sort of encyclopedia later on. For that reason, each chapter is organized to be as self-contained as possible. The list of references is extensive on purpose. Also, as a result, some chapters include sections that are at higher level than that the reader has learned up to that point. An asterisk * is attached to those sections and problems.

If the reader feels they are hard, there is no point in getting stuck there, skip it and come back later.

This book was based on a series of books originally written in Japanese and published by Asakura Shoten of Tokyo ten years ago. Since then, the author has reorganized and reformulated them with appropriate updating so that they can be published not as a translation but as new books.

Osaka, October, 2009

Yorikiyo Nagashima

Acknowledgements

The author wishes to thank Professor J. Arafune (Tokyo University) for valuable advice and Professors T. Kubota and T. Yamanaka (Osaka University) for reading part of the manuscript and for giving many useful suggestions. Needless to say, any mistakes are entirely the author's responsibility and he appreciates it if the reader notifies them to him whenever possible.

The author would like to express his gratitude to the authors cited in the text and to the following publishers for permission to reproduce various photographs, figures and diagrams:

- American Institute of Physics, publisher of *Physics Today* for permission to reproduce Figure 19.4;
- American Physical Society, publisher of the *Physical Review*, *Physical Review Letters* and the *Review of Modern Physics*, for permission to reproduce Figures 4.3, 5.2ab, 8.8, 9.2, 9.3ab, 12.18abc, 13.11a, 13.17, 13.19, 13.21, 13.23, 13.24ab, 14.6, 14.13–15, 14.20b, 14.25, 14.30, 15.3ab, 15.13, 15.15, 15.17a, 15.21, 16.5–7ab, 16.19, 16.21–23, 17.2b–17.7, 17.12c, 17.15a, 17.21ab, 18.10ab, 18.11abc;
- Annual Reviews, publisher of *Annual Review of Nuclear and Particle Science* for permission to reproduce Figures 12.23ab, 13.29, 14.18, 17.11;
- Elsevier Science Ltd., publisher of *Nuclear Physics*, *Physics Letters*, *Physics Report*, for permission to reproduce Figures 8.9ab, 12.24, 12.25, 12.29, 12.33ab, 12.35ab, 12.37, 12.40–42, 13.14abcd, 13.18, 13.30, 15.8, 15.17b, 15.22, 15.23, 16.1b, 16.8ab–16.10ab, 16.12, 16.13, 16.15–17ab, 17.19, 18.18;
- Institute of Physics Publishing Ltd., publisher of *Report on Progress of Physics* for permission to reproduce Figures 15.11 and 15.12;
- Japan Physical Society, publisher of *Butsuri* for permission to reproduce Figure 12.26b;
- Nature Publishing Group, publisher of *Nature* for permission to reproduce Figures 13.5 and 14.1;
- Particle Data Group, publisher of *Review of Particle Physics* for permission to reproduce Figures 12.10, 12.15–12.17, 12.20, 12.28b, 12.30ab, 12.38, 12.47, 13.25, 13.26, 14.11, 14.16, 14.24, 14.26, 17.17, 18.16;
- Royal Society, publisher of the *Proceedings of the Royal Society* for permission to reproduce Figure 13.4;

- Shokabo, publisher of the Cosmic Rays for permission to reproduce Figure 12.22;
- Springer Verlag, publisher of Zeitschrift für Physik and European Journal of Physics for permission to reproduce Figures 12.43, 16.18, 17.14c, 17.15b, 17.18ab, 17.24.

Part One A Field Theoretical Approach

1

Introduction

1.1

An Overview of the Standard Model

1.1.1

What is an Elementary Particle?

Particle physics, also known as high-energy physics, is the field of natural science that pursues the ultimate structure of matter. This is possible in two ways. One is to look for elementary particles, the ultimate constituents of matter at their smallest scale, and the other is to clarify what interactions are acting among them to construct matter as we see them. The exploitable size of microscopic objects becomes smaller as technology develops. What was regarded as an elementary particle at one time is recognized as a structured object and relinquishes the title of “elementary particle” to more fundamental particles in the next era. This process has been repeated many times throughout the history of science.

In the 19th century, when modern atomic theory was established, the exploitable size of the microscopic object was $\sim 10^{-10}$ m and the atom was “the elementary particle”. Then it was recognized as a structured object when J.J. Thomson extracted electrons in 1897 from matter in the form of cathode rays. Its real structure (the Rutherford model) was clarified by investigating the scattering pattern of α -particles striking a golden foil. In 1932, Chadwick discovered that the nucleus, the core of the atom, consisted of protons and neutrons. In the same year, Lawrence constructed the first cyclotron. In 1934 Fermi proposed a theory of weak interactions. In 1935 Yukawa proposed the meson theory to explain the nuclear force acting among them.

It is probably fair to say that the modern history of elementary particles began around this time. The protons and neutrons together with their companion pions, which are collectively called hadrons, were considered as elementary particles until ca. 1960. We now know that they are composed of more fundamental particles, the quarks. Electrons remain elementary to this day. Muons and τ -leptons, which were found later, are nothing but heavy electrons, as far as the present technology can tell, and they are collectively dubbed leptons. Quarks and leptons are the fundamental building blocks of matter. The microscopic size that can be explored

by modern technology is nearing 10^{-19} m. The quarks and leptons are elementary at this level. They may or may not be at a smaller level. The popular string theory regards the most fundamental matter constituent not as a particle but as a string at the Planck scale ($\sim 10^{-35}$ m), but in this book we limit ourselves to treating only experimentally established facts and foreseeable extensions of their models.

1.1.2

The Four Fundamental Forces and Their Unification

Long-Range Forces: Gravity and the Electromagnetic Force There are four kinds of fundamental forces in nature, the strong, electromagnetic, weak and gravitational forces. The strong force, as the name suggests, is the strongest of all; it is a thousand times stronger, and the weak force a thousand times weaker, than the electromagnetic force. The gravitational force or gravity is negligibly weak (that between the proton and electron is 10^{-42} times the electromagnetic force) at elementary particle levels. Its strength becomes comparable to the others only on an extremely small scale (Planck scale $\sim 10^{-35}$ m), or equivalently at extremely high energy (Planck energy $\sim 10^{19}$ GeV), not realizable on the terrestrial world. Despite its intrinsic weakness, gravity controls the macroscopic world. Indeed, galaxies hundred millions of light years apart are known to be mutually attracted by the gravitational force. The strength of the force decreases (only) inversely proportional to the distance squared, but it also scales as the mass, and the concentration of the mass more than compensates the distance even on the cosmic scale. We refer to it as a long-range force. The behavior of the electromagnetic force is similar to and far stronger than gravity, but, in general, the positive and negative charges compensate each other. However, its long-distance effect manifests itself in the form of the galactic magnetic field, which spans hundreds or possibly millions of light years. It binds nuclei and electrons to form atoms, atoms to form molecules and molecules to form matter as we observe them. Namely, it is the decisive force in the microworld where the properties of matter are concerned. Atomic, molecular and condensed matter physics need only consider the electromagnetic force. A mathematical framework, called quantum electrodynamics (QED), is used to describe the dynamical behavior of point-like charged particles, notably electrons, the electromagnetic field and their interactions. Mathematically, it is a combination of quantized Maxwell equations and relativistic quantum mechanics.

Short-Range Forces: The Strong and Weak Forces In the ultramicroscopic world of the nuclei, hadrons and quarks (at scales less than $\sim 10^{-15}$ m), both the strong and the weak force come in. The reason why they are important only at such a small scale is that they are short ranged reaching only a few hadron diameters, and accumulation of the mass does not help to make it stronger. They are referred to as short-range forces. The strong force (referred to as the color force at the most fundamental level) acts between quarks to bind them to form hadrons. Historically, the strong force was first discovered as the nuclear force to bind protons and neutrons. But as they were found to be composites of the quarks, the nuclear force is recog-

nized as a kind of molecular force (van der Waals force) that can be derived from the more fundamental color force. In 1935, Yukawa predicted the existence of the π meson as the carrier of the nuclear force. The idea that the force is transmitted by a “force carrier particle” was revolutionary and laid the foundation for present-day gauge theories. Later, it was clarified that the pion was a composite and cannot be a fundamental force carrier, but the basic idea remains valid. The weak force is known to act in the decay of hadrons, notably in nuclear β decays. It is also known to control the burning rate of the sun and to play a decisive role in the explosion of type II supernovae.

The electromagnetic force, though playing an essential role at the microscopic level, can also act at the macroscopic distance. We observe electromagnetic phenomena in daily life, for example electrically driven motors or the reaction of a tiny magnet to the earth’s magnetic field. The reason is that its strength is proportional to the inverse square of the distance (Coulomb’s law). In general, a long-range force decreases only as the inverse power of the distance ($\sim r^{-n}$) and can act on an object no matter how far away it is if there is enough of the force source. A short-range force, on the other hand, is a force whose strength diminishes exponentially with the distance. The nuclear force, for instance, behaves like $\sim e^{-r/r_0}/r^2$ ($r_0 \sim 10^{-15}$ m) and reaches only as far as $r \lesssim r_0$. The range of the weak force is even shorter, of the order of $\sim 10^{-18}$ m. Because of this, despite their far greater strength than gravity, their effects are not observed in the macroscopic world. However, if one can confine enough matter within the reach of the strong force and let it react, it is possible to produce a large amount of energy. Nuclear power and nuclear fusion are examples of such applications.

When a Long-Range Force Becomes a Short-Range Force The above description emphasizes the difference between the properties of the forces. Modern unified theories tell us, however, that all the fundamental forces, despite their variety of appearances, are essentially long-range forces and can be treated within the mathematical framework of gauge theories. The seemingly different behavior is a result of metamorphosis caused by environmental conditions. We introduce here one example of a long-range force becoming short range under certain circumstances. A train with a linear induction motor, the next generation supertrain, uses a phenomenon called the Meissner effect to magnetically levitate its body in the air. There, the magnetic field is completely repelled from the superconductor. This happens because, at very low temperatures near absolute zero, electron pairs in a special configuration called the “Cooper pair” collectively react to the magnetic field to form a strong eddy current to cancel the invading magnetic flux. The cancellation is complete except for a shallow depth at the surface, into which the flux can invade. In other words, the electromagnetic force, which is originally long range, becomes short range in the superconductor. This example inspired Nambu [287] to formulate an important concept known as “spontaneous symmetry breaking”, which played a crucial role in forming the Standard Model.

Higgs Mechanism The essence of modern unified theory is to recognize that “the vacuum in which we live is in a sort of superconducting phase”. What corresponds to the Cooper pair in the superconductor is called the Higgs particle, unseen by us but thought to permeate everywhere in the universe. The weak force is long range originally, but because of the ultracold vacuum that surrounds us becomes short range, owing to the pseudo-Meissner effect caused by the Higgs particles. The short-range force translates to a finite mass of the force carrier particle in quantum field theory, as we shall see in the following sections. The Higgs particle condenses at the ultracold temperature and causes a phase transition of the vacuum, a phenomenon similar to the vapor-to-water phase transition. Elementary particles, which have zero mass originally, acquire mass because of the dragging effect caused by swimming through the condensed sea of the Higgs particle. This mass-generating scheme is called the Higgs mechanism. It is thought that at high temperature (such as at the hot Big Bang) and before the Higgs condensation the weak force was a long-range force, just like the electromagnetic force, and both were just different aspects of the same force. The relation between them is similar to that between magnetic and electric forces. They are unified in the sense they are described by a single mathematical framework, just like the electric and magnetic forces are described by the Maxwell equations. After Glashow, Weinberg and Salam, who constructed the unified theory of the electromagnetic and the weak forces, it is often called as GWS theory or simply the electroweak theory.

Confinement and Screening of Color Contrary to the weak force, the strong force becomes short range for a different reason. A group of quarks bind together by the color force to become hadrons. Protons and pions are examples. At short distances, the color force between the quarks is inversely proportional to the square of the distance, just like any long-distance forces but at long distances, its behavior is different and has a component that does not diminish as the distance increases. Therefore, stored potential energy between them increases with distance, hence an infinite energy is required to separate them to a macroscopic distance. The quarks are, in a sense, connected by a string and, like the north and south poles of a magnet, cannot be isolated individually, a phenomenon called “confinement”. They tend not to separate from each other by more than $\sim 10^{-15}$ m, which is the typical size of hadrons. When the string is stretched beyond its limit, it is cut into many pieces, producing more hadrons. The color charges of the quarks come in three kinds dubbed “R, G, B”, for red, green and blue, but only a neutral color combination can form hadrons.¹⁾ In the vicinity of the hadron, quarks of different color are located at different distances and hence the effect of individual color can be felt. But from far away the color looks completely neutralised and the strong force between the hadrons vanishes, making it a short-range force. The force between the hadrons, including the nuclear force, is an action of partially compensated residual

1) The name “color” is used as a metaphor for three kinds of force-generating charge. We may equally call them by a number or any other index. Only a certain mixture of R, G, B, or color with anticolor (complementary color) that makes white can form hadrons. Hence in the color terminology we can conveniently say hadrons are color-neutral.

colors and is therefore secondary in its nature, like the van der Waals force acting between molecules. The action of color forces is, mathematically, almost identical to that of the electromagnetic force and is also described by a gauge theory, the difference being only in the number of charges (three).²⁾ This is the reason for the name quantum chromodynamics (QCD) to describe the mathematical framework of the strong force dynamics.

1.1.3

The Standard Model

The weak force can be described by a gauge theory that contains two kinds of charge. General relativity, the theory of gravity, was also recognized as a kind of gauge theory. Considering that the electromagnetic and weak forces are unified and all four fundamental forces work in the same mathematical framework, it is natural to consider that all the forces are unified but show different aspects in different environments. The Grand Unified Theory (GUT), to unify the electroweak and strong interactions, and the Super Grand Unified Theory,³⁾ to combine all the forces, including gravity, are currently active research areas. Among them “super-gravity” and “superstring theories” are the most popular. But only the electroweak theory and QCD are experimentally well established and are called collectively “the Standard Model” (SM) of elementary particles. The history and current situation of force unification are shown in Fig. 1.1.

The essence of the Standard Model can be summarized as follows:

1. The elementary particles of matter are quarks and leptons.
2. The mathematical framework for the force dynamics are gauge theories.
3. The vacuum is in a sort of superconducting phase.

Particle species in the Standard Model are given in Table 2.1 of Sect. 2.1. Matter particles, the quarks and leptons, are fermions having spin $1/2$, and the fundamental force carriers, also called gauge particles, are spin 1 bosons. The Higgs boson is a spin 0 particle and is the only undiscovered member. Whether it is elementary or a dynamical object remains to be seen. The Standard Model was proposed around 1970 and was firmly established experimentally by the end of the 1970s. The Standard Model was constructed on the torrent of data from the 1960s to 1980s, when large accelerators become operational (see Table 1.1). It marked the end of the era of the 80-year quest for the fundamental building blocks of matter that started with Rutherford at the beginning of the 20th century, and opened a new epoch. The effect on our consciousness (material view) was revolutionary and comparable to that of relativity and quantum mechanics at the beginning of the 20th century. The

- 2) Gluons, the equivalent of photons in electromagnetism, come in eight colors, which are combinations of the three basic colors, excluding white.
- 3) The prefix “super” is attached for theories that contain the supersymmetry that combines fermions and bosons to form one family. They are examples of theories beyond the Standard Model and are briefly described in Chap. 19.

Unification of the Forces

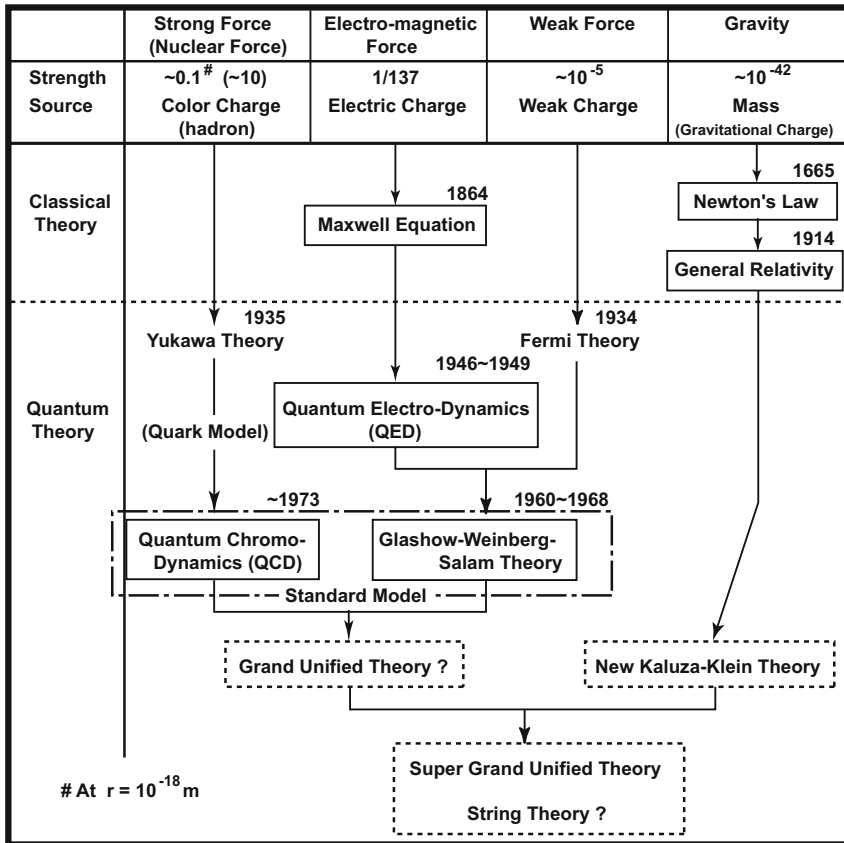


Figure 1.1 Unification of the forces. Those enclosed in dashed lines are not yet established. Only the strong force changes its strength appreciably within currently available energy ranges or, equivalently, distances. This is why the distance of the strong force is specified.

Standard Model is very powerful and can explain all the phenomena in the microworld of particles in a simple and unified way, at least in principle. Forty years have passed since the Standard Model was established, yet there has appeared only one phenomenon that goes beyond the Standard Model. Neutrino oscillation in which a neutrino of one kind is transformed to another while it propagates, does not happen if the mass of the neutrino vanishes, as is assumed in the Standard Model. But it is fair to say it only needs a small stretch and is not a contradiction to the Standard Model. The model is so powerful to the extent that it is not easy to think of an experiment with currently available accelerators that could challenge the Standard Model in a serious way. A quantum jump in the accelerator energy and/or intensity is required to find phenomena that go beyond the Standard Model and explore new physics.

Table 1.1 Chronicle of particle physics. Important discoveries and contributing accelerators. Items in parentheses (...) are theoretical works.

Year	Names	Discovered or proposed	Accelerator
1897	J.J. Thomson	Electron	Cathode tube
1911	E. Rutherford	Atomic model	α Isotope
1929	W. Heisenberg	(Quantum field theory)	
	W. Pauli		
1930	P.A.M. Dirac	(Dirac equation)	
1930	W. Pauli	(Neutrino)	
1932	J. Chadwick	Neutron	α Isotope
"	C.D. Anderson	Positron	Cosmic ray
"	E.O. Lawrence	Cyclotron	Cyclotron
1934	E. Fermi	(Weak interaction theory)	
1935	H. Yukawa	(Meson theory)	
1937	C.D. Anderson	Muon	Cosmic ray
	S.H. Neddermyer		
1947	G.F. Powell	Pion	Cosmic ray
"	G.O. Rochester	Strange particles	Cosmic ray
	C.C. Butler		
1946–1949	S. Tomonaga	(QED)	
	J. Schwinger		
	R. Feynman		
1950–1955		Strange particle production	Cosmotron
1953	K. Nishijima	(Nishijima–Gell-Mann law)	
	M. Gell-Mann		
1954	C.N. Yang	(non-Abelian gauge theory)	
	R.L. Mills		
1955	E. Segre	Antiproton	Bevatron
	O. Chamberlain		
1956	C.L. Cowan	Neutrino	Reactor
	F. Reines		
1956	C.N. Yang	(Parity violation)	
	T.D. Lee		
1958	R. Feynman	(V-A theory)	
	and others		
1960	Y. Nambu	(Spontaneous breaking of symmetries)	
1960–1970	L.W. Alvarez et al.	Resonance production by bubble chambers	Bevatron Cosmotron
1962	L. Lederman	2-neutrinos	BNL/AGS
	M. Schwarz		
	J. Steinberger		

Table 1.1 (continued)

Year	Names	Discovered or proposed	Accelerator
1964	M. Gell-Mann G. Zweig	(Quark model)	
"	V. Fitch J. Cronin	CP violation	BNL/AGS
"	P. Higgs	(Higgs mechanism)	
"	N. Samios et al.	Ω	BNL/AGS
1961–1968	S. Glashow S. Weinberg A. Salam	(Electro-weak unification)	
1969	J.I. Friedman H.W. Kendall R.E. Taylor	Parton structure of the nucleon	SLAC/Linac
~ 1971	G. 't Hooft M.J.G. Veltman	(Renormalization of EW theory)	
1972–1980		Neutrino experiments	Fermilab/PS CERN/SPS
1973	M. Kobayashi T. Maskawa	(KM model on CP)	
1973	H.D. Politzer D.J. Gross F. Wilczek	(Asymptotic freedom and QCD)	
"	A. Lagarrigue et al.	Neutral current	CERN/PS
1974	C.C. Ting B. Richter	J/ψ (charm quark)	BNL/AGS SPEAR
1974–1980		Charm spectroscopy	SPEAR, DORIS
1975	M. Pearl	τ lepton	SPEAR
1978	L. Lederman	Υ (bottom quark)	Fermilab/PS
1978–1979		Bottom spectroscopy	CESR/SPS
		Gluon	DESY/PETRA
1983	C. Rubbia van de Meer	W, Z^0	S $\bar{p}$ pS
1990–2000		$N_\nu = 3$	LEP
		Precision test of EW theory	LEP, TEVATRON
1994		Top quark	TEVATRON
1998	Y. Totsuka et al.	Neutrino oscillation	Cosmic rays
2000	SLAC/BaBar KEK/BELLE	Kobayashi–Maskawa model confirmed	SLAC/B KEK/B

The fact that the kinetic energy of particles in particle physics far exceeds the mass energy makes special relativity an indispensable tool to describe their behavior. Understanding the mathematical formalism that is built on the combination of special relativity and quantum mechanics, also known as quantum field theory, requires a fair amount of effort, and the ensuing results very often transcend human common sense. The only other fields that need both disciplines are cosmology and extreme phenomena in astrophysics, such as supernovae and black-hole dynamics.

1.2

The Accelerator as a Microscope

In order to investigate the microscopic structure of matter, we need a tool, a microscope, that magnifies the object so that we can see it, or more generally recognize it using our five senses. What corresponds to the light-collecting device in the microscope is a particle accelerator. Shining light on the object corresponds to making high-energy particles collide with target particles whose structure we want to investigate. The scattered light is collected by objective lenses and focused to an image, which we observe through an eyepiece. In the accelerator experiment, we use a radioisotope detector to measure patterns of the scattered particles. By computer processing they can be transformed to a pattern that we can recognize as a magnified image. The magnification of a microscope is determined by a combination of lenses, but they merely transform the image to help human eyes to recognize the scattered pattern of light. The resolution, the ability of the optical microscope to discriminate, is largely determined by the wavelength of the light. The more powerful electron microscope uses an electron beam and instead of optical lenses, quadrupole magnets to focus the beam. The scattered electrons are illuminated on a fluorescent board to make an image that human eyes can recognize. In other words, the electron microscope is a mini accelerator–detector complex.

The quantum principle dictates that particles have a dual wave/particle property and the wavelength is given by the Planck constant divided by the particle momentum (the de Broglie wavelength $\lambda = h/p$). Therefore, the resolving power is directly proportional to the momentum or the energy of the incoming particle. This is why we need high-energy accelerators to explore the structure of matter. The smaller the scale, the higher the energy. Therefore, to explore elementary particles, the highest energy technologically realizable at the era is required. This is why particle physics is often called high-energy physics.

Figure 1.2, commonly referred to as the Livingston chart, shows the history of major accelerators and their energy. On the right are indicated the approximate ranges of energy that are suitable for the investigation of the particle structure at the specified level. We summarize important discoveries and tools used for them in the history of particle physics in Table 1.1. We can see that the most effective tools for particle physics are accelerators of the highest energy available at the time. The one with the highest energy currently is the LHC (Large Hadron Collider; a proton–

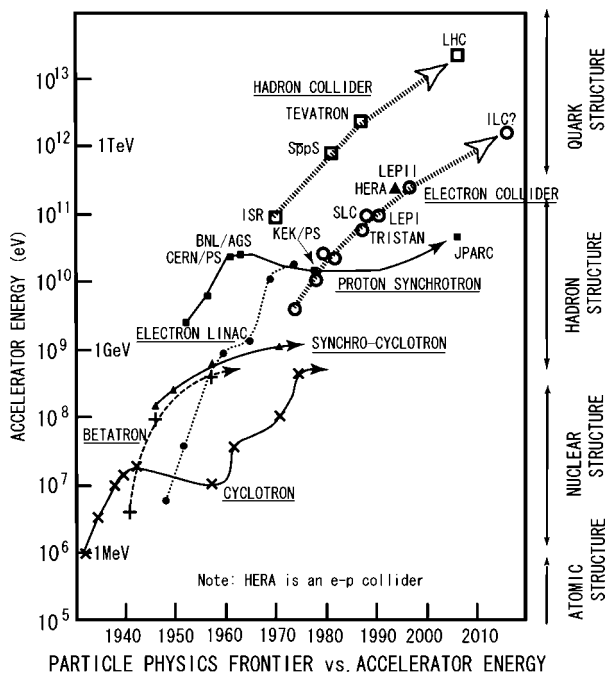


Figure 1.2 The Livingstone chart showing development of accelerators and the investigated constituent particles.

proton colliding accelerator of total energy 14 TeV) in Europe. It started operation in 2009. The LHC can explore a micro-size of $\sim 10^{-20}$ m and is expected to discover the Higgs particle.

2

Particles and Fields

In particle physics, a particle is defined as an energy quantum associated with a “field”. The field is an entity that is defined over all space and time (collectively called space-time) and is able to produce waves, according to classical physics, or quanta, in present day terminology, when excited or when some amount of energy is injected. In a microworld the quanta are observed as particles, but collectively they behave like waves. A quantum is the name given to an object that possesses the dual properties of both a particle and a wave. A typical example of a field is the electromagnetic field. It is created by a charge and extends all over space. It is static when the charge that creates it does not move, but it can be excited by vibrating the source of the field, the electric charge; then the vibrating field propagates. This is a charge radiating an electromagnetic wave. It is well known that historically quantum mechanics has its origin in Planck’s recognition that blackbody radiation is a collection of countable quanta, photons.

In 19th century physics, waves could only be transmitted by some kind of vibrating medium and the electromagnetic waves were considered to propagate in a medium called the “ether”. The existence of the ether was ruled out with the advent of special relativity, and people began to consider that the vacuum, though void of anything in the classical sense, has the built-in property of producing all kinds of fields, whose excitations are observed as quanta or particles. The Standard Model has advanced the idea that the vacuum is not an empty entity but rather filled with various kinds of exotic material, including the Higgs particle, and exhibits dynamical properties like those observed in ordinary matter. The vacuum as we view it nowadays is a kind of resurrected ether with strange attributes that nobody had thought of. In this chapter, we describe an intuitive picture of both particles and fields, define some associated variables and prepare basic tools and terminologies necessary for the treatment of quantum field theory.

2.1

What is a Particle?

There are several properties that characterize a particle.

(1) A particle is spatially localized at each instant and its number is countable.

In classical mechanics, particle's mass m and velocity $\mathbf{v}$ or energy E and momentum $\mathbf{p}$ can be defined and measured. In the nonrelativistic limit, where the particle's velocity is far smaller than the velocity of light c , these variables of a free moving particle are given by

$$\mathbf{p} = m\mathbf{v}, \quad E = \frac{1}{2}mv^2 = \frac{p^2}{2m}, \quad v = |\mathbf{v}|, \quad p = |\mathbf{p}| \quad (2.1)$$

where bold letters denote three-component vectors. However, when the velocity is close to c , Einstein's formulae

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - (v/c)^2}}, \quad E = \frac{mc^2}{\sqrt{1 - (v/c)^2}} \quad (2.2)$$

have to be used instead. In the microworld, the energy and the momentum are more important quantities directly related to observables. We can define the mass and velocity in terms of the energy and the momentum. The relation can be derived from Eq. (2.2):

$$E^2 = (\mathbf{p}c)^2 + (mc^2)^2, \quad \mathbf{v} = \frac{\mathbf{p}c^2}{E} \quad (2.3)$$

Note, Eq. (2.3) also holds for particles with zero mass, such as the photon. When the mass vanishes, $E = pc$ and v is always equal to c .

(2) The particle can be created or annihilated.

This results from Eq. (2.2). The mass of a particle is but one form of energy and can be created if enough energy exists, or the energy can be converted back to mass energy subject to selection rules of one kind or another. Examples of reactions are

$$e^+ + e^- \longrightarrow \gamma + \gamma \quad (2.4a)$$

$$\longrightarrow \mu^+ + \mu^- \quad (2.4b)$$

$$\longrightarrow p + \bar{p} \quad (2.4c)$$

$$p + \bar{p} \longrightarrow \text{any number of } \pi\text{'s} \quad (2.4d)$$

$$\pi^+ \longrightarrow \mu^+ + \nu_\mu \quad (2.4e)$$

$$\pi^- + p \longrightarrow K^0 + \Lambda \quad (2.4f)$$

As a result, the existence probability of a particular kind of particle is not conserved, and the one-particle treatment of the Schrödinger equation is not satisfactory for treating relativistic particles. Quantum field theory is needed to treat the production and annihilation of particles.

(3) A particle is not necessarily stable.

The statement is a corollary to statement (2). The Equation (2.4e) says that the pion is unstable and decays to a muon and a neutrino. Its average lifetime τ is measured to be approximately 2.6×10^{-8} s. This means there is a time uncertainty $\Delta t \sim \tau$ for the existence of the pion. Then the Heisenberg uncertainty principle

$$\Delta E \cdot \Delta t \gtrsim \hbar \quad (2.5)$$

tells us that there is also uncertainty in the energy or the mass in the rest frame. Here $\hbar = h/2\pi$ is Planck's constant divided by 2π and is also called Planck's constant when there is no confusion. In the case of the pion

$$\begin{aligned} \Delta mc^2 \sim \Delta E &\sim \frac{\hbar}{\Delta t} \sim \frac{6.58 \times 10^{-22} \text{ MeV} \cdot \text{s}}{2.6 \times 10^{-8} \text{ s}} \sim 2.5 \times 10^{-14} \text{ MeV} \\ \therefore \frac{\Delta m}{m} &\sim \frac{2.5 \times 10^{-14}}{140} \sim 2 \times 10^{-16} \end{aligned} \quad (2.6)$$

the uncertainty in the mass is, to all practical purposes, negligible. However, in the case of ρ meson, the lifetime is $\sim 4.2 \times 10^{-24}$ s¹⁾ and the mass uncertainty is

$$\Delta m \sim \frac{6.58 \times 10^{-22} \text{ MeV} \cdot \text{s}}{4.2 \times 10^{-24} \text{ s} \times c^2} \sim 160 \text{ MeV}/c^2 \quad (2.7)$$

which cannot be neglected compared to its mass $780 \text{ MeV}/c^2$. The relative uncertainty is very large:

$$\frac{\Delta m}{m} \sim 0.21 \quad (2.8)$$

Let us treat the situation a little more quantitatively. When the average lifetime is given by τ , the particle can die in a brief time interval dt with a probability dt/τ . If dN of N particles die in dt

$$\begin{aligned} \frac{dN}{N} &= -\frac{dt}{\tau} \\ \therefore N(t) &= N_0 e^{-t/\tau} \end{aligned} \quad (2.9)$$

In quantum mechanics, the existence probability of a particle is proportional to the square of its wave function ψ

$$|\psi|^2 \propto e^{-t/\tau} \quad (2.10)$$

and the wave function of the particle in its energy eigenstate has time dependence $\psi(x, t) \sim \psi(x) e^{-iEt/\hbar}$. By putting

$$E = E_0 - i\frac{\Gamma}{2} \quad (2.11)$$

1) In practice, such a short time interval cannot be measured. It is inversely derived from the mass width Eq. (2.16).

the time dependence of the probability is given by

$$|\psi(t)|^2 = |\psi(x)e^{-i(E_0 - i\Gamma/2)t/\hbar}|^2 = |\psi(x)|^2 e^{-\Gamma t/\hbar}$$

$$\therefore \Gamma = \frac{\hbar}{\tau}$$
(2.12)

From the equations, we know that being unstable is equivalent to having a complex energy with negative imaginary part $\Gamma > 0$. The next question is then what kind of mass distribution does the particle have? The wave function as a function of the energy can be expressed as a Fourier transform

$$\tilde{\psi}(E) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt \psi(t) e^{iEt/\hbar}$$
(2.13)

Assuming the particle was created at $t = 0$ and $x = 0$, $\psi(t) = 0$ for $t < 0$, then

$$\begin{aligned} \tilde{\psi}(E) &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} dt \psi(x=0, t) e^{iEt/\hbar} \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} dt \psi(0) e^{-i(E_0 - i\Gamma/2)t/\hbar} e^{iEt/\hbar} \\ &= \frac{\psi(0)}{\sqrt{2\pi}} \frac{i\hbar}{E - E_0 + i(\Gamma/2)} \end{aligned}$$
(2.14)

Therefore, the probability that the particle takes an energy value E is given by

$$P(E) \propto |\tilde{\psi}(E)|^2 = \frac{|\psi(0)|^2}{2\pi} \frac{\hbar^2}{(E - E_0)^2 + (\Gamma/2)^2}$$
(2.15)

The functional shape is shown in Figure 2.1 and is called a Lorentzian, the same function as the Breit–Wigner resonance formula, which describes unstable levels of nuclei. From the shape, Γ represents the full width at half maximum (FWHM) of the function. Γ can be regarded as the energy uncertainty of the unstable particle, hence the Heisenberg uncertainty principle

$$\Delta E \cdot \Delta t \sim \Gamma \cdot \tau = \hbar$$
(2.16)

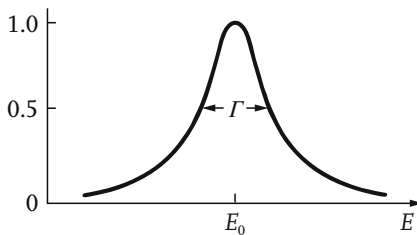


Figure 2.1 Lorentzian shape or Breit–Wigner resonance formula for the energy (or mass) distribution of unstable particles.

is reproduced. The unstable elementary particles such as ρ and $\Delta(1236)$ have the mass and width

$$m(\rho) = 776 \text{ MeV}/c^2, \quad \Gamma = 158 \text{ MeV} \quad (2.17)$$

$$m[\Delta(1236)] = 1236 \text{ MeV}/c^2, \quad \Gamma = 115 \text{ MeV} \quad (2.18)$$

Historically, they were observed as strong enhancements in the cross sections of $\pi\text{-}\pi$ and $\pi\text{-}p$ scatterings, namely as resonances. The process strongly suggests them to be of dynamical origin, and in the early days of hadron dynamics they were treated as such. Today, all hadrons, including nucleons, pions and resonances, are identified as the energy levels of quark bound states and are treated on an equal footing.

(4) A particle has spins and other degrees of freedom.

Energy and momentum are not the only attributes a particle possesses. It acts as if it were spinning on some axis. Spin means angular momentum and its magnitude is quantized to integers or half-integers in units of the Planck constant $\hbar$. Spin components or a projection of the spin onto a certain axis (z -axis) is also quantized in units of $\hbar$:

$$S_z = S, S-1, S-2, \dots, -(S-1), -S \quad (2.19)$$

If a charged particle has a finite spin, it has a magnetic moment.²⁾

Spin in quantum mechanics has, mathematically, all the properties of angular momentum, but the picture of a self-rotating object such as the rotating earth does not apply, because a point particle cannot have angular momentum in the classical sense ($\mathbf{L} = \mathbf{r} \times \mathbf{p}$). Therefore the spin component should rather be regarded as just a degree of freedom. For a particle of spin $1/2$ like an electron, only ($S_z = 1/2, -1/2$) is allowed. For the electron, energy, momentum, electric charge and spin orientation exhaust its degrees of freedom. This is evident as only two $S_z = \pm 1/2$ electrons can occupy the same orbit in a hydrogen atom. The Pauli exclusion principle forbids other electrons of the same attributes to occupy the same quantum state at the same space point. Other particles may have more attributes, such as isospin, color, strangeness, etc., which are generically called internal degrees of freedom.

(5) A particle has a corresponding antiparticle.

The particle and its antiparticle have identical mass, lifetime and spin values, but the spin-components and all the internal quantum numbers are opposite. The positron is the antiparticle of the electron. The photon is its own antiparticle, a situation that holds for other particles, such as neutral pions, which have no discriminating internal quantum numbers. The sum of spin components and internal quantum numbers of a particle and its antiparticle is zero, namely the same quantum number as that of the vacuum. When they meet, they annihilate pair-wise and become pure energy, usually in the guise of photons, gluons and other gauge

2) A composite particle such as the neutron can generally have magnetic moment even if it is electrically neutral. Think of two oppositely charged particles rotating in reverse directions. Even a neutral elementary particle like the neutrino can have it due to higher order radiative corrections.

Table 2.1 Elementary particles and their characteristics

	Spin	Electric charge ^a	I_3 ^b	Matter constituent particles		
				I	II	III
Quarks	1/2	+2/3	+1/2	u (up)	c (charm)	t (top)
	1/2	-1/3	-1/2	d (down)	s (strange)	b (bottom)
Leptons	1/2	0	+1/2	ν_e	ν_μ	ν_τ
	1/2	-1	-1/2	e^-	μ^-	τ^-

^a In units of proton charge.

^b The third component of weak isospin.

bosons. If sufficient energy is given to the pair, the produced energy instantly re-materializes to any number of particle–antiparticle pairs or bosons of heavier mass within the allowed energy limit.

(6) Matter particles and force carriers.

Minimum elements of matter that can be extracted fall into two classes. One exerts the strong force and the other does not. The former are called hadrons, the latter, leptons. Hadrons with half-integer spin are called baryons and include protons, neutrons and other heavier particles such as Λ , Σ , Ξ , $\dots$. Those with integer spin are called mesons and include π , K , ρ , $\dots$. All hadrons are composites of quarks that have spin 1/2, but the quarks cannot be isolated (confinement). One might doubt the real existence of such objects, but there is ample evidence for their existence inside the hadrons. Today, six kinds of quarks (u , d , s , c , b , t) are known with increasing mass, each having three color degrees of freedom and electric charge 2/3 or -1/3 in units of the proton charge e . The leptons have spin 1/2, too. Six kinds of leptons are known, three kinds of charged leptons (e , μ , τ) and their associated neutral leptons (neutrinos) (ν_e , ν_μ , ν_τ). In the Standard Model, the mass value of the neutrino is assumed zero.³⁾ To summarize, the constituent particles of matter are six kinds of quarks and leptons distinguished by “flavor”. They act pairwise in the weak interaction and are further divided into three generations. The particle species are summarized in Table 2.1.

Properties of the particles in the same row but in different generations are almost identical apart from their mass, which becomes heavier with the generation (see Table 2.2). Known matter in the universe is made mostly of the first gener-

3) In 1998, a phenomenon called neutrino oscillation, in which ν_μ is converted to another kind of neutrino, was discovered in atmospheric neutrinos. This implies a finite mass of the neutrino. Within the context of

this book, however, the neutrino mass may safely be assumed zero until later, when neutrino properties are discussed.

Table 2.2 Mass values of quarks and leptons [311]

quark	<i>u</i>	<i>d</i>	<i>s</i>	<i>c</i>	<i>b</i>	<i>t</i>
^a	~ 3 MeV	~ 5 MeV	~ 104 MeV	~ 1.27 GeV	~ 4.2 GeV	171 GeV
lepton	ν_1^b	ν_2^b	ν_3^b	<i>e</i>	μ	τ
^c	?	~ 0.009 eV	~ 0.05 eV	0.511 MeV	105.7 MeV	1777 MeV

^a Current mass which is determined by dynamical processes like scattering.

^b Observed ν_e, ν_μ, ν_τ are mixtures of mass eigenstates ν_1, ν_2, ν_3 .

^c For the neutrinos, only mass differences $\Delta m_{ij}^2 = |m_i^2 - m_j^2|$ are known. Consequently, $m_3 \gg m_2 \gg m_1$, is assumed. Note: MeV = 10^6 eV, GeV = 10^9 eV

ation.⁴⁾ Others are generated in high-energy phenomena or supposedly exist in extreme conditions such as the interior of neutron stars. The existence and role of the generation is an unsolved problem.⁵⁾ There is no clear evidence that higher generations do not exist, but the weak gauge boson Z^0 can decay into pairs of neutrinos and antineutrinos producing no others except the known three kinds. If the fourth generation exists, its neutrino has a mass greater than $m_Z/2 = 45 \text{ GeV}/c^2$. Considering it an unlikely situation, the Standard Model assumes only three generations.

Quantum field theory requires the existence of a particle that mediates the interaction, a force carrier. It is a quantum of the force fields. Gauge theory requires the force field to be a vector field, hence the force carrier must have spin 1.⁶⁾ Corresponding to the four fundamental forces in nature, there are four kinds of force carrier particles, gluons for the strong force, photons for the electromagnetic force, weak bosons ($W^\pm, Z^0$) and gravitons for gravity. A list of fundamental interactions and their characteristics are given in Table 2.3.

Vacuum particle? If the Higgs particle is discovered, there could be another category of spin 0 fundamental particle. It may be called a vacuum particle from the role of the Higgs. However, whether it is fundamental or of dynamical origin has yet to be determined. A typical example of the dynamical Higgs is if it is composed of a $t\bar{t}$ pair [277], just like the Cooper pair in superconductivity is composed of an electron pair of opposite momentum and spin orientation.

4) There is an unidentified matter form in the universe known as “dark matter”, which makes up ~ 22% of all cosmic energy. The ordinary matter of known type makes up only ~ 5%. The rest is unidentified vacuum energy called “dark energy”.

5) The muon, for instance, could be termed a heavy electron. Except for the fact that it can decay into an electron and neutrinos, all the reaction formulae for the muon may be reproduced from those of the electron simply by replacing its mass. The

seemingly mysterious existence of the muon is expressed in a phrase by I. Rabi: “Who ordered it?”. The generation puzzle is the extension of the muon puzzle.

6) This is true for the Standard Model force carriers. The quantization of the gravitational field produces a spin 2 graviton. This is due to the special situation that the source of the gravity is mass, or more generally energy-momentum, which itself is a second rank tensor in space-time.

Table 2.3 List of fundamental forces and their characteristics.

Force Source	Strong Color charge (RGB)	(Nuclear) ^a (hadron)	Electromagnetic Electric charge	Weak Weak charge ($I_3 = \pm 1/2$)	Gravity Mass ^b
Strength	0.1 ^c	(10)	1/137	10 ⁻⁵	10 ⁻⁴²
Range [m]	∞^d	(10 ⁻¹⁵)	∞	10 ⁻¹⁸	∞
Potential shape	$k_1 \frac{1}{r} + k_2 r$	$\frac{\exp(-m_\pi r)}{r} e$	$\frac{1}{r}$	$\frac{\exp(-m_W r)}{r} e$	$\frac{1}{r}$
Force carrier	g (gluon)	π meson	γ (photon)	$W^\pm, Z^0$	graviton
Spin	1	0	1	1	2
Theory	QCD	(Yukawa theory)	QED	$\underbrace{\hspace{1cm}}_{\text{GWS}^f \text{ Theory}}$	General relativity
Gauge symmetry	SU(3)	—	U(1) $\times$ SU(2)		Poincaré group

a Secondary strong force acting on hadrons.

b To be exact, energy-momentum tensor.

c At distance 10⁻¹⁵ m. The distance is specified as the color force changes its strength considerably within the measured range 10⁻¹⁵–10⁻¹⁸ m.

d The range beyond $\sim 10^{-15}$ m is meaningless because of confinement.

e m_π , m_W should be read as $m_\pi c/\hbar$, $m_W c/\hbar$, i.e. as inverse Compton wavelength.

f Glashow–Weinberg–Salam.

2.2

What is a Field?

2.2.1

Force Field

Force at a Distance, Contact Force and the Force Field

In classical mechanics, the gravitational force is expressed by Newton's formula

$$\mathbf{f} = -G_N \frac{mM}{r^2} \quad (2.20)$$

This means that an object with mass M exerts its force directly on another object with mass m at a distance r . Einstein considered such a “force at a distance” is unacceptable. To understand why, let's carry out a thought experiment. Suppose the sun suddenly disappears at $t = 0$ and infer what would happen to the earth. According to Newton's way of thinking, the sun's attraction would disappear instantly, because the mass M , the source of the force, has disappeared and the earth would be thrown off its orbit into cosmic space at the same instant. According to Einstein's principle of relativity, nothing can travel faster than light, and the information of the sun's disappearance would take at least 8 minutes to reach the earth. During the 8 minute's interval there is no way for the earth to know that the sun has vanished, which means that the gravitational force continues to act. It follows that something other than the sun must be there to exert the force on the earth. We call this something a “field”. The sun (mass) creates a gravitational field over all space and the force acting on the earth is the action of the local field at the same space point as where the earth is located. Namely, all forces must be “contact forces” that can act only upon contact with the object on which they exert. In this way of thinking, wherever a force acts there must be an associated field, a gravitational field for gravity and an electromagnetic field for the electromagnetic force.

Take a look at the Maxwell equations of the electric force, which are written as

$$\nabla \cdot \mathbf{E}(t, \mathbf{x}) = \frac{\rho(t, \mathbf{x})}{\epsilon_0}, \quad \mathbf{f}(t_y, \mathbf{y}) = q_1 \mathbf{E}(t_y, \mathbf{y}), \quad \nabla \times \mathbf{E}(t, \mathbf{x}) = 0 \quad (2.21)$$

This translates to the statement that an electric charge distribution ρ at a space point $\mathbf{x}$ at time t creates an electromagnetic field $\mathbf{E}(\mathbf{x}, t)$ over all space. A point charge q_1 at a spacetime point $(t_y, \mathbf{y})$ receives the force from the local electric field $\mathbf{E}$ at the same space-time where it is located. The third equation is an auxiliary condition that the static field must satisfy. If the static electric field is created by a point charge q_2 at point $\mathbf{r} = 0$

$$\rho = q\delta(\mathbf{r}), \quad \delta(\mathbf{r}) = \delta(x)\delta(y)\delta(z) \quad (2.22)$$

Solving Eq. (2.21) with the above input gives Coulomb's law

$$\mathbf{f} = \frac{q_1 q_2}{4\pi\epsilon_0} \frac{1}{r^2} \frac{\mathbf{r}}{r}, \quad r = |\mathbf{r}| \quad (2.23)$$

The result has the same form as Newton's law, but in the dynamical (i.e. time-dependent) solution the influence of a changing force never propagates faster than light. Namely, the Maxwell equations respect the principle of relativity. This is guaranteed by the local nature of the differential equations. Integral forms are useful for an intuitive interpretation, but the equations have to be written in differential forms. Historically, Maxwell's equations were obtained empirically (except for the displacement current), then the Lorentz transformation was derived to make the force action consistent in the rest frame and in a moving frame, which is a basis of the theory of relativity.

Photon

An oscillating electromagnetic field propagates through space and becomes an electromagnetic wave. It can be created by oscillating the source, but once created, it becomes a wave acting independently of the source. Quantum mechanics tells us that the wave behaves also as a particle. An entity exhibiting the dual behavior of a particle and a wave is called a quantum. The quantum of the electromagnetic field is called a photon. The above statement can be summarized as follows

The principle of relativity requires the force field and the quantum principle requires quanta of the field, i.e. force carrier particles.

The electromagnetic wave can carry energy and when it meets another electrically charged object it forces it to oscillate, because classically it is an oscillating force field. But as a particle, when it hits the object, it is either scattered or absorbed, giving it a kick. The energy and momentum of the quantum are given by the relations

$$E = h\nu = \hbar\omega, \quad p = \hbar k = \frac{h}{\lambda} \frac{k}{|k|} \quad (2.24)$$

where ν is the wave's frequency, $\omega = 2\pi\nu$ the angular frequency, λ the wavelength and $\mathbf{k}$ the wave vector. As we have seen, the creation of the electromagnetic (force) field is equivalent to the emission of a photon (force carrier particle) and the exertion of the force by the electromagnetic (force) field on the charge (source) is equivalent to absorption of the photon (quantum) by the charged object (another source). In quantum field theory, action of a force means an exchange of the force carrier particle. In other words, action/reaction of the force is synonymous to absorption/emission of the force quantum by the object on which the force acts. In a mathematical treatment of quantum mechanics, energy is transmitted often in momentum or energy space without ever referring to the space variables. The notion of a force is a bit different from that used in classical mechanics, where it refers to an energy variation associated with space displacement ($f = -\partial E/\partial x$). So instead of saying force exertion, the word "interaction" is used where any exchange of energy occurs.

Strength of the Force

In the following, we talk of (potential) energy instead of force. The electromagnetic interaction between, say, a proton and an electron in a hydrogen atom is expressed by the Coulomb potential and the strength of the force by α where

$$\begin{aligned} V(r) &= \frac{e^2}{4\pi\epsilon_0} \frac{1}{r} = \alpha \frac{\hbar c}{r} \\ \alpha &\equiv \frac{e^2}{4\pi\epsilon_0 \hbar c} \sim \frac{1}{137} \end{aligned} \quad (2.25)$$

Here, to define a dimensionless quantity α , we used the fact that $\hbar c$ has the dimension of [energy $\times$ distance]. e is the unit electric charge that both the electron and the proton have, which has a value of 1.6×10^{-19} Coulomb in rationalized MKS units. α is called the fine structure constant and represents the fundamental strength of the electromagnetic interaction. The dimensionless strength of the gravitational force can be defined similarly:

$$\begin{aligned} V(r) &= G \frac{m_1 m_2}{r} = \alpha_G \left(\frac{m_1 m_2}{m_p m_e} \right) \frac{\hbar c}{r} \\ \alpha_G &= G_N \frac{m_p m_e}{\hbar c} \sim 3 \times 10^{-42} \end{aligned} \quad (2.26)$$

where G_N is Newton's constant of universal gravitation and m_p, m_e are the proton and electron mass. In particle physics gravity can be neglected because $(\dots) \sim O(1)$ hence α_G is by many orders of magnitude smaller than α . The strength of the strong and weak force will be described in the following section.

Problem 2.1

- (1) Calculate and confirm the values of the fine structure constant and the gravitational strength.
- (2) Calculate the mass value which makes $\alpha_G = 1$. Express it in units of GeV. It is called the Planck energy.

Force Range

Generally a particle has finite mass. What will happen if a force carrier particle has nonzero mass? To exert the force, it has to be emitted and absorbed by interacting objects. The finite mass of the force carrier if emitted produces energy uncertainty of at least $\Delta m c^2$. The Heisenberg uncertainty principle dictates that the time interval Δt that allows energy violation should be less than $\sim \hbar / \Delta E$. If the force carrier is not absorbed within this time interval, the force cannot be transmitted. As the particle velocity is less than the light velocity c , the force range r_0 should be less than

$$r_0 \lesssim c \Delta t \sim \hbar / m c^2 = \hbar / m c \quad (2.27)$$

Therefore, the range of a force mediated by a force carrier having a finite mass is limited to its Compton wavelength.

Let us treat the problem a little more quantitatively. As will be explained soon, any relativistic particle should be a quantum of a field that satisfies the Klein–Gordon equation

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 + \mu^2 \right] \varphi = 0, \quad \mu = mc/\hbar \quad (2.28)$$

In analogy with the electric potential, the force field created by a point source would satisfy the equation

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 + \mu^2 \right] \varphi = g\delta^3(\mathbf{r}) \quad (2.29)$$

where g denotes the strength of the force corresponding to the electric charge of the electric force. For a static potential where $\partial/\partial t = 0$, the equation is easily solved to give

$$\varphi(r) = \frac{g}{4\pi} \frac{e^{-r/r_0}}{r}, \quad r_0 = \frac{1}{\mu} = \frac{\hbar}{mc} \quad (2.30)$$

Problem 2.2

Derive Eq. (2.30) from Eq. (2.29).

Hint: Use the Fourier transform

$$\varphi(\mathbf{r}) = \int \frac{d^3 p}{(2\pi)^3} e^{i\mathbf{p} \cdot \mathbf{r}} \tilde{\varphi}(\mathbf{p}), \quad \delta^3(\mathbf{r}) = \int \frac{d^3 p}{(2\pi)^3} e^{i\mathbf{p} \cdot \mathbf{r}}.$$

The reach of a short-range force is mathematically defined as Eq. (2.30), which is called the Yukawa potential. When $r \ll r_0$, the potential is of the Coulomb type, but for $r \gg r_0$, decreases very rapidly. For gravitation or the Coulomb force, $r_0 = \infty$ and they are called long-range forces. Based on the above formula, Yukawa predicted the mass value of the π meson, the nuclear force carrier, to be roughly 200 times that of the electron from the observation that the reach of the nuclear force is approximately $\sim 10^{-15}$ m. The strength of the nuclear force is measured to be $g^2/4\pi \simeq 10$, which is 1000 times the electromagnetic force. The source of the nuclear force is a nucleon, but as the nucleon as well as the pion turned out to be composites of quarks, the real source of the strong force was later identified as the color charge, which all quarks carry. The true strong force carrier is called the gluon and its mass is considered to be zero, just like the photon. The potential between the quarks is of the Coulomb type, but for reasons to be described later, has an extra term and is approximately (phenomenologically) given by

$$\phi(r) = k_1 \frac{1}{r} + k_2 r \quad (2.31)$$

The second term increases linearly with distance and is called the confining potential. k_1 decreases logarithmically at short distance, a phenomenon called asymptotic freedom and takes a value ~ 0.1 at $r \sim 10^{-18}$ m.

The weak force carriers come in three kinds, charged $W^\pm$ with mass of $80.4 \text{ GeV}/c^2$ and a neutral Z^0 with mass of $91.2 \text{ GeV}/c^2$. Their mass value means $r_0 \sim 10^{-18} \text{ m}$ for the weak force. The range of the weak force is so small that practically it could be replaced with a contact force before the advent of big accelerators. The coupling strength was expressed in terms of the dimensionful quantity $G_F \sim 10^{-5} m_p^2$ for historical reasons, where m_p is the mass of the proton.

Problem 2.3

From the mass values of $W^\pm$, Z^0 , calculate the weak force range.

Problem 2.4

From observation of stars and polarization of star light, it is known that a coherent magnetic field of $\sim 3 \mu (\mu = 10^{-6})$ gauss exists in the Crab nebula. This means that the electromagnetic force range is at least hundreds of light years. If the intergalactic magnetic field is confirmed, the range is even larger, millions of light years. Calculate the upper limit of the photon mass.

2.2.2

Relativistic Wave Equation

Physical quantities that characterize a particle are energy E and $\mathbf{p}$, but in quantum mechanics the particle is a quantum and shows wave behavior. The wave picture of a particle can be derived by replacing

$$E \rightarrow i\hbar \frac{\partial}{\partial t}, \quad \mathbf{p} \rightarrow -i\hbar \nabla \quad (2.32)$$

and regarding them as operators acting on a wave function. Inserting Eq. (2.32) into the nonrelativistic energy-momentum relation (2.1), we obtain a wave equation for the de Broglie wave:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi \quad (2.33)$$

If the particle is moving in a potential, one adds a potential $V(\mathbf{x})$ and the result is the well-known Schrödinger equation. To see if we can obtain a relativistic Schrödinger equation, we apply (2.32) to Eq. (2.3) and obtain

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 + \mu^2 \right] \varphi = 0, \quad \mu \equiv \frac{mc}{\hbar} \quad (2.34)$$

which is called the Klein–Gordon equation. Equation (2.34) has a plane wave solution

$$\varphi(r) \equiv \varphi(\mathbf{x}, t) = N e^{i(\mathbf{p} \cdot \mathbf{x} - Et)/\hbar} \quad (2.35)$$

where N is a normalization constant. A general solution can be expressed as a superposition of plane waves. Unfortunately, unlike the nonrelativistic equation (2.33), we cannot regard φ as a probability amplitude for a particle. If φ is to be regarded as a probability amplitude just like ψ in the Schrödinger equation, the probability density ρ and its flow (current) $\mathbf{j}$ have to be defined to satisfy the continuity equation (i.e. conservation of probability)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \quad (2.36)$$

The current $\mathbf{j}$ may assume the same form as the Schrödinger equation:

$$\mathbf{j} = -\frac{i\hbar}{2m} (\varphi^* \nabla \varphi - \varphi \nabla \varphi^*) \quad (2.37)$$

Then, in a relativistic treatment (to be discussed in Chap. 3), ρc has to be the fourth component of the four-dimensional current vector.⁷⁾

$$\rho c = \frac{i\hbar}{2mc} \left(\varphi^* \frac{\partial \varphi}{\partial t} - \varphi \frac{\partial \varphi^*}{\partial t} \right) \quad (2.38)$$

It is easily verified that ρ and $\mathbf{j}$ defined above satisfy the continuity equation Eq. (2.36).

One is tempted to interpret ρ as the probability density, but it does not work as it encounters difficulties.

Negative Energy The first difficulty stems from the fact that the energy-momentum equation $E^2 = (\mathbf{p}c)^2 + (mc^2)^2$ has two solutions

$$E = \pm \sqrt{(\mathbf{p}c)^2 + (mc^2)^2} \quad (2.39)$$

Namely, E can be negative. The particle with negative energy cannot occupy a stable state. Because it can be transferred to an indefinitely lower energy state by emitting a photon, hence is unstable. Historically, people tried to construct a theory without the negative energy solution. Mathematically, however, they are required to satisfy the completeness condition as well as causality and trying to construct a closed system without them has produced various paradoxes.

Negative probability The second difficulty stems from the fact that the Klein-Gordon equation is a second-rank differential equation and requires two initial conditions $\varphi(t = 0)$ and $\partial\varphi/\partial t(t = 0)$ to define its solution uniquely. In other words, φ and $\partial\varphi/\partial t$ can be chosen arbitrarily and do not always guarantee a positive value of ρ . In fact, if we calculate j^μ using the plane wave solution (2.35) we have

$$j^\mu = |N|^2 \frac{p^\mu}{m} = |N|^2 \left(\frac{E}{mc}, \frac{\mathbf{p}}{m} \right) \quad (2.40)$$

from which we see that a negative energy solution also gives negative probability. For these reasons, the Klein-Gordon equation was once abandoned historically.

7) c is multiplied to make the physical dimension the same as the space components.

2.2.3

Matter Field

However, we can still produce wave behavior by interpreting φ as a classical field. The four-dimensional current

$$qj^\mu \equiv (\rho c, \mathbf{j}) = i\hbar q(\varphi^* \partial^\mu \varphi - \varphi \partial^\mu \varphi^*), \quad \partial^\mu \equiv \left(\frac{1}{c} \frac{\partial}{\partial t}, -\nabla \right) \quad (2.41)$$

can now be interpreted as a current (charge flow) having positive as well as negative charges. From the classical field point of view, negative energy means negative frequency and simply means a wave propagating in the opposite direction. We shall see later that particles, or quanta to be exact, appear by quantizing the amplitude of the waves. From the wave amplitude of negative frequency appear particles with quantum number opposite to that of positive frequency, as suggested by the current flow. Both quanta have positive energy and particles with negative energy never appear. As it happened, history took a detour in coping with the notion of negative energy.

The above statement is equivalent to considering matter particles on an equal footing with the photon. Just as the photon is the quantum of the classical electromagnetic field, we regard a matter particle as a quantum of some kind of matter field. To see the parallelism between the matter field and the electromagnetic field, observe that the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ are expressed in terms of a scalar and a vector potential

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A} \quad (2.42a)$$

$$A^\mu \equiv \left(\frac{\phi}{c}, \mathbf{A} \right) \quad (2.42b)$$

In the Lorentz gauge, the equations that the potentials satisfy are given as

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = q \frac{\rho}{\epsilon_0} \quad (2.43a)$$

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} = \mu_0 q \mathbf{j} \quad (2.43b)$$

where q denotes electric charge. In vacuum, where $\rho = \mathbf{j} = 0$, Eq. (2.43) are equations for a freely propagating electromagnetic field, namely, the photon. Equation (2.43) are nothing but the Klein–Gordon equation with $m = 0$. Note that φ which we have introduced is a scalar field while the photon is a vector field. The difference results in the spin of the field quantum, but does not affect the discussion here.

The particle behavior lost by treating φ as a classical field can be recovered by quantizing the field (a process sometimes called as the second quantization), just as the photon is the quantized electromagnetic field. Quantization of the field is an

inevitable consequence of fusing quantum mechanics with the theory of relativity.⁸⁾ Since the field is a collection of an infinite number of harmonic oscillators (see the following arguments), the quantization of the field can be achieved by quantizing its harmonic components.

2.2.4

Intuitive Picture of a Field and Its Quantum

Before we jump into the mathematics of field quantization, it may help to have a physical image of what the field is made of. We use an analogy to visualize the field to help the reader to understand the otherwise formal and mathematical procedure of quantization. Of course, there is always a danger of stretching the analogy too far and applying it where it should not be. True understanding is obtained only by interpreting the mathematics correctly. With those caveats, I proceed and pick up an example from a familiar phenomenon in condensed matter physics.

The example we describe is metal, a form of crystal. In the metal, positively ionized atoms are distributed regularly lattice wise. A mechanical model of point masses connected by springs is a good approximation (Fig. 2.2). Atoms in the lattice can vibrate and transmit their vibration to neighboring lattices. Namely, the lattice of positively ionized atoms has the property of the field we want to talk about. The quantum of the lattice vibration is called a phonon. In the hypothetical conditions where the temperature is absolute zero and all the lattice vibrations come to rest, the electric field made by the lattice is uniform, at least at a scale larger than the lattice spacing, and therefore no net forces exist to act on the electrons. This is because the direct Coulomb forces between the electrons are cancelled by negative Coulomb forces acting between the electron and the lattice.

Put an electron at some space point in the metal, and its electric field would deform the regular lattice configuration (or field) in its neighborhood. Another electron placed at some distance would feel the field distortion. Namely, the two electrons can exert forces on each other through coupling with the field. If the elec-

8) The field could be relativistic or nonrelativistic. In this book, we talk about only relativistic fields.

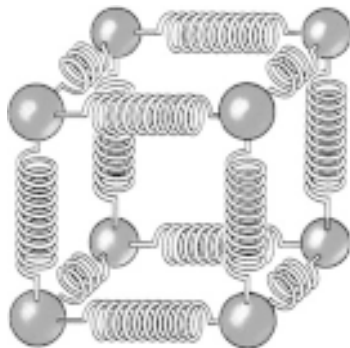


Figure 2.2 Mechanical oscillator model of a field.

tron is in motion, the changing field it produces can excite the lattice vibration or, in quantum mechanics, create a phonon. The vibration can propagate through the field and disturbs the other electron, or equivalently the electron is scattered by absorbing the phonon.

This classical treatment of the electron coupling with the lattice vibration can be rephrased as an interaction of the electron and the phonon. The quantum treatment of the phonon–electron interaction and that of the photon–electron interaction in particle physics are very much alike, except that a relativistic treatment is hardly necessary in condensed matter physics. Historically, the latter framework (quantum electrodynamics, QED) was established first, with subsequent application to condensed matter. The phonon is the quantum of the field or a wave excited in a real medium made of atoms, and so it cannot be identified as a real particle that travels in the vacuum. Suppose the medium is infinitely large, the lattice spacing so small that we cannot realize its discreteness by any technology we can use, and the lattice is invisible. Then the phonon has every reason to qualify as a real particle flying in the vacuum. Whatever the properties of the vacuum that surrounds us may be, we may consider it to be made of such a superlattice,⁹⁾ and has the ability to produce waves and transmit them.

The caveats I warned about are as follows. The mechanical field model can create a longitudinal wave whereas the electromagnetic wave in vacuum is always transverse. The fields we deal with in particle physics have infinite degrees of freedom while the lattice degrees of freedom are many but finite because of the finite size and spacing, though for all practical purposes it can be treated as infinite. Absolute zero temperature cannot be realized, and the lattice is perpetually in thermal excitation. The metal environment provides electrons with a hot bath of phonons. When an electron travels, it is constantly scattered by the vibrating lattice or phonons. The net effect is a velocity-dependent resistive force, and the velocity saturates when the resistive force cancels the local electric force.

2.2.5

Mechanical Model of a Classical Field

Here, we show that when the field is a collection of an infinite number of connected harmonic oscillators distributed continuously in space, it satisfies the Klein–Gordon equation, which is the fundamental equation of motion for relativistic fields:

$$\left[\frac{\partial^2}{\partial t^2} - \nabla^2 + \mu^2 \right] \varphi(t, \mathbf{x}) \equiv [\partial_\mu \partial^\mu + \mu^2] \varphi(x) = 0 \quad (2.44)$$

In order to simplify the mathematics, we consider a one-dimensional oscillator (Figure 2.3) consisting of $N + 1$ point masses m with interval a spanning total length $D = (N + 1)a$. Let u_{n-1} , u_n , u_{n+1} be the displacements of the $(n - 1)$ th, n th and $(n + 1)$ th mass points, respectively. Then their equation of motion is

9) We use this as a metaphor. In condensed matter physics superlattice has its own meaning.

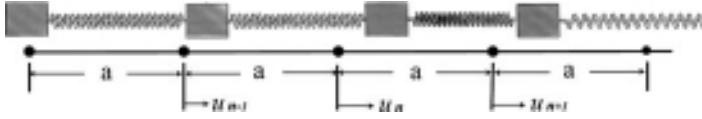


Figure 2.3 $N + 1$ mass points are connected by N springs of length a . u_{n-1} , u_n , u_{n+1} are displacements of the $(n - 1)$ th, n th and $(n + 1)$ th points, respectively.

expressed as

$$m \frac{d^2 u_n}{dt^2} = T \left(\frac{u_{n+1} - u_n}{a} \right) - T \left(\frac{u_n - u_{n-1}}{a} \right) \quad (2.45)$$

Here, we have assumed the springs had been prestretched to a length a to give tension T and that the additional increase of the tension is proportional to its relative stretch. Equation (2.45) can be derived from the Lagrangian

$$L = K - V = \sum_n \left[\frac{m}{2} \left(\frac{du_n}{dt} \right)^2 - \frac{T}{2a} (u_{n+1} - u_n)^2 \right] \quad (2.46)$$

where K is the kinetic energy and V is the potential energy, by using Euler's equation

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{u}_n} \right) - \left(\frac{\partial L}{\partial u_n} \right) = 0 \quad (2.47)$$

where $\dot{u} = du/dt$.

Making N large and keeping the tension constant, we put $D = (N+1)a$, $x = na$, $\Delta x = a$, and rename u_{n-1} , u_n , u_{n+1} as $u(x - \Delta x)$, $u(x)$, $u(x + \Delta x)$. Then the equation becomes

$$m \frac{\partial^2 u}{\partial t^2}(x, t) = T \left[\frac{u(x + \Delta x, t) - u(x, t)}{\Delta x} - \frac{u(x, t) - u(x - \Delta x, t)}{\Delta x} \right] \quad (2.48)$$

Divide both sides by $a = \Delta x$, define linear mass density $\sigma = m/a$ and use $\{u(x + \Delta x) - u(x)\}/\Delta x \simeq \partial u(x)/\partial x$. Then we obtain a wave equation for a wave propagating with velocity v :

$$\sigma \frac{\partial^2 u}{\partial t^2}(x, t) = T \frac{1}{\Delta x} \left[\frac{\partial u}{\partial x}(x, t) - \frac{\partial u}{\partial x}(x - \Delta x, t) \right] \simeq T \frac{\partial^2 u}{\partial x^2}(x, t) \quad (2.49)$$

$$\therefore \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2}(x, t) - \frac{\partial^2 u}{\partial x^2}(x, t) = 0, \quad v = \sqrt{\frac{T}{\sigma}} \quad (2.50)$$

If we consider N oscillators under the influence of gravity as in Figure 2.4, the Lagrangian (2.46) becomes [Problem 2.5]

$$L = \sum_n \left[\frac{m}{2} \left(\frac{du_n}{dt} \right)^2 - \frac{T}{2a} (u_{n+1} - u_n)^2 - \frac{mg}{2\ell} u_n^2 \right] \quad (2.51)$$

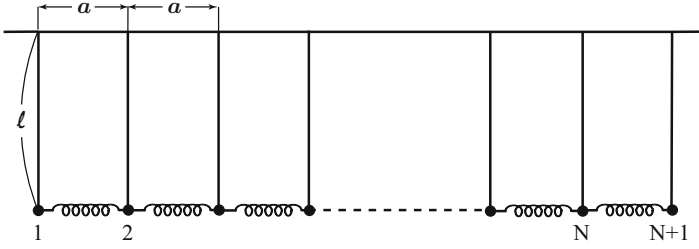


Figure 2.4 Connected pendulums under the influence of gravity (ℓ = length of the pendulums).

where g = gravitational acceleration, and in the continuum limit

$$\frac{1}{v^2} \frac{\partial^2 u}{\partial t^2}(x, t) - \frac{\partial^2 u}{\partial x^2}(x, t) + \frac{\omega_0^2}{v^2} u(x, t) = 0, \quad \omega_0^2 = \frac{g}{\ell} \quad (2.52)$$

The equation has exactly the same form as the Klein–Gordon equation, as one can see by the substitution $v \rightarrow c$, $\omega_0/v \rightarrow \mu$.

Problem 2.5

Prove that the Lagrangian (2.46) is modified to Eq. (2.51) when the oscillators are connected pendulums of length ℓ (Fig. 2.4). Prove, in the continuum limit, that the equation of motion becomes Eq. (2.52).

Problem 2.6

Show that the total energy of $N + 1$ pendulums is given by

$$E = \sum_n \left[\frac{m}{2} \left(\frac{\partial u_n}{\partial t} \right)^2 + \frac{1}{2} \frac{T}{a} (u_{n+1} - u_n)^2 + \frac{m}{2} \frac{g}{\ell} u_n^2 \right] \quad (2.53)$$

In the continuum limit, the Lagrangian is given by

$$L = \int_0^L dx \left[\frac{\sigma}{2} \left\{ \frac{\partial u}{\partial t}(x, t) \right\}^2 - \frac{T}{2} \left\{ \frac{\partial u}{\partial x}(x, t) \right\}^2 - \frac{\sigma \omega_0^2}{2} (u(x, t))^2 \right] \quad (2.54)$$

Redefining $\varphi = \sqrt{T}u$, $\mu = \sqrt{\omega_0^2/v^2}$ and recovering original 3-dimensionality, the Lagrangian and the energy become

$$L \equiv \int d^3x \mathcal{L} = \iiint dx dy dz \frac{1}{2} \left[\frac{1}{v^2} \left(\frac{\partial \varphi}{\partial t} \right)^2 - (\nabla \varphi \cdot \nabla \varphi) - \mu^2 \varphi^2 \right] \quad (2.55)$$

$$E = \int d^3x \frac{1}{2} \left[\frac{1}{v^2} \left(\frac{\partial \varphi}{\partial t} \right)^2 + (\nabla \varphi \cdot \nabla \varphi) + \mu^2 \varphi^2 \right] \quad (2.56)$$

The corresponding Hamiltonian can be obtained by

$$H = \int d^3x \mathcal{H} = \int d^3x [p\dot{\phi} - \mathcal{L}] \quad p = \frac{\delta \mathcal{L}}{\delta \dot{\phi}} \quad (2.57)$$

which of course reproduces Eq. (2.56).

Taking the limit $D \rightarrow \infty$ and putting $v \rightarrow c$ reproduces the Lagrangian for the Klein–Gordon equation:

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi + \mu^2 \phi = 0 \quad (2.58)$$

The solution to the Klein–Gordon equation may be expressed as a superposition of plane waves (mathematically as the Fourier transform)

$$\phi(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int d^3k \tilde{\phi}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \quad \omega^2 = c^2(|\mathbf{k}|^2 + \mu^2) \quad (2.59)$$

When the Klein–Gordon equation is quantized as described in Sect. 5.3, μ^2 becomes the mass of the quantum.

2.3

Summary

We have visualized the field that obeys the Klein–Gordon equation as three-dimensional connected springs distributed over all space. They have the following characteristics.

1. A classical field is a collection of an infinite number of harmonic oscillators numbered by space coordinate $\mathbf{x} = (x, y, z)$ interacting with neighboring oscillators.
2. The amplitude difference in the discrete oscillators becomes differential in the continuum limit. This means an oscillator at point $\mathbf{x}$ interacts with neighboring oscillators through spatial derivatives. However, as is well known in classical mechanics, an infinite number of connected (i.e. interacting) harmonic oscillators is equivalent to the same number of “normal modes”, which oscillate independently of each other. The normal modes are plane waves, given in (2.59), and the interaction with the neighboring oscillators in the original mode is converted simply to propagation of the wave in the normal mode. An infinite number of interacting harmonic oscillators is equivalent to a collection of freely propagating waves numbered by the wave number $\mathbf{k}$. In short, an alternative way of viewing the field is to regard it as a collection of an infinite number of harmonic oscillators numbered by the wave number $\mathbf{k}$.
3. Plane waves show particle behavior when they are quantized. As their energy-momentum is given by $E = \hbar\omega$, $\mathbf{p} = \hbar\mathbf{k}$, the mass m of the quantum is given by

$$m = \frac{\sqrt{E^2 - (\mathbf{p}c)^2}}{c^2} = \frac{\hbar\mu}{c} \quad (2.60)$$

It is interesting to note the origin of the mass term μ or ω_0 in the mechanical oscillator model. It did not exist in the medium that excites the wave. It was given as an external force effect acting on all the oscillators uniformly. This has a far-reaching implication. In the Standard Model, we shall learn that all the fundamental particles in nature have vanishing mass originally and acquire finite mass through interactions with the Higgs field, which pervades the surrounding space. Namely, the mass of an elementary particle is not an intrinsic property of its own but is an acquired attribute as a result of environmental changes. This concept served as a major breakthrough in elucidating the deep roots of gauge theory and so reaching unified theories.

2.4

Natural Units

Universal constants such as $\hbar$ and c appear frequently in equations. In order to avoid writing these constants repeatedly and risking mistakes in the expression, we introduce “natural units”. All physical quantities have dimensions. They are expressed in terms of length, mass and time. The familiar MKS units we use are meter (m) for the length L , kilogram (kg) for the mass M and second (s or sec) for the time T . The dimension of any physical quantity Q is expressed as

$$[Q] = [L^a M^b T^c] \quad (2.61)$$

Velocity, for instance, has the dimension

$$[v] = [LT^{-1}] = \text{m/s} \quad (2.62)$$

Similarly, the dimensions of other mechanical quantities, e.g. acceleration α , force f and energy E , are expressed as

$$[\alpha] = [LT^{-2}] = \text{m/s}^2 \quad (2.63a)$$

$$[f] = [MLT^{-2}] = \text{kg} \cdot \text{m/s}^2 \equiv \text{Newton} \quad (2.63b)$$

$$[E] = [ML^2T^{-2}] = \text{kg} \cdot \text{m}^2/\text{s}^2 \equiv \text{Joule} \quad (2.63c)$$

One need not, however, necessarily adopt MLT as the basic units to construct the dimensions of physical quantities. Any three independent physical quantities are equally eligible. Besides, MKS units are too large for applications in particle physics. We may choose energy as one of the basic units and express it in electron volts (eV), an amount of energy obtained when an electron or proton of unit charge is accelerated by 1 V of electric potential. As the unit electric charge is 1.6021×10^{-19} Coulomb

$$1 \text{ eV} = 1.6021 \times 10^{-19} \text{ Coulomb} \cdot \text{Volt} = 1.6021 \times 10^{-19} \text{ Joules} \quad (2.64)$$

Moreover as the extension of electronvolts

$$\begin{aligned}
 1 \text{ keV} &= 10^3 \text{ eV} \\
 1 \text{ MeV} &= 10^6 \text{ eV} \\
 1 \text{ GeV} &= 10^9 \text{ eV} \\
 1 \text{ TeV} &= 10^{12} \text{ eV}
 \end{aligned} \tag{2.65}$$

are also used, depending on the kind of phenomena concerned. For the time being, we will use MeV as the standard energy unit. In quantum mechanics and relativity theory, $\hbar = h/2\pi$, where h is Planck's constant and c the velocity of light, frequently appears. Adopting c , $\hbar$ and energy as the three basic units, the description of physical quantities is simplified, as we describe now. Since

$$\hbar = 1.055 \times 10^{-34} \text{ J s (Joules} \cdot \text{s)} = 6.582 \times 10^{-22} \text{ MeV s} \tag{2.66a}$$

$$\hbar c = 1.9733 \text{ MeV} \times 10^{-13} \text{ m} \tag{2.66b}$$

Length and time can be expressed as

$$\begin{aligned}
 [L] &= [\hbar c] \times [E^{-1}] \\
 [T] &= [\hbar] \times [E^{-1}]
 \end{aligned} \tag{2.67}$$

Namely, length is measured by inverse energy E^{-1} in units of $1.9733 \times 10^{-13} \text{ m}$. Time is also measured by inverse energy E^{-1} in units of $6.582 \times 10^{-22} \text{ s}$. Now we define natural units by putting $\hbar = c = 1$. Then the constants $\hbar$ and c disappear from all the equations. The velocity v becomes a dimensionless quantity measured in units of the velocity of light, which we had to write as $\beta = v/c$ previously. Every dimensionful quantity is expressed in terms of energy or inverse energy. The mass multiplied by c^2 and the momentum multiplied by c become energy, so they are all expressed in MeV. Sometimes the mass is expressed in MeV/c^2 , and the momentum in MeV/c , to differentiate them from real energy. To recover the original physical dimension, we only need to reconstruct the right dimension by multiplying or dividing by $\hbar$ and c . For instance, the pion mass m_π is 140 MeV, and we talk about distance or time as $1/m_\pi$. Then

$$\begin{aligned}
 \frac{1}{m_\pi} &= \frac{\hbar c [\text{MeV} \cdot \text{m}]}{140 \text{ MeV}} = \frac{1.97 \times 10^{-13} \text{ m}}{140} = 1.4 \times 10^{-15} \text{ m} \\
 \frac{1}{m_\pi} &= \frac{\hbar [\text{MeV} \cdot \text{s}]}{140 \text{ MeV}} = \frac{6.58 \times 10^{-22} \text{ s}}{140} = 4.7 \times 10^{-24} \text{ s}
 \end{aligned} \tag{2.68}$$

When we talk of a cross section σ being $1/m_p^2$, its value in MKS units is given by

$$\sigma = \frac{1}{m_p^2} = \frac{(\hbar c)^2}{(m_p c^2)^2} = \left\{ \frac{1.97 \times 10^{-13} [\text{MeV} \cdot \text{m}]}{938.3 \text{ MeV}} \right\}^2 = 4.4 \times 10^{-28} \text{ cm}^2 \tag{2.69}$$

Problem 2.7

Show the value of the classical electron radius r_e is 2.82×10^{-15} m, where r_e is given by α/m_e . $\alpha = 1/137$ is the fine structure constant and m_e , the electron mass, is 0.511 MeV.

Problem 2.8

The neutrino changes its flavor periodically while it propagates, a phenomenon called neutrino oscillation. The probability of a ν_e changing into ν_μ at time t is given by

$$P(\nu_e \rightarrow \nu_\mu : t) \propto \sin^2 \frac{\Delta m^2}{4E} t \quad (2.70a)$$

$$\Delta m^2 = |m_1^2 - m_2^2| \quad (2.70b)$$

If Δm^2 is given in units of $(\text{eV})^2$, E in MeV, the length in meters, show that the oscillation wavelength L is given by

$$L[\text{m}] = \frac{4\pi E}{\Delta m^2} = 2.4 \times \frac{E [\text{MeV}]}{\Delta m^2 [\text{eV}^2]} \quad (2.71)$$

In electromagnetism, we have used, so far, rationalized MKS units. From now on we will use Heaviside–Lorentz units, which put $\epsilon_0 = \mu_0 = 1$ and natural units instead of MKS units.

3

Lorentz Invariance

In particle physics, most particles we treat are relativistic, i.e. $E \approx pc \gg mc^2$, and the theory of special relativity is an indispensable mathematical tool. A notable feature is that space and time mix and cannot be treated independently as in Newtonian mechanics. Physical objects that were considered as an independent three-component vector and a scalar in nonrelativistic physics mix in high-energy phenomena. They have to be combined to make a four-component Lorentz vector that transforms like a time and space coordinate (its difference to be exact). For their consistent and unified treatment, we have to rely on Einstein's theory of relativity, which is based on two principles.

1. Invariance of the velocity of light, which postulates that the velocity of light always remains as the constant c in any inertial frame.
2. Relativity principle, which requires covariance of the equations, namely the physical law should keep its form invariant in any inertial frame. In mathematical language this amounts to the fact that physical laws have to be expressed in Lorentz tensors.

The relativity principle applies to Galilei transformation and is valid in Newtonian mechanics also, but the invariance of the velocity of light necessitates Lorentz transformation in changing from one inertial system to another that are moving relative to each other with constant speed.

3.1

Rotation Group

Before introducing Lorentz transformation operators, it is useful first to review the familiar rotation operator in three dimensions, which makes a group $O(3)$. For the definition of a group, see Appendix A. The basic operations are very similar, and in fact the Lorentz group contains the rotation group as its subgroup. Let $\mathbf{r}_B = (x_B, y_B, z_B)$ be a point in a coordinate frame L that can be obtained from point $\mathbf{r}_A$

by rotation through angle θ . Then in matrix form

$$\mathbf{r}_B \equiv \begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} = [R_{ij}] \begin{bmatrix} x_A \\ y_A \\ z_A \end{bmatrix} \equiv \mathbf{R} \mathbf{r}_A \quad (3.1)$$

As the rotation does not change the length $r = \sqrt{x^2 + y^2 + z^2}$,

$$r_B^2 = \mathbf{r}_B^T \mathbf{r}_B = \mathbf{r}_A^T \mathbf{R}^T \mathbf{R} \mathbf{r}_A = r_A^2 = \mathbf{r}_A^T \mathbf{r}_A \quad (3.2a)$$

$$\therefore \quad \mathbf{R}^T \mathbf{R} = \mathbf{1} \quad (3.2b)$$

Namely, $\mathbf{R}$ is a 3×3 orthogonal matrix. A set of three variables that transforms similarly by rotation is called a vector and is expressed as $\mathbf{V} = (V_x, V_y, V_z)$. As an example, let us write down a rotation around the z -axis by angle θ (Figure 3.1a).

$$\begin{bmatrix} V'_x \\ V'_y \\ V'_z \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} \Leftrightarrow \mathbf{V}' = R_z(\theta) \mathbf{V} \quad (3.3)$$

Here, we have defined the rotation as moving the vector from point A to point B referred to as active transformation. Some authors use the word rotation to mean transformation of the coordinate frame without actually moving the vector, which we refer to as passive transformation. Mathematically, the active rotation is equivalent to passive rotation by angle $-\theta$ (Figure 3.1b), i.e. the passive transformation is an inverse active transformation. We use active transformation throughout this book. What we are concerned with is the effect of the operation on a “vector field $\mathbf{V}(\mathbf{x})$ ”, a vector defined at every space coordinate $\mathbf{x}$. The effect can be expressed as

$$\mathbf{V}'(\mathbf{x}) = D(\mathbf{R}) \mathbf{V}(\mathbf{R}^{-1} \mathbf{x}) \quad (3.4a)$$

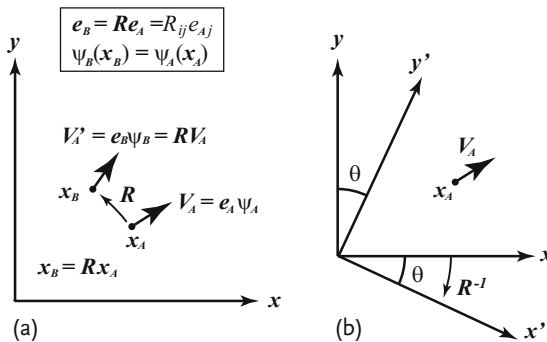


Figure 3.1 (a) Active transformation moves the vector. (b) Passive transformation moves the coordinate frame.

where $D(\mathbf{R})$ acts on the components of the vector. Expressed in terms of components

$$V'_i(x_k) = \sum_j R_{ij} V_j \left(\sum_l R_{kl}^{-1} x_l \right) \quad (3.4b)$$

To prove the above equation, let $\mathbf{V}'(\mathbf{x})$ be a new vector field obtained from the old vector field $\mathbf{V}(\mathbf{x})$ by rotation. Both fields are defined at all space points $\mathbf{x}$. Let $\mathbf{V}_B = \mathbf{V}(\mathbf{x}_B)$ and $\mathbf{V}_A = \mathbf{V}(\mathbf{x}_A)$ and $\mathbf{V}'_A = \mathbf{V}'(\mathbf{x}_B)$. The vector $\mathbf{V}'_A$ is obtained by rotating $\mathbf{V}_A$ from point $\mathbf{x}_A$ to point $\mathbf{x}_B$ (Figure 3.1a). Note that the vector $\mathbf{V}_B$ is the old vector field evaluated at $\mathbf{x}_B$ and generally differs from $\mathbf{V}_A$ or $\mathbf{V}'_A$ in both magnitude and direction. We can write down three-component vectors in factorized form, namely as a product of their magnitude and a direction vector

$$\mathbf{V}_A = |\mathbf{V}_A| \mathbf{e}_A, \quad \mathbf{V}'_A = |\mathbf{V}'_A| \mathbf{e}_B \quad (3.5)$$

where $\mathbf{e}_A = \mathbf{V}_A/|\mathbf{V}_A|$ and $\mathbf{e}_B = \mathbf{V}_B/|\mathbf{V}_B|$ are two base vectors of unit length connected by the rotation. $|\mathbf{V}|$ is the magnitude or length of the vector. The base vectors are specified by their direction only. Since the magnitude of the vector does not change by rotation, but the direction of the base vector changes, we have

$$|\mathbf{V}'_A| = \mathbf{R}|\mathbf{V}_A| = |\mathbf{V}_A| = |\mathbf{V}(\mathbf{x}_A)| = |\mathbf{V}(\mathbf{R}^{-1}\mathbf{x}_B)| \quad (3.6a)$$

$$\mathbf{e}_B = \mathbf{R}\mathbf{e}_A \quad \text{i.e.} \quad (e_B)_k = \sum_l R_{kl}(e_A)_l \equiv D(\mathbf{R})\mathbf{e}_A \quad (3.6b)$$

$$\begin{aligned} \therefore \mathbf{V}'(\mathbf{x}_B) &= |\mathbf{V}'_A| \mathbf{e}_B = |\mathbf{V}_A| \mathbf{R}\mathbf{e}_A = \mathbf{R}\{|\mathbf{V}(\mathbf{x}_A)| \mathbf{e}_A\} = \mathbf{R}\mathbf{V}(\mathbf{R}^{-1}\mathbf{x}_B) \\ &= R_{kl}[\mathbf{V}(\mathbf{R}^{-1}\mathbf{x}_B)]_l = D(\mathbf{R})\mathbf{V}(\mathbf{R}^{-1}\mathbf{x}_B) \end{aligned} \quad (3.6c)$$

Rewriting $\mathbf{x}_B$ as $\mathbf{x}$, we obtain Eqs. (3.4). When the rotation angle around the z -axis is infinitesimal, we can obtain $D(R_z)(\delta\theta)$ from Eq. (3.3).

$$D(R_z)(\delta\theta) = \begin{bmatrix} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 1 - i\delta\theta \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 1 - i\delta\theta S_z \quad (3.7a)$$

$$S_z \equiv i \frac{dR_z}{d\theta} \Big|_{\theta=0} \quad (3.7b)$$

We can also extract an operator that carries out the reverse rotation on the arguments as follows. For any function of space point $\mathbf{x}$

$$\begin{aligned} f(\mathbf{R}_z^{-1}\mathbf{x}) &= f(\mathbf{x} + \delta\theta \gamma, \gamma - \delta\theta \mathbf{x}, z) \\ &= f(\mathbf{x}, \gamma, z) + \delta\theta (\gamma \partial_x - \mathbf{x} \partial_\gamma) f(\mathbf{x}, \gamma, z) \\ &= (1 - i\delta\theta L_z) f(\mathbf{x}, \gamma, z) \\ L_z &= -i(\mathbf{x} \partial_\gamma - \gamma \partial_x) \end{aligned} \quad (3.8)$$

L_z is the familiar orbital angular momentum operator and S_z is the spin angular momentum operator. Combining the operations on the coordinates and components

$$\begin{aligned} R_z(\delta\theta)V(R_z^{-1}\mathbf{x}) &= (1 - i\delta\theta J_z)V(\mathbf{x}) \\ J_z &= L_z + S_z \end{aligned} \quad (3.9)$$

Rotation through a finite angle θ can be obtained by repeating the infinitesimal rotation $\delta\theta = \theta/n$ n times and making $n \rightarrow \infty$:

$$R_z(\theta) = \lim_{n \rightarrow \infty} \prod \left(1 - iJ_z \frac{\theta}{n} \right)^n = e^{-i\theta J_z} \quad (3.10)$$

J_z is called a “generator” of rotation on axis z . Other spin generators are defined similarly.

$$S_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad S_y = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} \quad S_z = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (3.11)$$

They satisfy the following commutation relations:

$$[J_i, J_j] = i\epsilon_{ijk}J_k \quad (3.12a)$$

$$[L_i, L_j] = i\epsilon_{ijk}L_k \quad (3.12b)$$

$$[S_i, S_j] = i\epsilon_{ijk}S_k \quad (3.12c)$$

$$\epsilon_{ijk} = \begin{cases} +1 & ijk = \text{even permutation of } 123 \\ -1 & ijk = \text{odd permutation of } 123 \end{cases} \quad (3.13)$$

The relation Eq. (3.12) is called Lie algebra and ϵ_{ijk} is called the structure function of the rotation group $O(3)$. The rotation can be specified by a set of three parameters, a unit vector to define the rotation axis and the rotation angle θ , and is expressed as a rotation vector $\boldsymbol{\theta}$. Generally, the 3×3 rotation matrix has nine variables, but they are reduced to three because of the orthogonality relation (3.2b). It is known that essential properties¹⁾ of the rotational group $O(3)$ (or more generally those of a Lie group) is completely determined by the three generators and their commutation relations (i.e. structure constant). The rotation operator is expressed as

$$\mathbf{R}(\mathbf{n}, \theta) = e^{-i\mathbf{J} \cdot \boldsymbol{\theta}} = e^{-i\mathbf{J} \cdot \mathbf{n}\theta} \quad (3.14)$$

1) Meaning local properties. Global properties may differ. For instance $SU(2)$ and $O(3)$ have the same algebra but there is 2:1 correspondence between the two. See Appendix A.

3.2

Lorentz Transformation

3.2.1

General Formalism

Introducing a coordinate variable $x^0 = ct$, we define two kinds of Lorentz vector

$$dx^\mu = (dx^0, dx^1, dx^2, dx^3) = (cdt, dx, dy, dz) \quad (3.15a)$$

$$dx_\mu = (dx_0, dx_1, dx_2, dx_3) = (cdt, -dx, -dy, -dz) \quad (3.15b)$$

Lorentz transformation is defined as one which keeps the Minkowski metric, the scalar product of the two vectors defined by

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \equiv dx^\mu dx_\mu \equiv g_{\mu\nu} dx^\mu dx^\nu \quad (3.16)$$

invariant. Here we adopted the convention (Einstein's contraction) that when the same Greek index appears simultaneously as a superscript and a subscript, summation from 0 to 3 is assumed. " $g_{\mu\nu}$ " defined by Eq. (3.16) is called the metric tensor and is expressed in matrix form as

$$\mathbf{G} \equiv [g_{\mu\nu}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad (3.17)$$

$$\mathbf{G}^{-1} \equiv [g^{\mu\nu}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Minkowski space is defined as a space having the metric (length) defined by Eq. (3.16). In Minkowski space the inverse of the metric tensor is the same as the metric itself, a feature which does not hold in general.²⁾ In general relativity, the metric tensor is a function of space-time coordinates and the Einstein field equation is a differential equation for $g_{\mu\nu}$. Introducing a column vector $\mathbf{x}$ for the vector with upper index and row vector $\mathbf{x}^T$ for the vector with lower index, the (homogeneous) Lorentz transformation can be expressed in the form

$$x^{\mu'} = L^\mu_\nu x^\nu, \quad \mathbf{x}' = \begin{bmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{bmatrix} = \mathbf{L}\mathbf{x}, \quad \mathbf{L} = [L^\mu_\nu] = \frac{\partial x^{\mu'}}{\partial x^\nu} \quad (3.18)$$

- 2) The distinction of the upper and lower index plays a minor role in Minkowski space. In fact, in the old convention, $x^4 = ict$ was used to keep the metric in the same form as the Euclidian definition of length without distinction of upper and lower indexes. But it is essential in general relativity.

The general homogeneous Lorentz transformation includes space rotation and constant velocity parallel transformation, called (Lorentz) boost, which are specified by two vectors, $\boldsymbol{\omega}$ and $\mathbf{v}$. The magnitude of the rotation vector $\boldsymbol{\omega}$ represents a rotation angle, and its direction the rotation axis. Consider the special case of Lorentz boost in the x direction. Here, a particle at (t, x, y, z) in a coordinate frame L is boosted to (t', x', y', z') with velocity v . This statement is equivalent to changing to another coordinate frame L' which is moving in the x direction at velocity $-v$. L' is assumed to coincide with L at $t = t' = 0$. Then the two coordinates are related by

$$t \rightarrow t' = \frac{t + (v/c^2)x}{\sqrt{1 - (v/c)^2}} \Rightarrow x'^0 = \gamma(x^0 + \beta x) \quad (3.19a)$$

$$x \rightarrow x' = \frac{x + vt}{\sqrt{1 - (v/c)^2}} \Rightarrow x'^1 = \gamma(\beta x^0 + x) \quad (3.19b)$$

$$x'^2 = x^2, \quad (3.19c)$$

$$x'^3 = x^3, \quad \beta = \frac{v}{c}, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}} \quad (3.19d)$$

or in matrix form

$$\begin{bmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{bmatrix} = \begin{bmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \quad (3.20)$$

A combination of four variables that transforms like the coordinate x^μ (more rigorously the difference of the coordinate dx^μ) is called a contravariant vector and is denoted as A^μ . A covariant vector is defined as

$$A_\mu = g_{\mu\nu} A^\nu = (A^0, -A^1, -A^2, -A^3) \equiv (A^0, -\mathbf{A}) \quad (3.21)$$

Conversely, the contravariant vector can be defined using the inverse of the metric tensor

$$A^\mu = g^{\mu\nu} A_\nu = (A_0, -A_1, -A_2, -A_3), \quad g^{\mu\rho} g_{\rho\nu} = \delta^\mu_\nu \quad (3.22)$$

In the nonrelativistic limit, the 0th (time) component and the space components are separated, the time component behaving as a scalar under space rotation, hence there is no distinction between the two types of vectors. However, under the Lorentz transformation, the distinction has to be made. The contra- and covariant vectors in Minkowski space can be exchanged simply by changing the sign of the space components, an operation called parity transformation. The Lorentz transformation of the covariant vector is expressed as

$$A'_\mu = A_\nu M^\nu_\mu, \quad \text{or} \quad \mathbf{A}'^T = \mathbf{A}^T \mathbf{M} \quad (3.23)$$

Putting $A_\mu = dx_\mu$,

$$dx'_\mu dx'^{\mu'} = (dx^T \mathbf{M})(\mathbf{L} dx) = dx_\mu dx^\mu = dx^T dx \quad (3.24a)$$

$$\therefore \mathbf{M} = \mathbf{L}^{-1}, \quad M_\mu^\nu = \frac{\partial x^\nu}{\partial x'^\mu} \quad (3.24b)$$

which means the Lorentz transformation of the covariant vector is given by the inverse transformation. In a boost like Eq. (3.19), it is obtained simply by the substitution $\mathbf{v} \rightarrow -\mathbf{v}$.

Problem 3.1

Show that a boost L_v in general direction (along $\mathbf{v}$) is given by

$$\begin{aligned} L_v &= \begin{bmatrix} \gamma & \boldsymbol{\beta}\gamma \\ \boldsymbol{\beta}\gamma & 1 + \hat{\boldsymbol{\beta}} \otimes \hat{\boldsymbol{\beta}}(\gamma - 1) \end{bmatrix} \\ &= \begin{bmatrix} \gamma & \beta_x \gamma & \beta_y \gamma & \beta_z \gamma \\ \beta_x \gamma & \hat{\beta}_x^2 \gamma + \hat{\beta}_y^2 + \hat{\beta}_z^2 & \hat{\beta}_x \hat{\beta}_y (\gamma - 1) & \hat{\beta}_x \hat{\beta}_z (\gamma - 1) \\ \beta_y \gamma & \hat{\beta}_y \hat{\beta}_x (\gamma - 1) & \hat{\beta}_x^2 + \hat{\beta}_y^2 \gamma + \hat{\beta}_z^2 & \hat{\beta}_y \hat{\beta}_z (\gamma - 1) \\ \beta_z \gamma & \hat{\beta}_z \hat{\beta}_x (\gamma - 1) & \hat{\beta}_z \hat{\beta}_y (\gamma - 1) & \hat{\beta}_x^2 + \hat{\beta}_y^2 + \hat{\beta}_z^2 \gamma \end{bmatrix} \\ (\hat{\boldsymbol{\beta}} \otimes \hat{\boldsymbol{\beta}})_{ij} &\equiv \hat{\beta}_i \hat{\beta}_j, \quad \hat{\boldsymbol{\beta}} = \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \end{aligned} \quad (3.25)$$

Hint: Separate the three-dimensional vector in the longitudinal ($\parallel \boldsymbol{\beta}$) and the transverse directions ($\perp \boldsymbol{\beta}$)

$$\mathbf{x} = [x - \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}} \cdot \mathbf{x})] + \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}} \cdot \mathbf{x}) \quad (3.26)$$

and apply the boost to the longitudinal part.

3.2.2

Lorentz Vectors and Scalars

We define Lorentz vectors and Lorentz scalars, which will appear frequently in quantum field theory. An important physical variable is the proper time τ . It is the time an observer feels in the observer's rest frame.

Proper Time The proper time $d\tau \equiv dt \sqrt{1 - \beta^2}$ is a Lorentz-invariant scalar.

Proof: $ds^2 = (cdt)^2 - dx^2 - dy^2 - dz^2 = c^2 dt^2 \{1 - (dx/dt)^2 - (dy/dt)^2 - (dz/dt)^2\} = (cdt)^2(1 - \beta^2) = (cd\tau)^2$

is Lorentz invariant by definition. $\square$

Since a vector multiplied or divided by a scalar is also a vector, it immediately follows that the variables in Table 3.1 are Lorentz vectors.

Table 3.1 Useful Lorentz vectors.

Name	Contravariant vector	Covariant vector	
Coordinate difference	dx^μ	dx_μ	(3.27a)
Velocity	$u^\mu = dx^\mu/d\tau = (\gamma c, \gamma \mathbf{v})$	$u_\mu = dx_\mu/d\tau$	(3.27b)
Acceleration	$\alpha^\mu = du^\mu/d\tau$	$\alpha_\mu = du_\mu/d\tau$	(3.27c)
Energy-momentum	$p^\mu = m u^\mu = (E/c, \mathbf{p})$	$p_\mu = (E/c, -\mathbf{p})$	(3.27d)
Wave vector	$k^\mu = (\omega/c, \mathbf{k})$	$k_\mu = (\omega/c, -\mathbf{k})$	(3.27e)
Current	$j^\mu = \rho_0 u^\mu = (\rho c, \mathbf{j})^a$	$j_\mu = (\rho c, -\mathbf{j})$	(3.27f)
Electromagnetic potential	$A^\mu = (\phi/c, \mathbf{A})$	$A_\mu = (\phi/c, -\mathbf{A})$	(3.27g)
Differential operator	$\partial^\mu = \frac{\partial}{\partial x_\mu} = (\frac{1}{c} \frac{\partial}{\partial t}, -\nabla)^b$	$\partial_\mu = \frac{\partial}{\partial x^\mu} = (\frac{1}{c} \frac{\partial}{\partial t}, \nabla)$	(3.27h)

a ρ_0 is (charge) distribution in its rest frame.

b Notice the minus sign of the space components of the differential vector.

Notice, space components of the four-velocity u^μ differ from conventional ones by the factor γ .

Problem 3.2

Prove that (3.28e), (3.28g), (3.28h) are Lorentz vectors.

Scalar Product If the scalar product of a given Lorentz vector A^μ and a four-component variable B^μ satisfy the equation

$$A'^\mu B'_\mu = A^\mu B_\mu = A_\mu B^\mu = A'_\mu B'^\mu = \text{constant} \quad (3.28)$$

then B^μ is a Lorentz vector. We shall denote the scalar product as $A \cdot B \equiv A_\mu B^\mu$.³⁾
Proof:

$$A'^\mu B'_\mu = L^\mu_\rho A^\rho B'_\mu = A^\rho B_\rho \quad \therefore \quad B'_\mu L^\mu_\rho = B_\rho$$

Multiplying by $M^\rho_\sigma = [L^{-1}]^\rho_\sigma$ from the right

$$B'_\mu L^\mu_\rho M^\rho_\sigma = B'_\mu \delta^\mu_\sigma = B'_\sigma = B_\rho M^\rho_\sigma$$

which shows B_σ is a covariant vector. □

By making scalar products with two vectors, it is shown that variables in Table 3.2 are Lorentz-invariant quantities.

3) We denote the usual scalar product in three-dimensional space by bold letters, i.e. $\mathbf{A} \cdot \mathbf{B}$, and the Lorentz scalar product by normal letters, $A \cdot B$.

Table 3.2 Useful Lorentz scalars.

Name	Scalar
Einstein's relation	$p^2 = p^\mu p_\mu = (E/c)^2 - \mathbf{p}^2 = (mc)^2$ (3.29a)
d'Alembertian	$\partial_\mu \partial^\mu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 = \square$ (3.29b)
Phase of a wave	$kx \equiv k_\mu x^\mu = \omega t - \mathbf{k} \cdot \mathbf{x}$ (3.29c)
Current conservation law	$\partial_\mu j^\mu = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0$ (3.29d)
Phase space volume	$\theta(p^0) \delta(p^2 - (mc)^2) d^4 p = \frac{d^3 p}{2E/c}$ (3.29e)
Volume element in Minkowski space	$d^4 x = dx^0 dx^1 dx^2 dx^3$ (3.29f)

Equation (3.29f) can be proved by calculating the Jacobian of the transformation:

$$d^4 x' = \left| \frac{\partial x'}{\partial x} \right| d^4 x = (\det \mathbf{L}) d^4 x = d^4 x \quad [\text{See Eq. (3.36)}] \quad (3.30)$$

Comment 3.1 In high-energy experiments, where $E \sim pc$, the production cross section is often expressed in approximately Lorentz-invariant form

$$\frac{d^3 \sigma}{c^3 d^3 p / E} \simeq \frac{1}{E} \frac{d^2 \sigma}{dE d\Omega} \quad (3.31)$$

3.3

Space Inversion and Time Reversal

Since the Lorentz transformation keeps the metric ds^2 invariant in any inertial coordinate frame,

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu = g_{\mu\nu} dx^{\mu'} dx^{\nu'} = g_{\mu\nu} L_\rho^\mu L_\sigma^\nu dx^\rho dx^\sigma \\ &= g_{\rho\sigma} L_\mu^\rho L_\nu^\sigma dx^\mu dx^\nu \\ \therefore g_{\mu\nu} &= g_{\rho\sigma} L_\mu^\rho L_\nu^\sigma \end{aligned} \quad (3.32)$$

where in the second line, indices were interchanged since it does not make any difference when summation is carried out. It can be rewritten in matrix representation as

$$\mathbf{G} = \mathbf{L}^T \mathbf{G} \mathbf{L} \quad (3.33)$$

Multiplying by $\mathbf{G}^{-1}$ from the left

$$\mathbf{1} = \mathbf{G}^{-1} \mathbf{L}^T \mathbf{G} \mathbf{L} = [\mathbf{G} \mathbf{L} \mathbf{G}^{-1}]^T \mathbf{L} \Rightarrow \mathbf{L}^{-1} = [\mathbf{G} \mathbf{L} \mathbf{G}^{-1}]^T \quad (3.34)$$

where we have used the relations $\mathbf{G}^T = \mathbf{G}$, $[\mathbf{G}^{-1}]^T = \mathbf{G}^{-1}$. From Eq. (3.34)

$$\det \mathbf{L}^{-1} = \det \mathbf{G} \det \mathbf{L} \det \mathbf{G}^{-1} \quad (3.35)$$

Using $\det \mathbf{L}^{-1} = 1/\det \mathbf{L}$, we obtain

$$\det \mathbf{L} = \pm 1 \quad (3.36)$$

We see there are two kinds of Lorentz transformation, depending on the sign of $\det \mathbf{L}$. Taking the (0-0) component of Eq. (3.32),

$$1 = (L_0^0)^2 - \sum_i (L_0^i)^2 \Rightarrow |L_0^0| > 1 \quad (3.37)$$

The transformation with $L_0^0 > 1$ is called orthochronous and the one with $L_0^0 < 1$ is called non-orthochronous, meaning “changes or does not change the time direction”. The Lorentz transformation can be classified into four classes depending on the sign of $\det \mathbf{L}$ and L_0^0 , and that with $\det \mathbf{L} > 0$ and $L_0^0 > 0$ is called the proper Lorentz transformation $\mathbf{L}_p$.

The proper Lorentz transformation becomes a non-transformation or identity transformation in the limit when the transformation parameter becomes 0, namely it is continuously connected with $\mathbf{L} = \mathbf{1}$. It includes boost as well as rotation $\mathbf{R}$ in three-dimensional space. The rotation is expressed as

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbf{R} & \\ 0 & & & \end{bmatrix} \quad (3.38)$$

Defining $\mathbf{P}$, $\mathbf{T}$ and $\mathbf{PT}$ by

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & -1 & \\ 0 & & & \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & 1 & \\ 0 & & & \end{bmatrix},$$

$$\mathbf{PT} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & -1 & \\ 0 & & & \end{bmatrix} \quad (3.39)$$

then $\det \mathbf{L}_p = \det(\mathbf{PT}) = 1$, $\det \mathbf{P} = \det \mathbf{T} = -1$. $\mathbf{P}$ is called space inversion or parity transformation, $\mathbf{T}$ time reversal and $\mathbf{PT}$ total reflection. Their transformation is discrete and is not connected smoothly with the identity transformation. Any Lorentz transformation can be expressed as a combination of the proper transformation with the three discrete transformations and needs six independent parameters (velocity and rotation vector) to specify it.

3.4

Covariant Formalism

3.4.1

Tensors

Products of two vectors transform under Lorentz transformation as

$$A^{\mu'} B^{\nu'} = L^{\mu}_{\rho} L^{\nu}_{\sigma} A^{\rho} B^{\sigma} \quad (3.40a)$$

$$A_{\mu'} B_{\nu'} = M^{\rho}_{\mu} M^{\sigma}_{\nu} A_{\rho} B_{\sigma} \quad (3.40b)$$

$$A^{\mu'} B_{\nu'} = L^{\mu}_{\rho} M^{\sigma}_{\nu} A^{\rho} B_{\sigma} \quad (3.40c)$$

and make contravariant, covariant and mixed tensors of rank 2, which are called the direct product of two vectors. More generally, any set of 16 variables that behaves in the same way as the product of two vectors is called a rank 2 Lorentz tensor. The concept can be extended to arbitrary rank r

$$T^{\mu\nu\cdots}_{\rho\sigma\cdots} \sim A^{\mu} B^{\nu} \cdots P_{\rho} Q_{\sigma} \cdots \quad \text{product of } r \text{ vectors} \quad (3.41)$$

Here, $\sim$ means “transforms like”. By multiplying a tensor by δ^{ρ}_{μ} , $g_{\mu\nu}$ or $g^{\rho\sigma}$ and making a contraction (summation of the same upper and lower indices), the rank r tensor can be converted to a new tensor of rank $r - 2$:

$$\begin{aligned} T^{\mu\nu\cdots}_{\rho\sigma\cdots} \quad (\text{rank } r) \rightarrow & \left. \begin{aligned} \delta^{\rho}_{\mu} T^{\mu\nu\cdots}_{\rho\sigma\cdots} &= T'^{\nu\cdots}_{\sigma\cdots} \\ g_{\mu\nu} T^{\mu\nu\cdots}_{\rho\sigma\cdots} &= T'^{\cdots}_{\rho\sigma\cdots} \\ g^{\rho\sigma} T^{\mu\nu\cdots}_{\rho\sigma\cdots} &= T'^{\mu\nu\cdots}_{\cdots} \end{aligned} \right\} \\ & = \text{rank } r - 2 \text{ tensor} \end{aligned} \quad (3.42)$$

We introduce another useful tensor, called the Levi-Civita tensor, defined as

$$\begin{aligned} \epsilon^{\mu\nu\rho\sigma} &= \begin{cases} \pm 1, & \mu\nu\rho\sigma = \text{even (odd) permutation of } 0123 \\ 0, & \text{any pair of } \mu\nu\rho\sigma \text{ are equal} \end{cases} \\ \epsilon^{0123} &= -\epsilon_{0123} = 1 \end{aligned} \quad (3.43)$$

The Levi-Civita tensor is a Lorentz tensor that transforms to itself because

$$\epsilon^{\mu\nu\rho\sigma'} = L^{\mu}_{\alpha} L^{\nu}_{\beta} L^{\rho}_{\gamma} L^{\sigma}_{\delta} \epsilon^{\alpha\beta\gamma\delta} = \epsilon^{\mu\nu\rho\sigma} \det L = \epsilon^{\mu\nu\rho\sigma} \quad (3.44)$$

3.4.2

Covariance

When equations are made of tensors of the same rank, they keep the same form in any inertial frame. For example,

$$A^\mu = B^\mu \xrightarrow{L} A^{\mu'} = B^{\mu'} \quad (3.45a)$$

$$A_\mu B^{\mu\nu} = C^\nu \xrightarrow{L} A_{\mu'} B^{\mu'\nu'} = C^{\nu'} \quad (3.45b)$$

They are said to be (Lorentz) covariant. By transformation the values of both sides of equations may change, but the functional form is kept invariant. Note, Lorentz covariance is not to be confused with Lorentz invariance in strict sense, which means that the value itself does not change under Lorentz transformation and which applies only to Lorentz scalars. Note, however, many authors including myself often use Lorentz invariance to mean Lorentz covariance. The principle of relativity requires covariance under transformations. In special relativity, the transformation is Lorentz transformation; in general relativity it is general coordinate transformation.

To get used to the usage of covariant formalism, let us take a look at the familiar Maxwell equations:

$$\nabla \cdot \mathbf{E} = q \frac{\rho}{\epsilon_0}, \quad \nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \mu_0 q \mathbf{j} \quad (3.46a)$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \quad \nabla \cdot \mathbf{B} = 0 \quad (3.46b)$$

The electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ can be expressed in terms of potentials:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A} \quad (3.47)$$

Since ∂^μ and $A^\mu = (\phi/c, \mathbf{A})$ are both Lorentz vectors, Equation (3.47) may be rewritten as an antisymmetric tensor of rank 2

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu = \begin{bmatrix} 0 & -E_1/c & -E_2/c & -E_3/c \\ E_1/c & 0 & -B_3 & B_2 \\ E_2/c & B_3 & 0 & -B_1 \\ E_3/c & -B_2 & B_1 & 0 \end{bmatrix} \quad (3.48)$$

This means that electromagnetic fields look like two sets of vectors but actually are different components of a single second-rank tensor. Using $F^{\mu\nu}$, the Maxwell equations (3.46) can be expressed in two sets of tensor equations

$$\partial_\mu F^{\mu\nu} = \mu_0 q j^\nu \quad (3.49a)$$

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \quad (F_{\mu\nu} = g_{\mu\rho} g_{\nu\sigma} F^{\rho\sigma}) \quad (3.49b)$$

Problem 3.3

Confirm Eq. (3.48) and Eq. (3.49).

Problem 3.4

Prove that a dual tensor defined by

$$\tilde{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} \quad (3.50)$$

exchanges $\mathbf{E} \rightarrow \mathbf{B}$, $\mathbf{B} \rightarrow -\mathbf{E}$, and Eq. (3.49b) can be rewritten as

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad (3.51)$$

From the definition of the electromagnetic tensor $F^{\mu\nu}$, the Maxwell equation without sources (3.49b) is actually an identity equation. Namely Eq. (3.46b) is equivalent to saying that the electromagnetic potential A^μ is a Lorentz vector. In Eq. (3.49), the two Maxwell equations are written as rank 1 and rank 3 tensor equations. In this expression, Lorentz covariance of the Maxwell equations is quite evident. Historically the converse held. The Maxwell equations were derived first, then Lorentz transformation was devised to make the equation covariant, which led Einstein to special relativity.

Problem 3.5

By its construction, $f^\mu = q F^{\mu\nu} u_\nu$, where u^μ is defined in Eq. (3.28b), is a contravariant vector. Show that in the nonrelativistic limit its space component reduces to the Lorentz force $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ and hence f^μ is its relativistic extension.

3.4.3**Supplementing the Time Component**

In constructing a covariant equation, sometimes it happens that we know only the space component of a four-vector; the time (zeroth) component is unknown. We did not know the time component of the Lorentz force, but the four-vector defined in Problem 3.5 naturally supplements it. In the following we show two more examples to explicitly construct the yet unknown component of the four-vectors.

Example 3.1: Newton's Second Law

Here, we construct a four-vector f^μ which represents a force whose space component, in the nonrelativistic limit, reduces to $\mathbf{F}$ that appears in Newton's second law. Referring to the list of Lorentz vectors Table 3.1, a Lorentz-covariant extension of

the second law may be expressed as

$$m \frac{du^\mu}{d\tau} = f^\mu, \quad f^\mu = (f^0, \mathbf{f}) \equiv (f^0, \gamma \mathbf{F}) \quad (3.52)$$

It is straightforward to show that the space component of Eq. (3.52) reproduces Newton's second law

$$m \frac{d\mathbf{v}}{dt} = \mathbf{F} \quad (3.53)$$

To derive the time component of the four-vector f^μ , we differentiate a Lorentz scalar $u_\mu u^\mu = 1$ and use Eq. (3.52) to obtain

$$\begin{aligned} m \left(u_\mu \frac{du^\mu}{d\tau} \right) &= m \left(u_0 \frac{du^0}{d\tau} - \mathbf{u} \cdot \frac{d\mathbf{u}}{d\tau} \right) = u^0 f^0 - \mathbf{u} \cdot \mathbf{f} = 0 \\ \therefore f^0 &= \frac{\mathbf{u} \cdot \mathbf{f}}{u_0} = \gamma \frac{\mathbf{v} \cdot \mathbf{F}}{c} \xrightarrow{\gamma \rightarrow 0} \frac{\mathbf{v} \cdot \mathbf{F}}{c} \end{aligned} \quad (3.54)$$

Then the time component of Eq. (3.52) becomes

$$m \frac{du^0}{d\tau} = \gamma \frac{d(m\gamma c)}{dt} = f^0 = \frac{\gamma \mathbf{v} \cdot \mathbf{F}}{c}, \Rightarrow \frac{d}{dt} \left(\frac{mc^2}{\sqrt{1 - (v/c)^2}} \right) = \mathbf{v} \cdot \mathbf{F} \quad (3.55)$$

The right hand side of the equation is the rate of energy increase, and the above relation led Einstein to derive the famous relation $E = mc^2$.

Example 3.2: Equation of Spin Precession*

We derive here the spin precession frequency of a particle moving in a homogeneous electromagnetic field. Although the quantum mechanical motion should be derived using the Dirac equation (see Chap. 4) with inclusion of a Pauli term, the expectation value of the spin operator is known to follow the same time dependence as that would be obtained from a classical equation of motion [47, 216]. The spin operator is an antisymmetric tensor $S^{\mu\nu}$ (see Sect. 3.5), but a formulation in terms of a covariant (axial) four-vector would be much more convenient and we will try to construct it. We define the spin four-vector S^μ which reduces to the conventional spin $S^\mu \stackrel{R}{=} (0, \mathbf{s})$ in the particle's rest frame (R).⁴⁾ It obeys the equation of motion

$$\frac{ds}{d\tau} \stackrel{R}{=} g \frac{q}{2m} \mathbf{s} \times \mathbf{B}, \quad q = \pm e \quad (3.56)$$

where $\mu = gq\hbar/2m$ ($\hbar$ and c omitted in natural units) is the magnetic moment of the particle and also defines the g -factor. Denote the particle's four-velocity by

4) Actually, S^μ is the dual of the tensor $S^{\mu\nu}$ in the sense that $S^\mu = (1/2c)\epsilon^{\mu\nu\rho\sigma} u_\nu S_{\rho\sigma}$, where u^μ is the particle's four velocity.

5) We assume the equation is linear in the spin S^μ and the external field $F^{\mu\nu}$.

$u^\mu = (u^0, \mathbf{u}) = \gamma(1, \mathbf{v})$, which obeys the equation of motion

$$m \frac{du^\mu}{d\tau} = q F^{\mu\nu} u_\nu = q(\mathbf{u} \cdot \mathbf{E}, u^0 \mathbf{E} + \mathbf{u} \times \mathbf{B}) \quad (3.57)$$

where we used Eq. (3.52) and the Lorentz force $q F^{\mu\nu} u_\nu$ supplied by the electromagnetic field. To obtain the zeroth component S^0 of the spin vector, we consider a Lorentz scalar $\mathbf{u} \cdot \mathbf{S}$ that vanishes in the rest frame and hence in any frame. Then

$$\mathbf{S} \cdot \mathbf{u} = S^0 u^0 - \mathbf{u} \cdot \mathbf{S} = 0 \rightarrow S^0 = \mathbf{v} \cdot \mathbf{S} \quad (3.58)$$

The equation of motion for S^0 in the rest frame is readily obtained

$$\frac{dS^0}{d\tau} = \frac{d\mathbf{v}}{d\tau} \cdot \mathbf{S} + \mathbf{v} \cdot \frac{d\mathbf{S}}{d\tau} \stackrel{R}{=} \frac{q}{m} \mathbf{s} \cdot \mathbf{E} \quad (3.59)$$

Eq. (3.56) and Eq. (3.59) may be combined and generalized to a Lorentz covariant form:⁹⁾

$$\begin{aligned} \frac{dS^\mu}{d\tau} &= g \frac{q}{2m} [F^{\mu\nu} S_\nu + (S_\mu F^{\mu\nu} u_\nu) u^\mu] - \frac{q}{m} \left(\frac{d\mathbf{u}}{d\tau} \cdot \mathbf{S} \right) u^\mu \\ &= g \frac{q}{2m} F^{\mu\nu} S_\nu + \left(\frac{g-2}{2} \right) \frac{q}{m} (S_\mu F^{\mu\nu} u_\nu) u^\mu \end{aligned} \quad (3.60a)$$

$$F^{\mu\nu} S_\nu = (\mathbf{S} \cdot \mathbf{E}, S^0 \mathbf{E} + \mathbf{S} \times \mathbf{B}) \stackrel{R}{=} (\mathbf{s} \cdot \mathbf{E}, \mathbf{s} \times \mathbf{B}) \quad (3.60b)$$

$$S_\mu F^{\mu\nu} u_\nu = S^0(\mathbf{u} \cdot \mathbf{E}) - u^0(\mathbf{S} \cdot \mathbf{E}) - (\mathbf{S} \cdot \mathbf{u} \times \mathbf{B}) \stackrel{R}{=} -\mathbf{s} \cdot \mathbf{E} \quad (3.60c)$$

where (3.57) is used in obtaining the second line. Equation (3.60a) is the relativistic equation of motion of the spin vector. It can easily be shown that Eq. (3.60a) reduces to Eq. (3.56) and Eq. (3.59) in the rest frame if one makes use of Eq. (3.60b) and (3.60c). For experimental use, we are interested in the spin precession frequency. The formula Eq. (3.60a) is used to determine a precise value of the $g - 2$ factor of the muon, which will be treated in more detail in Sect. 8.2.2.

3.4.4

Rapidity

Successive Lorentz boost in the same direction is represented by a single boost where the transformation velocity is given by

$$\beta'' = |\mathbf{v}/c|'' = \frac{\beta + \beta'}{1 + \beta\beta'} \quad (3.61)$$

Proof: Assume velocity $\mathbf{v}'$ in frame L is observed as $\mathbf{v}''$ in frame L' , where the frame L' is traveling in the x direction with $-v$ in frame L . Space coordinates $(t', x^{1'})$ are expressed in terms of (t, x^1) using transformation of Eq. (3.19). We

omit other coordinates for simplicity:

$$x^{0'} = \gamma(x^0 + \beta x^1) \quad (3.62a)$$

$$x^{1'} = \gamma(\beta x^0 + x^1) \quad (3.62b)$$

$$\beta' = \frac{v'}{c} = \frac{dx^1}{dx^0} \quad (3.62c)$$

Then

$$\beta'' = \frac{dx^{1'}}{dx^{0'}} = \frac{\gamma(\beta dx^0 + dx^1)}{\gamma(dx^0 + \beta dx^1)} = \frac{\beta + \beta'}{1 + \beta\beta'} \quad (3.63)$$

□

Equation (3.63) shows the velocity is not an additive parameter, i.e. nonlinear, in the successive transformation. A more convenient additive parameter “rapidity” can be defined by

$$\beta = \tanh \eta, \quad \text{or} \quad \eta = \frac{1}{2} \ln \frac{1 + \beta}{1 - \beta} \quad (3.64)$$

Problem 3.6

Show that the rapidity is additive, i.e. $\eta'' = \eta + \eta'$.

Using the rapidity, a Lorentz transformation with finite η , can be decomposed into N successive transformations with rapidity $\Delta\eta = \eta/N$, which we discuss more in the following section. Solving β, γ in terms of η we have

$$\beta = \tanh \eta, \quad \gamma = \cosh \eta, \quad \beta\gamma = \sinh \eta \quad (3.65)$$

Lorentz boost (3.19) can be rewritten as

$$x^{0'} = \cosh \eta x^0 + \sinh \eta x^1 \quad (3.66a)$$

$$x^{1'} = \sinh \eta x^0 + \cosh \eta x^1 \quad (3.66b)$$

Compare this with rotation in the x - y plane.

$$x' = x \cos \theta - y \sin \theta \quad (3.67a)$$

$$y' = x \sin \theta + y \cos \theta \quad (3.67b)$$

As Eq. (3.66) can be obtained from Eq. (3.67) by the substitution $\theta \rightarrow -i\eta$, $x \rightarrow ix^0$, $y \rightarrow x^1$, Lorentz boost (in the x direction) is formally a rotation by angle $-i\eta$ in the x and imaginary time (ix^0) plane.

Comment 3.2 In high-energy collider experiments, produced secondary particles are boosted in the z direction (along the incoming particle axis). The boosted angular distribution is better expressed as rapidity distribution. $dN/d\eta \sim \text{const.}$ is observed experimentally. At high energies, each particle has $E \sim pc$, $p_{\parallel} = p \cos \theta$, and its rapidity is approximated by so-called pseudo-rapidity

$$\eta' = \frac{1}{2} \ln \frac{1 + \beta_{\parallel}}{1 - \beta_{\parallel}} = \frac{1}{2} \ln \frac{E + p_{\parallel}c}{E - p_{\parallel}c} \sim -\ln \tan \frac{\theta}{2} \quad (3.68)$$

This fact is taken into account in designing detectors, which are divided into modules that span the same solid angle in the η - ϕ (azimuthal angle) plane.

3.5

Lorentz Operator

In order to obtain a Lorentz transformation operator that will act on a field ψ , where ψ can have multiple components, the Lorentz operation of the field can be defined similarly to rotation [see Eq. (3.4)]

$$\psi(x) \xrightarrow{L} \psi'(x) = D(L)\psi(L^{-1}x) \quad (3.69)$$

where Lx means for small boost

$$x^{0'} = x^0 + \Delta\eta x^1, \quad x^{1'} = x^1 + \Delta\eta x^0 \quad (3.70)$$

where we have considered boost only in the x direction for simplicity.

First we consider Lorentz transformation operations of the scalar field. Here only the coordinate variables are affected:

$$\begin{aligned} \psi'(x) &= \psi(L^{-1}x) = \psi(x^0 - \Delta\eta x^1, x^1 - \Delta\eta x^0) \\ &= \psi(x^0, x^1) - \Delta\eta(x^1\partial_0\psi + x^0\partial_1\psi) \equiv (1 - i\Delta\eta N_1)\psi(x) \end{aligned} \quad (3.71)$$

As a transformation with finite η can be achieved by N successive infinitesimal transformations with $\Delta\eta = \eta/N$:

$$\psi'(x) = \lim_{N \rightarrow \infty} \{1 - i(\eta/N)N_1\}^N \psi = \exp(-i\eta N_1)\psi \quad (3.72)$$

N_1 is called a generator of Lorentz boost and is defined by the formula

$$N_1 = -i(x^0\partial_1 + x^1\partial_0) = i(x^0\partial^1 - x^1\partial^0) \quad (3.73a)$$

Similarly, considering boosts in the x^2 and x^3 directions

$$N_2 = i(x^0\partial^2 - x^2\partial^0) \quad (3.73b)$$

$$N_3 = i(x^0\partial^3 - x^3\partial^0) \quad (3.73c)$$

As is well known, orbital angular momentum operators, which are generators of space rotation, are given as

$$L_i = -i(x^j \partial_k - x^k \partial_j) = i(x^j \partial^k - x^k \partial^j), \quad (i, j, k) = (1, 2, 3) \text{ and cyclic} \quad (3.73d)$$

It is easily verified that N_i and L_j satisfy the following commutation relations:

$$[L_i, L_j] = i\epsilon_{ijk} L_k \quad (3.74a)$$

$$[N_i, N_j] = -i\epsilon_{ijk} L_k \quad (3.74b)$$

$$[N_i, L_j] = i\epsilon_{ijk} N_k \quad (3.74c)$$

$$\epsilon_{ijk} = \begin{cases} +1 & ijk = \text{even permutation of } 123 \\ -1 & = \text{odd permutation of } 123 \end{cases} \quad (3.74d)$$

Now define Lorentz generators $M^{\mu\nu}$ and Lorentz variables $\omega_{\mu\nu}$ by

$$M^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu) \quad (3.75a)$$

$$M^{\mu\nu} = -M^{\nu\mu}, \quad \omega_{\mu\nu} = -\omega_{\nu\mu} \quad (3.75b)$$

$$(M^{23}, M^{31}, M^{12}) = (L_1, L_2, L_3) \quad (3.75c)$$

$$(M^{01}, M^{02}, M^{03}) = (N_1, N_2, N_3) \quad (3.75d)$$

$$(\omega_{23}, \omega_{31}, \omega_{12}) = (\theta_1, \theta_2, \theta_3) \quad (3.75e)$$

$$(\omega_{01}, \omega_{02}, \omega_{03}) = (\eta_1, \eta_2, \eta_3) \quad (3.75f)$$

Then the general Lorentz transformation operators that act on coordinate variables can be written as

$$L \equiv \exp[-i(N \cdot \boldsymbol{\eta}) - i(L \cdot \boldsymbol{\theta})] = \exp\left(-\frac{i}{2}\omega_{\mu\nu} M^{\mu\nu}\right) \quad (3.76)$$

Next we consider Lorentz transformations on multi-component fields and include spin operators. Then a general Lorentz boost operator is expressed as

$$M^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu) + S^{\mu\nu} \quad (3.77a)$$

$$(S^{23}, S^{31}, S^{12}) = (S_1, S_2, S_3) \quad (3.77b)$$

$$(S^{01}, S^{02}, S^{03}) = (K_1, K_2, K_3) \quad (3.77c)$$

where $\mathbf{S}$ and $\mathbf{K}$ satisfy the same commutation relations as $\mathbf{L}$ and $\mathbf{N}$ in Eq. (3.74). In the following we consider a Lorentz transformation of the vector fields.⁶⁾ It can be expressed as

$$V^\mu(x) \xrightarrow{L} V^{\mu'}(x) = L^\mu{}_\nu V^\nu(L^{-1}x) \quad (3.78)$$

Problem 3.7

(a) By inspecting Eq. (3.66) and Eq. (3.67) and generalizing to other directions, show that infinitesimal proper Lorentz transformation with $\boldsymbol{\theta} = \delta\boldsymbol{\theta}$, $\boldsymbol{\eta} = \delta\boldsymbol{\eta}$ acting on vector components is given by

$$L^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon^\mu{}_\nu \quad (3.79a)$$

$$\epsilon^\mu{}_\nu = \begin{bmatrix} 0 & \delta\eta_1 & \delta\eta_2 & \delta\eta_3 \\ \delta\eta_1 & 0 & -\delta\theta_3 & \delta\theta_2 \\ \delta\eta_2 & \delta\theta_3 & 0 & -\delta\theta_1 \\ \delta\eta_3 & -\delta\theta_2 & \delta\theta_1 & 0 \end{bmatrix} \quad (3.79b)$$

(b) Show also that

$$\epsilon_{\mu\nu} \equiv g_{\mu\rho}\epsilon^\rho{}_\nu = -\epsilon_{\nu\mu} \quad (3.79c)$$

(c) Show that Eq. (3.32) alone constrains $\epsilon_{\mu\nu}$ to be antisymmetric.

As the Lorentz vector transforms in the same way as the space-time coordinates, we can derive the spin part of the Lorentz generator $S^{\mu\nu}$ by using Eq. (3.66) or (3.79). Let us derive K_1 , the Lorentz boost generator in the x direction. Writing

$$L^\mu{}_\nu = [1 - i\delta\boldsymbol{\eta} \cdot \mathbf{K} - i\delta\boldsymbol{\theta} \cdot \mathbf{S}]^\mu{}_\nu \quad (3.80)$$

we can derive K_1 by

$$[K_1]^\mu{}_\nu = i \frac{\partial}{\partial\eta_1} L^\mu{}_\nu \Big|_{\eta_1=0} = \begin{bmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3.81)$$

Other generators are obtained similarly.

Problem 3.8

Show

$$K_1 = \begin{bmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad K_2 = \begin{bmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad K_3 = \begin{bmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix} \quad (3.82a)$$

6) For the treatment of particles with half-integer spin, see Appendix A.

$$S_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{bmatrix}, \quad S_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3.82b)$$

Show also that K_i and S_i satisfy the same commutation relations Eq. (3.74) as N_i and L_i .

3.6

Poincaré Group*

So far we have considered only homogeneous Lorentz transformation. However, the reader should be aware that the Minkowski metric Eq. (3.16) is invariant under more general Lorentz transformations known as the Poincaré transformation:

$$x'^{\mu} = L^{\mu}_{\nu} x^{\nu} + a^{\mu}, \quad \mathbf{x}' = \mathbf{L}\mathbf{x} + \mathbf{a}, \quad \mathbf{L} = [L^{\mu}_{\nu}] = \frac{\partial x'^{\mu}}{\partial x^{\nu}} \quad (3.83)$$

If we write the transformation as $T(L, a)$, it satisfies

$$T(L', a')T(L, a) = T(L'L, L'a + a') \quad (3.84)$$

and makes a Poincaré (or inhomogeneous Lorentz) group. If we consider an infinitesimal transformation, the transformation matrix can be written as

$$L^{\mu}_{\nu} = \delta^{\mu}_{\nu} + \omega^{\mu}_{\nu}, \quad a^{\mu} = \varepsilon^{\mu} \quad (3.85)$$

The infinitesimal translation can be expressed as

$$\begin{aligned} f'(x^{\mu} + \varepsilon^{\mu}) &= f(x^{\mu} - \varepsilon^{\mu}) = f(x^{\mu}) - \varepsilon^{\mu} \partial_{\mu} f(x^{\mu}) = [1 + i\varepsilon_{\mu}(i\partial^{\mu})]f(x^{\mu}) \\ &= (1 + i\varepsilon_{\mu}P^{\mu})f(x^{\mu}) \\ P^{\mu} &= i\partial^{\mu} = i(\partial_0, -\nabla) \end{aligned} \quad (3.86)$$

The homogeneous Lorentz transformation is already given by Eq. (3.75). Therefore the inhomogeneous Lorentz transformation operator can be written as

$$f(L^{-1}x - a) = \left(1 - i\frac{1}{2}\omega_{\mu\nu}M^{\mu\nu} + i\varepsilon_{\mu}P^{\mu}\right)f(x) \quad (3.87)$$

The finite transformation is obtained as usual by successive operation and is given by

$$U(L, a) = \exp\left[-\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} + ia_{\mu}P^{\mu}\right] \quad (3.88)$$

Fields are transformed by the Poincaré group operation as

$$\psi'(x) = D(L, a)\psi(L^{-1}x - a) = U(L, a)\psi(x) \quad (3.89)$$

where $M^{\mu\nu}$ now includes the spin operator. Definitions Eq. (3.86) and Eq. (3.75) show that the operators P^μ and $M^{\mu\nu}$ satisfy the relations

$$[P^\mu, P^\nu] = 0 \quad (3.90a)$$

$$[M^{\mu\nu}, P^\rho] = i(g^{\nu\rho}P^\mu - g^{\mu\rho}P^\nu) \quad (3.90b)$$

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(g^{\nu\rho}M^{\mu\sigma} - g^{\mu\rho}M^{\nu\sigma} - g^{\nu\sigma}M^{\mu\rho} + g^{\mu\sigma}M^{\nu\rho}) \quad (3.90c)$$

Note, the last commutation relation is the same as Eq. (3.74), but re-expressed in terms of $M^{\mu\nu}$ instead of L_i and K_i . They constitute a closed Lie algebra, as they should to make a group, and thus are generators of the Poincaré group. From Eq. (3.75) and Eq. (3.90), we see that the momentum as well as the angular momentum operators, $\mathbf{J} = (M^{23}, M^{31}, M^{12})$, commute with the Hamiltonian ($H = P^0$) and hence are constants of motion, but the Lorentz boost operators $\{\mathbf{K} = (M^{01}, M^{02}, M^{03})\}$ are not. We can make two Casimir operators that commute with all the generators:

$$m^2 = P_\mu P^\mu \quad (3.91a)$$

$$W_\mu W^\mu = -J(J+1)m^2, \quad W_\mu = -\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}M^{\nu\rho}P^\sigma \quad (3.91b)$$

W_μ is called the Pauli–Lubanski operator. Equation (3.91) can be used to define the mass and spin of a particle. It can be verified that

$$[P^\mu, W^\nu] = 0 \quad (3.92)$$

Therefore we can diagonalize the mutually commutable operators simultaneously,

$$P^\mu, W_3, P^2, W^2 \quad (3.93)$$

and use their eigenvalues to specify a one particle state. Particle states may be classified by the value of m^2 and P^μ as follows:

1. $P^2 = m^2 > 0$ ordinary massive particle
 2. $P^2 = m^2 = 0$ photon, graviton, etc.
 3. $P^2 < 0$ tachyon or virtual particle
 4. $P^\mu = 0$ vacuum
- (3.94)

Only particles in categories 1 and 2 are physical particles and we consider them only. For particles with $m^2 > 0$, we can also specify their spin S . In their rest frame, $P^\mu = (m, 0, 0, 0)$, W_i becomes

$$W_i = -\frac{1}{2}\varepsilon_{i\nu\rho\sigma}J^{\nu\rho}P^\sigma = -\frac{m}{2}\varepsilon_{ijk0}J^{jk} = mS_i \quad (3.95)$$

and the state can also be specified by its spin components. However, for particles with $m^2 = 0$ including photons and gravitons, which fly always at the speed of light, the notion of rotation in the rest frame cannot be applied. As

$$P^2 = W^2 = P \cdot W = 0 \quad (3.96)$$

P^μ and W^μ are orthogonal and both are on the light cone. So if W^μ is a real vector, it means they are proportional:

$$W^\mu = \lambda P^\mu \quad (3.97)$$

Therefore, the particle state with $m^2 = 0$ is characterized by one number, which is proportional to W^μ/P^μ and has the dimension of angular momentum. It is called helicity ($\lambda = J_3$ along the momentum direction). With the parity operation (if conserved), the helicity can be flipped. It can be shown that $J_3 = \pm|J|$. The photon has $J = 1$ and its interaction preserves parity. Therefore the photon can have $J_3 = \pm 1$, but a state with $J_3 = 0$ is not allowed, which means the photon is a quantum of a transverse wave. In summary, a massless particle can have only two states, specified by its helicity. For the case of a spin 1/2 massless particle, which is known as the Weyl particle, we shall give an explicit expression of the state in Sect. 4.1.2.

In summary, in the framework of special relativity, a particle state can be specified by its energy-momentum, its spin value s , and spin components s_z . If the particle is massive, s_z has $2S + 1$ components but if it is massless $s_z = \pm S$ has only two components.

4

Dirac Equation

4.1

Relativistic Schrödinger Equation

4.1.1

Dirac Matrix

We have seen in Chap. 2, that the relativistic wave equation encounters problems, such as negative probability and negative energy, if we try to interpret the wave function as a probability amplitude. The difficulties can be avoided by interpreting it as an equation for a (classical) field. Historically, a relativistic version of the Schrödinger equation without negative probability was first obtained by Dirac. It also turned out to be the basic wave equation for the spin 1/2 particle. Since quarks and leptons, the fundamental particles of matter, have spin 1/2, we have to understand the structure of the spin 1/2 field before we quantize the Dirac fields. For the moment, we follow Dirac's argument to derive a relativistic Schrödinger equation. His heuristic argument helps in understanding the physical meaning of otherwise mathematical equations. Dirac's logic was as follows.

1. The negative probability originates in the fact that the Klein–Gordon equation contains second time derivatives. Therefore, the equation must be a linear function of the first time derivative. Relativistic covariance requires that space derivatives have to be linear, too.
2. The wave function ψ should have more than one component. Because the equation of motion generally includes second derivatives and they can only be included by introducing coupled multicomponent equations.¹⁾ If we define the

1) This can be understood, for instance, by considering a simple harmonic oscillator equation $\partial^2 q / \partial t^2 = -\omega^2 q$ which includes a second derivative. By introducing $q_1 = q$, $q_2 = -i(\partial q / \partial t) / \omega$, it can be converted to two simultaneous equations, $\partial q_1 / \partial t = i\omega q_2$, $\partial q_2 / \partial t = i\omega q_1$. In matrix form, they can be

combined to a single linear equation.

$$\frac{d\psi}{dt} = H\psi, \quad \psi = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & i\omega \\ i\omega & 0 \end{bmatrix} \quad (4.1)$$

probability as

$$\rho = \sum_r \psi_r^* \psi_r = \psi^\dagger \psi, \quad \psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{bmatrix} \quad (4.2)$$

ρ stays positive definite.

3. The Klein–Gordon equation (i.e. the Einstein relation $E^2 = \mathbf{p}^2 + m^2$) has to be satisfied.

Applying (1) and (2), the equation must be of the form

$$i \frac{\partial \psi}{\partial t} = H \psi = \left(-i \sum_i \alpha_i \frac{\partial}{\partial x^i} + \beta m \right) \psi = (\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m) \psi \quad (4.3)$$

Here, α_i and β are matrices. $\mathbf{p}$ is a differential operator but becomes a c-number when we apply it to a plane wave $\psi(x) = A e^{-i p x} = A e^{-i E t + \mathbf{p} \cdot \mathbf{x}}$. Since any wave can always be decomposed into plane waves, we shall use $(E, \mathbf{p})$ interchangeably between differential operators and c-numbers unless there is confusion between the two:

$$p_\mu = (E, -\mathbf{p}) \Leftrightarrow i \partial_\mu, \quad p^\mu = (E, \mathbf{p}) \Leftrightarrow i \partial^\mu = \left(i \frac{\partial}{\partial t}, -i \nabla \right) \quad (4.4)$$

The Schrödinger equation (4.3) can be expressed in the form

$$E \psi = (\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m) \psi \quad (4.5)$$

Statement (3) requires

$$E^2 = \mathbf{p}^2 + m^2 = \sum_{i,j=1}^3 (\alpha_i \alpha_j + \alpha_j \alpha_i) p^i p^j + m \sum_{i=1}^3 (\beta \alpha_i + \alpha_i \beta) p^i + \beta^2 m^2 \quad (4.6a)$$

$$\therefore \quad \alpha_i^2 = \beta^2 = 1, \quad \alpha_i \alpha_j + \alpha_j \alpha_i = 0 \quad (i \neq j), \quad \beta \alpha_i + \alpha_i \beta = 0 \quad (4.6b)$$

Matrices that satisfy Eq. (4.6b) are called Dirac matrices. Their dimension has to be an even number.

Proof:

(a) Eigenvalues of α_i, β are ± 1 .

$$\therefore \quad \alpha_i \phi = \lambda \phi \rightarrow \alpha_i^2 \phi = \phi = \lambda^2 \phi, \quad \therefore \quad \lambda = \pm 1 \quad (4.7)$$

(b) α_i, β are traceless.

$$\therefore \quad \text{Tr}[\alpha_i] = \text{Tr}[\alpha_i \beta^2] = \text{Tr}[\beta \alpha_i \beta] = -\text{Tr}[\alpha_i \beta^2] = -\text{Tr}[\alpha_i] \quad (4.8)$$

The third equality follows because a cyclic order change does not affect the trace value and the fourth equality uses the anticommutativity of α_i and β . (a) and (b) together restrict the dimension to be an even number. $\square$

4.1.2

Weyl Spinor

If $m = 0$, $\alpha_i = \pm \sigma_i$ satisfy the Dirac condition (4.6b), where σ_i are Pauli matrices:

$$i \frac{\partial \psi}{\partial t} = \pm \boldsymbol{\sigma} \cdot (-i \nabla) \psi \quad (4.9)$$

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (4.10)$$

This particular form is called the Weyl equation. We shall consider a plane-wave solution of the type $\psi(\mathbf{x}) \sim u(\mathbf{p}) e^{-i p x}$, where $u(\mathbf{p})$ is a two-component wave function. We distinguish two solutions ϕ_L and ϕ_R , corresponding to different signs of the Weyl equation (4.9). They satisfy

$$E \phi_L = -\boldsymbol{\sigma} \cdot \mathbf{p} \phi_L, \quad E \phi_R = +\boldsymbol{\sigma} \cdot \mathbf{p} \phi_R \quad (4.11a)$$

$$\phi_L = -\frac{|\mathbf{p}|}{E} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{|\mathbf{p}|} \phi_L, \quad \phi_R = \frac{|\mathbf{p}|}{E} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{|\mathbf{p}|} \phi_R \quad (4.11b)$$

Here,

$$\mathbf{h} \equiv \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{|\mathbf{p}|} = \boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \quad (4.12)$$

is a spin projection operator in the direction of motion and is called a helicity operator. $\hat{\mathbf{p}}$ is a unit vector in the direction of $\mathbf{p}$. Classically, the helicity eigenstate with $h = (+)$ represents a moving state rotating clockwise²⁾ and the particle is said to be right handed. Similarly, the helicity $(-)$ state is left handed.

Problem 4.1

Let $\hat{\mathbf{p}}$ be in the direction (θ, ϕ) . Then the helicity operator is expressed as

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{p}} = \begin{bmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{bmatrix} \quad (4.13)$$

Show that eigenstates $\chi_{\pm}$ of the helicity operator are expressed as

$$\chi_+ = \begin{bmatrix} \cos \frac{\theta}{2} e^{-i\phi/2} \\ \sin \frac{\theta}{2} e^{i\phi/2} \end{bmatrix}, \quad \chi_- = \begin{bmatrix} -\sin \frac{\theta}{2} e^{-i\phi/2} \\ \cos \frac{\theta}{2} e^{i\phi/2} \end{bmatrix} \quad (4.14)$$

As $m = 0$, $|\mathbf{p}|/E = \pm 1$. Choosing χ_- for ϕ_L gives

$$\phi_L = \chi_- \quad E > 0, \quad \phi_L = \chi_+ \quad E < 0 \quad (4.15)$$

- 2) The orbital angular momentum is defined as $\mathbf{l} = \mathbf{r} \times \mathbf{p}$ and $l_z > 0$ means clockwise rotation viewed from upstream or in the $-z$ direction. We consider the spin operator to have the same attributes.

In other words, the solution ϕ_L to the Weyl equation is a pure helicity state of a particle but flipped for its antiparticle ($E < 0$). The exact definition of an antiparticle will be given in the next section. We shall use it to denote the negative energy particle. ϕ_L does not contain a χ_+ component for the particle and no χ_- components for the antiparticle. We have now justified the use of L/R for the subscript. We call it the chirality ($-$) state, an extended helicity concept applicable to the antiparticle. It will be discussed in more detail in Sect. 4.3.5. ϕ_R is defined similarly. ϕ_L and ϕ_R are solutions of different equations and hence are particles independent of each other. They are called Weyl particles.

However, the two equations (4.11a) and hence ϕ_L and ϕ_R interchange by a parity transformation:

$$\mathbf{x} \leftrightarrow -\mathbf{x}, \quad \mathbf{p} \leftrightarrow -\mathbf{p} \quad (4.16)$$

Therefore the solutions cannot be applied to particles that have strong or electromagnetic interactions, as they are known to conserve parity. The Weyl solution could describe the neutrino, which interacts only weakly apart from the universal gravity.³⁾ To describe particles in the real world, we have to combine the two solutions. Let us extend ψ from a two-components to 4-component wave function and define

$$\psi = \begin{bmatrix} \phi_L \\ \phi_R \end{bmatrix}, \quad \alpha \equiv \begin{bmatrix} -\sigma & 0 \\ 0 & \sigma \end{bmatrix}, \quad \beta \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (4.17)$$

β is a matrix to exchange ϕ_R and ϕ_L . The above α and β satisfy the Dirac condition (4.6b) and so the equation using α and β defined above satisfies all the requirements for describing relativistic particles having finite mass. It is called the Dirac equation. Note that the way to determine α and β is not unique. Different α' and β' defined by transformation matrix T

$$\alpha' = T\alpha T^{-1}, \quad \beta' = T\beta T^{-1} \quad (4.18)$$

satisfy the same Dirac conditions. The matrix form in Eq. (4.17) is called the Weyl representation. Another representation commonly used in many textbooks is the Pauli–Dirac representation, which is given by

$$\begin{aligned} \psi_D &= T\psi_{\text{Weyl}}, \quad T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ \alpha_D &= \begin{bmatrix} 0 & \sigma \\ \sigma & 0 \end{bmatrix}, \quad \beta_D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{aligned} \quad (4.19)$$

The physics content does not depend on which representation is used, but it is easier to proceed with an explicit representation. In this book we use the Weyl representation unless otherwise noted because it diagonalizes the energy-momentum

3) However, the phenomenon of neutrino oscillation, discovered in 1998 [357], is evidence for the finite mass of the neutrino, hence the possibility of the neutrino being a Weyl particle is denied.

term and is more suitable for describing relativistic ($E, p \gg m$) particles. The Pauli–Dirac representation diagonalizes the mass term and is convenient for description in the nonrelativistic energy region. Let us proceed. The Dirac equation is given by

$$i \frac{\partial \psi}{\partial t} = (-\boldsymbol{\alpha} \cdot i \nabla + \beta m) \psi \Leftrightarrow E \psi = [\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m] \psi \quad (4.20)$$

Now we try to construct a probability current by assigning

$$j^\mu(x) = (\rho(x), \mathbf{j}(x)) \quad (4.21a)$$

$$\rho(x) = \psi^\dagger(x) \psi(x), \quad \mathbf{j}(x) = \psi^\dagger(x) \boldsymbol{\alpha} \psi(x) \quad (4.21b)$$

Using the Dirac equation (4.20)

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \frac{\partial \psi^\dagger}{\partial t} \psi + \psi^\dagger \frac{\partial \psi}{\partial t} = -i \{ -(H \psi)^\dagger \psi + \psi^\dagger H \psi \} \\ &= -i \{ -\{ (-i \boldsymbol{\alpha} \cdot \nabla + \beta m) \psi \}^\dagger \psi + \psi^\dagger \{ (-i \boldsymbol{\alpha} \cdot \nabla + \beta m) \psi \} \} \\ &= -\nabla \cdot (\psi^\dagger \boldsymbol{\alpha} \psi) = -\nabla \cdot \mathbf{j} \end{aligned} \quad (4.22)$$

so the current defined above respects the equation of continuity. Since ρ is always positive definite, the interpretation of ψ as a probability amplitude wave function is possible. Dirac seems to have succeeded in constructing a relativistic version of the Schrödinger equation.

Problem 4.2

Consider Eq. (4.20) as expressed in the Heisenberg picture. Using the Heisenberg equation of motion

$$\frac{dO}{dt} = i[H, O] \quad (4.23)$$

and setting $O = \mathbf{x}$, show that $d\mathbf{x}/dt = \boldsymbol{\alpha}$.

Therefore $\boldsymbol{\alpha}$ can be considered as a velocity operator. This also means $j^\mu = (\rho, \mathbf{j})$ can be interpreted as a classical four-component current $(\rho, \rho \mathbf{v})$.

Problem 4.3

Prove that $H = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m$ does not commute with $\mathbf{L}$ but does with $\mathbf{J}$, where

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} = \mathbf{x} \times (-i \nabla), \quad \mathbf{J} = \mathbf{L} + \mathbf{S} = \mathbf{L} + \frac{1}{2} \boldsymbol{\Sigma}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{bmatrix} \quad (4.24)$$

$\mathbf{S}$ satisfies the commutation relation of the angular momentum. It can be considered as a four-component version of the spin operator with $s = 1/2$.

Problem 4.4

Show that the four-component helicity operator $h = \boldsymbol{\Sigma} \cdot \hat{\mathbf{p}}$ commutes with the Dirac Hamiltonian.

Therefore, the helicity eigenstate is also an energy eigenstate of the Hamiltonian, and its quantum number can be used to specify the quantum state of the spin 1/2 particle.

4.1.3

Interpretation of the Negative Energy**Spin of the Dirac Particle**

There still remains the negative energy problem, but first we will try to understand what kind of freedom the wave function possesses. In the limit of $\mathbf{p} \rightarrow 0$, the Dirac equation, Eq. (4.20), becomes

$$E\psi = m\beta\psi = m \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \psi \quad (4.25)$$

Solutions are easy to find:

$$\psi_{+,r} = \begin{bmatrix} \xi_r \\ \xi_r \end{bmatrix}, \quad \psi_{-,r} = \begin{bmatrix} \xi_r \\ -\xi_r \end{bmatrix} \quad (4.26)$$

where ξ_r is a two-component spin wave function

$$\xi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \xi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (4.27)$$

$\psi_{\pm,r}$ are eigenstates of the Hamiltonian as well as the four-component spin operator $\mathbf{S}$ defined in Eq. (4.24). Namely,

$$E\psi_{+,r} = m\psi_{+,r} \quad E\psi_{-,r} = -m\psi_{-,r} \quad (4.28a)$$

$$S_3\psi_{\pm,+} = \frac{1}{2}\psi_{\pm,+} \quad S_3\psi_{\pm,-} = -\frac{1}{2}\psi_{\pm,-} \quad (4.28b)$$

We now have proved that the Dirac equation describes a particle of spin 1/2, but also has a freedom corresponding to the sign of the energy. The negative energy state seems unavoidable.

Consequences of the Dirac Equation

Dirac applied his equation to the problem of energy levels in the hydrogen atom. By solving the Dirac equation including a static Coulomb potential, he obtained the energy spectrum given in Fig. 4.1 and explained the fine and hyperfine structure of the energy levels, which were all degenerate according to the nonrelativistic Schrödinger equation.⁴⁾ He also investigated the property of the electron in a mag-

4) Actually Dirac did not calculate the energy levels of the hydrogen when he published the paper on his equation. Later, in his Oppenheimer memorial talk in 1969, he confessed that he did not do it because he was afraid his equation might be disproved by detailed calculations. The author is grateful to Professor T. Kubota for pointing out this historically interesting anecdote.

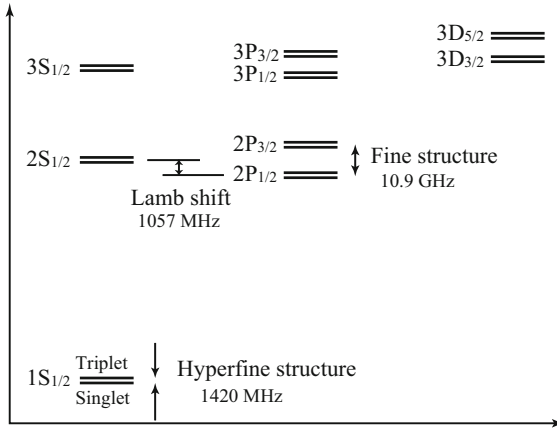


Figure 4.1 Low-lying energy spectrum of the hydrogen atom. Degenerate states with the principal quantum number n are separated by spin-orbit coupling (fine structure) and spin-spin coupling (hyperfine structure). The Lamb shift is due to a higher order QED effect. The diagram is not drawn

to scale. Note that the hyperfine splitting of $1S^{1/2}$ corresponds to an energy splitting of $5.89 \times 10^{-6} \text{ eV} = 1420 \text{ MHz} = 21 \text{ cm radio wave}$, which is extensively used in radio astronomy to probe the hydrogen abundance in the universe.

netic field and succeeded in explaining the $g = 2$ value of the moment, which is described in Sect. 4.3. By his great success, he must have convinced himself of the correctness of his equation. Then the negative energy puzzle must correspond to some physical reality and its apparent contradictory property should have some explanation. The contradiction stems from the fact that if a particle is in a negative energy state, it can lose further energy by emitting photons and falls down without limit and hence cannot be stable.

Dirac considered that the vacuum must be a state where all the negative energy levels are occupied. Then, because of the Pauli exclusion principle, electrons cannot fall down to negative energy levels and thus become stable. However, the concept opened a new possibility. When an energy greater than twice the electron mass is given to the vacuum, an electron in the negative-energy sea is excited to a positive energy level and a hole is created in the sea (Fig. 4.2a). When an electric field is applied, a negative energy electron next to the hole moves in to fill the hole and creates another hole where it was before (Fig. 4.2b). The process is repeated and the net effect is that the hole moves in the opposite direction to the electron. It behaves as an ordinary particle with positive energy, but having charge and momentum opposite to the electron. Its spin component is also opposite, because when the hole is filled, the quantum number is that of the vacuum. Thus, Dirac predicted the existence of the antiparticle and that a pair creation of a particle and antiparticle from vacuum will result if enough energy is injected. Thus materialization of pure energy became a possibility, as was already implicit in Einstein's equation $E = mc^2$.

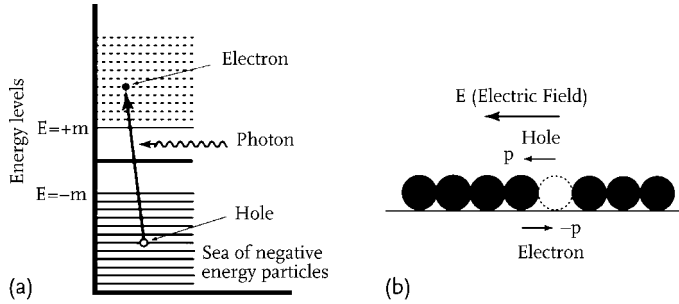


Figure 4.2 (a) The vacuum is a sea of negative-energy electrons. A hole is created by exciting an electron to a positive energy level. (b) The hole moves in the opposite direction to the electron.

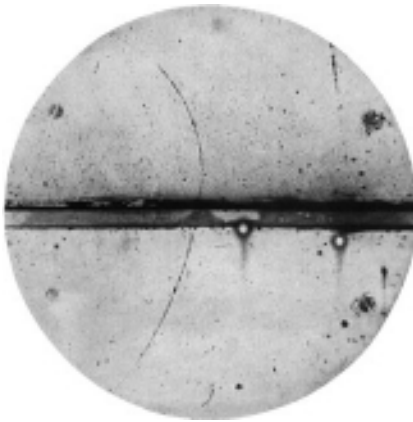


Figure 4.3 Anderson's discovery of the positron in a Wilson cloud chamber [21]. The incident charged particle in the cosmic rays leaves a curved track in the magnetic field. The particle loses energy after passing through a 9 mm thick lead plate, leading to a larger curvature (smaller radius). Therefore it has passed through the plate from below.

From the orientation of the magnetic field and the direction of the bend, the charge of the particle is determined to be positive. From the difference of the curvature before and after the passage of the plate, the energy loss, hence the mass, of the particle can be deduced. It was concluded that it had much lighter mass than the proton. Photo from [22].

Dirac considered first that the antiparticle might be the proton, but it soon became clear that the antiparticle must have the same mass, same spin, but opposite spin component as well as the opposite magnetic moment. In 1932, C.D. Anderson discovered the positron [21] (see Fig. 4.3), a positively charged particle that has the same mass as the electron, in cosmic rays using a cloud chamber and confirmed Dirac's theory.

Particles Traveling in the Opposite Time Direction

Although the existence of the antiparticle was confirmed experimentally, Dirac's hole theory cannot be applied to bosons, for which the Pauli exclusion principle

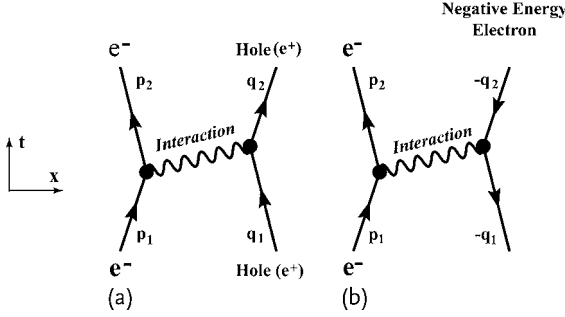


Figure 4.4 The antiparticle is (a) a hole in the sea of negative electrons or (b) an electron traveling in the reverse time direction.

does not apply.⁵⁾ To avoid the use of the Pauli exclusion principle, let us take a minute to consider how an electron with negative energy behaves in the hole theory. Suppose an electron and an antiparticle having energy-momentum p_1 and q_1 are scattered into a state with p_2 and q_2 (Fig. 4.4a). The process can be expressed by

$$p_1 + q_1 \rightarrow p_2 + q_2 \quad (4.29)$$

In the hole theory, a hole with q_1 became a hole with q_2 , as depicted in Figure 4.4a, but this is a reinterpretation of an event where an electron with $-q_2$ was scattered into an electron with $-q_1$. Writing the equation in terms of the original electron gives

$$p_1 + (-q_2) \rightarrow p_2 + (-q_1) \quad (4.30)$$

Here the scattering of the electron with $(-q_2)$ precedes the filling of the hole $(-q_1)$. Equation (4.29) and Eq. (4.30) are equivalent mathematically. What actually happens is Eq. (4.29), but we can consider it as Eq. (4.30), but with the negative particle traveling backward in time. The interpretation is that when a negative-energy particle of energy-momentum $-p^\mu = (-E, -\mathbf{p})$, $E > 0$, travels backward in time, it is observed as an antiparticle of positive energy-momentum p^μ traveling in the normal time direction, as is depicted in Figure 4.4b. The arrow indicates the time direction.

The interpretation is valid for bosons, where no Pauli exclusion principle can be applied. Of course, in reality, there are no backward-traveling particles. This is an example of how one has to make a physically correct interpretation of a mathematical equation that exhibits an apparent contradiction. Feynman and Stückelberg abandoned the idea of the negative-energy sea and only applied the concept of backward traveling to the negative-energy particle.

5) The concept of holes is useful and is routinely used in today's condensed matter physics. In the ground state, at zero absolute temperature, all the electrons occupy states from the lowest level up to the Fermi level, which may be called the vacuum state. Then, at finite temperature, some electrons are excited to above the Fermi level and holes are created. Many phenomena are successfully described using particle-hole interactions.

Let us consider a wave field instead of a wave function, obeying the spinless Klein–Gordon equation for simplicity. The wave with negative frequency, i.e. $p^\mu = (-E, -\mathbf{p})$, $E > 0$, can be written as

$$e^{-i(p^0 t - \mathbf{p} \cdot \mathbf{x})} \Big|_{p^\mu} \rightarrow e^{i\{(-E)(-t) - (-\mathbf{p}) \cdot (-\mathbf{x})\}} = e^{i(Et - \mathbf{p} \cdot \mathbf{x})} \quad (4.31)$$

This means a wave with negative frequency, i.e. $p^\mu = (-E, -\mathbf{p})$, $E > 0$, traveling backward in time as well as in space direction gives a wave traveling in the same direction as that with positive frequency. A current created by a plane wave field $\phi = N e^{-i p x}$ is expressed as

$$j^\mu = i e (\phi^* \partial^\mu \phi - \partial^\mu \phi^* \phi) \sim e |N|^2 p^\mu = e |N|^2 (E, \mathbf{p}) \quad (4.32)$$

The plane wave with negative frequency $p^\mu = (-E, -\mathbf{p})$ gives a current

$$j^\mu = e |N|^2 (-|E|, -\mathbf{p}) = -e |N|^2 (|E|, \mathbf{p}) \quad (4.33)$$

which is the same current as the positive energy current, only the charge is opposite. If we interpret ψ not as a relativistic probability amplitude but as a classical wave field, the negative energy simply means negative frequency and presents no difficulty in its interpretation.

The concept of backward time traveling, though unnecessary once quantized field theory is introduced, brings in a new concept about scattering phenomena.

Consider a particle A proceeding in time. It will behave as described in Fig. 4.5a, where kinks denote interactions. In the conventional way of thinking, the particle never travels backward in time. Using the new concept of backward travel for a

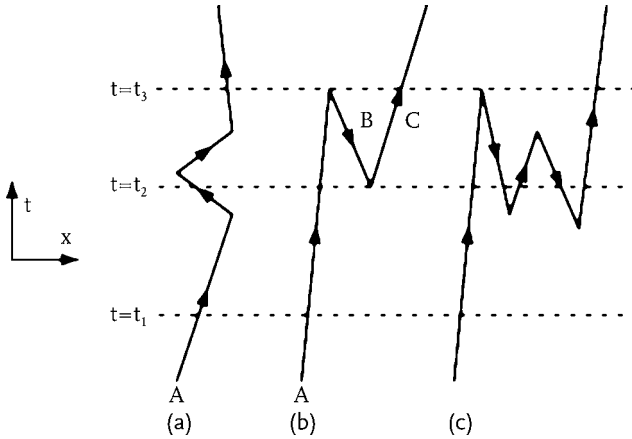


Figure 4.5 If backward time travel is allowed, a particle–antiparticle pair appears. (a) Ordinary scattering. (b) When backward travel is allowed, a scattering backwards in time occurs. From the observational point of view, an antiparticle B and a particle C appear at time

$t = t_2$. The antiparticle and the original particle A annihilate each other and disappear at time $t = t_3$. (c) Within an interval allowed by the Heisenberg uncertainty principle, any number of particles can appear.

negative-energy particle, a process like Fig. 4.5b is allowed. Observationally no particle travels backward in time. What really happens is that an antiparticle B and a particle C appear at time $t = t_2$, and subsequently the particle B annihilates with the original particle A and disappears at time $t = t_3$. The particle C takes over the part of A and proceeds onwards. A pair creation from vacuum has occurred. As a matter of fact, any number of particles can be created in a time interval allowed by the Heisenberg uncertainty principle. The charge has to be conserved, which means the backward-traveling particle must have the opposite charge.

What we have learned is, whether a sea of negative energy or backward travel in time is considered, a negative-energy state forces us to go from a one-particle description to a multiparticle world. A backward-traveling particle induces vacuum fluctuation. The vacuum is no longer a stable static state, but rather like boiling water, constantly creating and annihilating pairs of particles and antiparticles.

4.1.4

Lorentz-Covariant Dirac Equation

The Schrödinger equation involves special treatment of the time variable or energy and its form is not manifestly covariant. For a covariant treatment, we multiply both sides of Eq. (4.20) by β and rearrange the equation to get the Dirac equation in a covariant form:

$$(\gamma^\mu p_\mu - m) \psi = (\gamma^\mu i \partial_\mu - m) \psi(x) = 0 \quad (4.34)$$

where we have defined the Dirac γ matrices by

$$\gamma^0 = \beta = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{\gamma} = (\gamma^1, \gamma^2, \gamma^3) = \beta \boldsymbol{\alpha} = \begin{bmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{bmatrix} \quad (4.35)$$

They satisfy the following relations

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = g^{\mu\nu} \quad (4.36)$$

$$\gamma^{0\dagger} = \gamma^0, \quad \boldsymbol{\gamma}^\dagger = -\boldsymbol{\gamma} \quad (4.37)$$

Equation (4.36) is called the Dirac condition for gamma matrices. To obtain a conjugate form of the Dirac equation, we take the hermitian conjugate of Eq. (4.34) and multiply by γ^0 from the right. Using $\gamma^0 \gamma^{\mu\dagger} \gamma^0 = \gamma^\mu$ we obtain

$$\overline{\psi} (\gamma^\mu i \overleftarrow{\partial}_\mu + m) = 0, \quad \overline{\psi} = \psi^\dagger \gamma^0 \quad (4.38)$$

$\overline{\psi}$ is called the adjoint of ψ . The equation of continuity is now expressed as

$$\partial_\mu j^\mu = \partial_\mu (\overline{\psi} \gamma^\mu \psi) = 0 \quad (4.39)$$

The form suggests that the current $\bar{\psi}\gamma^\mu\psi$ is a Lorentz 4-vector. Indeed if we define $S^{\mu\nu}$ by

$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu], \quad S^{\mu\nu} = -S^{\nu\mu} \quad (4.40a)$$

$$S^{ij} = \frac{1}{2} \begin{bmatrix} \boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma} \end{bmatrix} \equiv \frac{1}{2} \boldsymbol{\Sigma}, \quad S^{0i} = \frac{i}{2} \alpha_i = \frac{i}{2} \begin{bmatrix} -\boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma} \end{bmatrix} \quad (4.40b)$$

they satisfy the commutation relations Eq. (3.74) for Lorentz group generators. Therefore

$$S(L) = \exp\left(-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}\right) = \exp(\boldsymbol{\alpha} \cdot \boldsymbol{\eta}/2 - i\boldsymbol{\Sigma} \cdot \boldsymbol{\theta}/2) \quad (4.41)$$

is a Lorentz operator acting on $\psi(x)$. The above formula shows that the rotation operator is unitary but the boost operator is not.⁶ Using Eq. (4.41) and the Dirac condition Eq. (4.36) for the gamma matrices, we obtain

$$S^{-1} = \gamma^0 S^\dagger \gamma^0 \quad (4.42)$$

A Lorentz-transformed spinor takes the form [see Eq. (3.69)]

$$\psi(x) \xrightarrow{L} \psi'(x) = D(L)\psi(L^{-1}x) \quad (4.43)$$

or equivalently, expressed in terms of $S(L)$,

$$\psi(x) \rightarrow \psi'(x') = S(L)\psi(x), \quad \bar{\psi}(x) \rightarrow \bar{\psi}'(x') = \bar{\psi}(x)S^{-1}(L) \quad (4.44)$$

A bit of algebra shows $\bar{\psi}\psi$ and $\bar{\psi}\gamma^\mu\psi$ are a Lorentz scalar and a vector, respectively. Although ψ itself is not a vector, its bilinear form with constant matrices sandwiched in between behaves like a vector (or a tensor) and is called a spinor.

Problem 4.5

For Lorentz boost in the x direction Eq. (4.41) becomes

$$S(L_x) = \exp(\alpha_x \eta_x/2) \quad (4.45)$$

Show $\bar{\psi}\gamma^\mu\psi$ follows the same transformation law as given in Eq. (3.62) and that $\bar{\psi}\psi$ remains unchanged.

More detailed treatment of the spinor is given in Appendix A. Loosely speaking, we may say the spinor is like the square root of the vector. Since

$$\bar{\psi}\gamma^\mu\psi \xrightarrow{L} \bar{\psi}'\gamma^\mu\psi' = \bar{\psi}S^{-1}\gamma^\mu S\psi = L^\mu_\nu \bar{\psi}\gamma^\nu\psi \quad (4.46a)$$

6) This is related to the fact that the Lorentz group is not a compact group. The range of the Lorentz parameter is $-1 < v < 1$ or $-\infty < \eta < \infty$. There is no finite-dimensional unitary representation. As the argument becomes very mathematical, we only note that it is possible to realize states that transform as unitary representations of the Poincaré group.

we may simply write

$$\gamma^\mu \xrightarrow{L} \gamma^{\mu'} = S^{-1} \gamma^\mu S = L^\mu_\nu \gamma^\nu \quad (4.46b)$$

In this sense we also refer to γ^μ as a Lorentz 4-vector, but it is meaningful only when it is sandwiched between spinors.

Now we prove the Lorentz covariance of the Dirac equation. By a Lorentz transformation Eq. (4.34) should take the form

$$(\gamma^\mu i \partial'_\mu - m) \psi'(x') = 0 \quad (4.47)$$

Inserting (4.44) and using

$$\partial'_\mu = (L^{-1})^\rho_\mu \partial_\rho \quad (4.48)$$

Eq. (4.47) becomes

$$[\gamma^\mu (L^{-1})^\rho_\mu i \partial_\rho - m] S \psi(x) = 0 \quad (4.49)$$

Multiplying by S^{-1} from left and using Eq. (4.46b), we obtain

$$\begin{aligned} \left[(S^{-1} \gamma^\mu S) (L^{-1})^\rho_\mu i \partial_\rho - m \right] \psi(x) &= [L^\mu_\sigma (L^{-1})^\rho_\mu \gamma^\sigma i \partial_\rho - m] \psi(x) = 0 \\ \therefore (\gamma^\mu i \partial_\mu - m) \psi(x) &= 0 \end{aligned} \quad (4.50)$$

Thus the Dirac equation has been proved to keep its form under the Lorentz transformation.

4.2

Plane-Wave Solution

Rewriting the four-component Dirac equation again in terms of two-component wave functions, we obtain

$$i(\partial_0 - \boldsymbol{\sigma} \cdot \nabla) \phi_L(x) = m \phi_R(x) \quad (4.51a)$$

$$i(\partial_0 + \boldsymbol{\sigma} \cdot \nabla) \phi_R(x) = m \phi_L(x) \quad (4.51b)$$

See Eq. (4.17), (4.34) and (4.35) to obtain the above equations. We see that the mass is the cause of mixing of ϕ_L and ϕ_R .⁷⁾ For plane-wave solutions, ϕ_L and ϕ_R can be

7) The reason why the two helicity states mix when the mass is finite can be understood as follows. The velocity of a particle with finite mass is always less than the speed of light c . Therefore it is possible to transform to a coordinate system that is flying faster than the particle. In that coordinate system,

the particle is moving backwards but rotating in the same direction. Therefore the helicity status is inverted. The amplitude for mixing is proportional to $\sim m/p$. If $m = 0$, there is no such coordinate system and hence the helicity does not mix.

expressed as

$$\phi_L(x) = N\phi_L(\mathbf{p})e^{-ipx} \quad (4.52a)$$

$$\phi_R(x) = N\phi_R(\mathbf{p})e^{-ipx} \quad (4.52b)$$

They satisfy the following coupled equations:

$$(p^0 + \boldsymbol{\sigma} \cdot \mathbf{p})\phi_L(\mathbf{p}) = m\phi_R(\mathbf{p}) \quad (4.53a)$$

$$(p^0 - \boldsymbol{\sigma} \cdot \mathbf{p})\phi_R(\mathbf{p}) = m\phi_L(\mathbf{p}) \quad (4.53b)$$

From Eq. (4.53a), we can express ϕ_R in terms of ϕ_L

$$\phi_R = \frac{1}{m}(p^0 + \boldsymbol{\sigma} \cdot \mathbf{p})\phi_L \quad (4.54)$$

Then inserting this equation into Eq. (4.53b), we obtain

$$\frac{1}{m}(p^0 - \boldsymbol{\sigma} \cdot \mathbf{p})(p^0 + \boldsymbol{\sigma} \cdot \mathbf{p})\phi_L = \frac{1}{m}(p^{02} - \mathbf{p}^2)\phi_L = m\phi_L \quad (4.55)$$

Therefore, ϕ_L satisfies the Klein–Gordon equation. Similarly ϕ_R satisfies the Klein–Gordon equation too. To obtain the solution to Eq. (4.53), we assume $p^0 = E > 0$ and rewrite $E, \mathbf{p}$ in terms of rapidity. Using $E = m\gamma = m \cosh \eta$, $\mathbf{p} = m\beta\gamma \hat{\mathbf{p}} = m \sinh \eta \hat{\mathbf{p}}$ [see Eq. (3.65)], we have

$$E \pm \boldsymbol{\sigma} \cdot \mathbf{p} = m (\cosh \eta \pm \sinh \eta \boldsymbol{\sigma} \cdot \hat{\mathbf{p}}) = m \exp(\pm \boldsymbol{\sigma} \cdot \boldsymbol{\eta}) \quad (4.56a)$$

$$\eta = \frac{1}{2} \ln \frac{E + p}{E - p}, \quad \hat{\mathbf{p}} = \frac{\mathbf{p}}{|\mathbf{p}|}, \quad p = |\mathbf{p}|, \quad \boldsymbol{\eta} = \eta \hat{\mathbf{p}} \quad (4.56b)$$

Note: When dealing with a transcendental function of matrices, we can always expand it in series. Thus

$$\exp(\boldsymbol{\sigma} \cdot \boldsymbol{\eta}) = 1 + (\boldsymbol{\sigma} \cdot \boldsymbol{\eta}) + \frac{(\boldsymbol{\sigma} \cdot \boldsymbol{\eta})^2}{2!} + \frac{(\boldsymbol{\sigma} \cdot \boldsymbol{\eta})^3}{3!} + \dots \quad (4.57a)$$

Using

$$(\boldsymbol{\sigma} \cdot \boldsymbol{\eta})^{2n} = \eta^{2n} \mathbf{1}, \quad (\boldsymbol{\sigma} \cdot \boldsymbol{\eta})^{2n+1} = \eta^{2n+1} (\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}) \quad (4.57b)$$

and rearranging the series, we obtain

$$\begin{aligned} &= \left(1 + \frac{\eta^2}{2!} + \dots\right) + \left(\eta + \frac{\eta^3}{3!} + \dots\right) (\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}) \\ &= \cosh \eta + \sinh \eta (\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}) \end{aligned} \quad (4.57c)$$

Let us introduce four-component spin vectors σ^μ , $\bar{\sigma}^\mu$ and dot products by

$$\sigma^\mu = (\mathbf{1}, \boldsymbol{\sigma}) = \bar{\sigma}_\mu, \quad \bar{\sigma}^\mu = (\mathbf{1}, -\boldsymbol{\sigma}) = \sigma_\mu \quad (4.58a)$$

$$p \cdot \sigma = p^\mu \sigma_\mu = E - \boldsymbol{\sigma} \cdot \mathbf{p}, \quad (4.58b)$$

$$p \cdot \bar{\sigma} = p^\mu \bar{\sigma}_\mu = E + \boldsymbol{\sigma} \cdot \mathbf{p} \quad (4.58c)$$

Then multiplying by $\sqrt{p \cdot \sigma} = \sqrt{m} \exp(-\boldsymbol{\sigma} \cdot \boldsymbol{\eta}/2)$ on both sides of Eq. (4.53a) and using $(E - \boldsymbol{\sigma} \cdot \mathbf{p})(E + \boldsymbol{\sigma} \cdot \mathbf{p}) = E^2 - \mathbf{p}^2 = m^2$, we obtain

$$\sqrt{p \cdot \sigma} \phi_L = \sqrt{p \cdot \sigma} \phi_R \quad (4.59)$$

which has solutions

$$\phi_L = N \sqrt{p \cdot \sigma} \xi_r = N \sqrt{m} \exp\left(-\frac{1}{2} \boldsymbol{\sigma} \cdot \boldsymbol{\eta}\right) \xi_r \quad (4.60a)$$

$$\phi_R = N \sqrt{p \cdot \bar{\sigma}} \xi_r = N \sqrt{m} \exp\left(+\frac{1}{2} \boldsymbol{\sigma} \cdot \boldsymbol{\eta}\right) \xi_r \quad (4.60b)$$

Combining the two two-component wave functions to obtain a four-component solution

$$u_r(p) \equiv \begin{bmatrix} \phi_L \\ \phi_R \end{bmatrix}_{E>0} = N \begin{bmatrix} \sqrt{p \cdot \sigma} \xi_r \\ \sqrt{p \cdot \bar{\sigma}} \xi_r \end{bmatrix} = N \sqrt{m} \exp\left(\frac{1}{2} \boldsymbol{\alpha} \cdot \boldsymbol{\eta}\right) \begin{bmatrix} \xi_r \\ \xi_r \end{bmatrix} \quad (4.61a)$$

where ξ_r , ($r = 1, 2$) is a basis of two-component spin eigenstates, and N is a normalization constant taken as $N = 1$ because we adopt normalizations such that there are $2E$ particles per volume. (See Sect. 6.5.2 for detailed discussion.) Comparing with Eq. (4.41), we see the plane-wave solution $\psi(x) \sim u(\mathbf{p})e^{-ipx}$ is nothing but a Lorentz-boosted spinor from the rest frame. For spin functions, it is often convenient to use the helicity eigenstate $\chi_\pm$ instead of ξ_r . In that case, $p \cdot \sigma$ and $p \cdot \bar{\sigma}$ acting on them are pure c-numbers.

For the solution with $p^0 = -E < 0$, we use $e^{ipx} = e^{i(Et - \mathbf{p} \cdot \mathbf{x})}$ for a plane-wave expression instead of $e^{-ipx}|_{p^0=-E} = e^{i(E + \mathbf{p} \cdot \mathbf{x})}$ and invert the momentum together to represent a particle propagating in the same direction as that of positive energy. Replacing $(p^0 \pm \boldsymbol{\sigma} \cdot \mathbf{p}) \rightarrow -(E \pm \boldsymbol{\sigma} \cdot \mathbf{p})$ in Eq. (4.53), the solution becomes

$$v_r(p) \equiv \begin{bmatrix} \phi_L \\ \phi_R \end{bmatrix}_{E<0} = \begin{bmatrix} \sqrt{p \cdot \sigma} \eta_r \\ -\sqrt{p \cdot \bar{\sigma}} \eta_r \end{bmatrix}, \quad \eta = i\sigma_2 \xi^* = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \xi^* \quad (4.61b)$$

where η , a flipped spin state, was chosen instead of ξ for reasons which will later become clear (see Sect. 4.3.4). Sometimes, a somewhat different expression for $u_r(p)$ and $v_s(\mathbf{p})$ may be useful. Going back to Eq. (4.60) and using

$$\begin{aligned} \exp\left(\mp \frac{1}{2} \boldsymbol{\sigma} \cdot \boldsymbol{\eta}\right) &= \cosh \frac{\eta}{2} \mp \sinh \frac{\eta}{2} \boldsymbol{\sigma} \cdot \hat{\boldsymbol{\eta}} = \cosh \frac{\eta}{2} \left(1 \mp \tanh \frac{\eta}{2} \boldsymbol{\sigma} \cdot \hat{\mathbf{p}}\right) \\ \cosh \frac{\eta}{2} &= \sqrt{\frac{E+m}{2m}} \quad \sinh \frac{\eta}{2} = \sqrt{\frac{E-m}{2m}} \\ \tanh \frac{\eta}{2} &= \sqrt{\frac{E-m}{E+m}} = \frac{p}{E+m} \end{aligned} \quad (4.62)$$

the plane-wave solutions are expressed as

$$u_r(\mathbf{p}) = \begin{bmatrix} \sqrt{p \cdot \vec{\sigma}} \xi_r \\ \sqrt{p \cdot \vec{\sigma}} \xi_r \end{bmatrix} = \sqrt{\frac{E+m}{2}} \begin{bmatrix} \left(1 - \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m}\right) \xi_r \\ \left(1 + \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m}\right) \xi_r \end{bmatrix} \quad (4.63a)$$

$$v_r(\mathbf{p}) = \begin{bmatrix} \sqrt{p \cdot \vec{\sigma}} \eta_r \\ -\sqrt{p \cdot \vec{\sigma}} \eta_r \end{bmatrix} = \sqrt{\frac{E+m}{2}} \begin{bmatrix} \left(1 - \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m}\right) \eta_r \\ -\left(1 + \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m}\right) \eta_r \end{bmatrix} \quad (4.63b)$$

These solutions satisfy the following relations:

$$\bar{u}_r(\mathbf{p}) u_s(\mathbf{p}) = 2m \delta_{rs}, \quad u_r^\dagger(\mathbf{p}) u_s(\mathbf{p}) = 2E \delta_{rs} \quad (4.64a)$$

$$\bar{v}_r(\mathbf{p}) v_s(\mathbf{p}) = -2m \delta_{rs}, \quad v_r^\dagger(\mathbf{p}) v_s(\mathbf{p}) = 2E \delta_{rs} \quad (4.64b)$$

$$\bar{u}_r(\mathbf{p}) v_s(\mathbf{p}) = \bar{v}_r(\mathbf{p}) u_s(\mathbf{p}) = 0 \quad (4.64c)$$

$$u_r^\dagger(\mathbf{p}) v_s(-\mathbf{p}) = v_r^\dagger(-\mathbf{p}) u_s(\mathbf{p}) = 0 \quad (4.64d)$$

$$\text{Note: } u_r^\dagger(\mathbf{p}) v_s(\mathbf{p}) \neq 0, \quad v_r^\dagger(\mathbf{p}) u_s(\mathbf{p}) \neq 0 \quad (4.64e)$$

Problem 4.6

Prove Eq. (4.64).

If we take the z -axis in the direction of the momentum, $\chi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\chi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, all the plane-wave solutions can be written down concisely.

Problem 4.7

Using c-number p^μ , prove

$$(\not{p} - m)u(\mathbf{p}) = 0, \quad \bar{u}(\mathbf{p})(\not{p} - m) = 0 \quad (4.65a)$$

$$(\not{p} + m)v(\mathbf{p}) = 0, \quad \bar{v}(\mathbf{p})(\not{p} + m) = 0 \quad (4.65b)$$

$$\not{p} \equiv \gamma^\mu p_\mu \quad (4.65c)$$

The notation $\not{p}$ is referred as the Feynman slash.

Problem 4.8

Denoting the a -th component of $u_r(\mathbf{p})$ as $\{u_r(\mathbf{p})\}_a$, show

$$\begin{aligned} [A_+]_{ab} &\equiv \frac{1}{2m} \sum_{r=1,2} \{u_r(\mathbf{p})\}_a \{\bar{u}_r(\mathbf{p})\}_b = \left[\frac{\not{p} + m}{2m} \right]_{ab} \\ [A_-]_{ab} &\equiv \frac{1}{2m} \sum_{r=1,2} \{v_r(\mathbf{p})\}_a \{\bar{v}_r(\mathbf{p})\}_b = \left[\frac{\not{p} - m}{2m} \right]_{ab} \end{aligned} \quad (4.66)$$

are projection operators for picking up positive and negative energy solutions respectively. Namely

$$\begin{aligned} \mathcal{A}_+ u_r(\mathbf{p}) &= u_r(\mathbf{p}), \quad \mathcal{A}_+ v_r(\mathbf{p}) = 0, \quad \mathcal{A}_- u_r(\mathbf{p}) = 0, \\ \mathcal{A}_- v_r(\mathbf{p}) &= v_r(\mathbf{p}) \end{aligned} \quad (4.67)$$

The properties of projection operators are

$$\mathcal{A}_+ + \mathcal{A}_- = 1, \quad \mathcal{A}_+^2 = \mathcal{A}_+, \quad \mathcal{A}_-^2 = \mathcal{A}_-, \quad \mathcal{A}_\pm \mathcal{A}_\mp = 0 \quad (4.68)$$

Equations (4.66) mean

$$\sum_{r=1,2} u_r(\mathbf{p}) \bar{u}_r(\mathbf{p}) = \not{p} + m, \quad \sum_{r=1,2} v_r(\mathbf{p}) \bar{v}_r(\mathbf{p}) = \not{p} - m \quad (4.69)$$

which will appear frequently in the calculation of the transition amplitude.

4.3

Properties of the Dirac Particle

4.3.1

Magnetic Moment of the Electron

To understand the nonrelativistic properties of the Dirac particle, let us derive a low-energy approximation of the Dirac equation. We apply it to the electron, which has electric charge $q = -e$. Setting $\psi = \begin{bmatrix} u_A \\ u_B \end{bmatrix}$, first we write down the equation in terms of two-component wave functions. We also make the substitution $p_\mu \rightarrow p_\mu + eA_\mu$ to include the electromagnetic interaction:

$$\begin{aligned} \{\gamma^\mu (i\partial_\mu + eA_\mu) - m\} \psi(x) &= \{(i\partial_0 + e\phi)\gamma^0 - \boldsymbol{\gamma} \cdot (-i\nabla + e\mathbf{A}) - m\} \psi(x) \\ &= \begin{bmatrix} -m & i\partial_0 + e\phi - \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A}) \\ i\partial_0 + e\phi + \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A}) & -m \end{bmatrix} \begin{bmatrix} u_A \\ u_B \end{bmatrix} = 0 \end{aligned} \quad (4.70a)$$

This is a set of two equations for the two-component spinors u_A, u_B .

$$\{i\partial_0 + e\phi - \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A})\} u_B = m u_A \quad (4.70b)$$

$$\{i\partial_0 + e\phi + \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A})\} u_A = m u_B \quad (4.70c)$$

In the high-energy limit, where the mass can be neglected, equations for u_A and u_B are disconnected and represent two independent particles ($u_A \sim \phi_L, u_B \sim \phi_R$). But in the low-energy limit, the difference disappears and $u_A \approx u_B$. For this reason

we introduce⁸⁾

$$u_+ = \frac{1}{\sqrt{2}}(u_A + u_B) \quad u_- = \frac{1}{\sqrt{2}}(-u_A + u_B) \quad (4.71)$$

with the expectation $u_- \rightarrow 0$ in the nonrelativistic limit. Then the Dirac equation reads

$$(i\partial_0 + e\phi)u_+ - \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A})u_- = mu_+ \quad (4.72a)$$

$$(i\partial_0 + e\phi)u_- - \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A})u_+ = -mu_- \quad (4.72b)$$

Up to now the equation is exact. At this point we choose the $p^0 > 0$ solution and set

$$u_+ = \varphi e^{-i(m+T)t} \quad u_- = \chi e^{-i(m+T)t} \quad (4.73)$$

Then we can make the substitution $i\partial_0 \rightarrow m + T$, where T now represents the kinetic energy. Substituting Eq. (4.73) in Eq. (4.72) and applying the nonrelativistic approximation $T, e\phi \ll m$, Eq. (4.72b) becomes

$$\chi \simeq \frac{1}{2m} \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A})\varphi \quad (4.74)$$

Substituting this expression in Eq. (4.72a), we obtain

$$T\varphi = \left[-e\phi + \frac{1}{2m} \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A}) \boldsymbol{\sigma} \cdot (-i\nabla + e\mathbf{A}) \right] \varphi \quad (4.75)$$

Then we make use of the relations

$$\begin{aligned} (\boldsymbol{\sigma} \cdot \mathbf{C})(\boldsymbol{\sigma} \cdot \mathbf{D}) &= \mathbf{C} \cdot \mathbf{D} + i\boldsymbol{\sigma} \cdot (\mathbf{C} \times \mathbf{D}) \\ (-i\nabla + e\mathbf{A}) \times (-i\nabla + e\mathbf{A})\varphi &= -ie(\nabla \times \mathbf{A} + \mathbf{A} \times \nabla)\varphi \\ &= -ie(\langle\langle \nabla \times \mathbf{A} \rangle\rangle \varphi + \langle\langle \nabla \varphi \rangle\rangle \times \mathbf{A} + \langle\langle \mathbf{A} \times \nabla \varphi \rangle\rangle) = -ie\mathbf{B}\varphi \end{aligned} \quad (4.76)$$

Here $\langle\langle \dots \rangle\rangle$ means that the action of the nabla does not go beyond the double bracket and $\mathbf{B}$ is the magnetic field. After inserting Eq. (4.76) in Eq. (4.75) and rearranging, we obtain

$$T\varphi = \left[-\frac{1}{2m}(\nabla + ie\mathbf{A})^2 - e\phi + \frac{e}{2m} \boldsymbol{\sigma} \cdot \mathbf{B} \right] \varphi \quad (4.77)$$

which is the Schrödinger equation including the electromagnetic potential plus an extra magnetic interaction term $-\boldsymbol{\mu} \cdot \mathbf{B}$. This means that the Dirac particle has a magnetic moment

$$\boldsymbol{\mu} = -\frac{e}{2m} \boldsymbol{\sigma} = g \frac{q}{2m} \mathbf{s} \quad (g = 2) \quad (4.78)$$

8) This is equivalent to adopting the Pauli–Dirac representation in Eq. (4.19).

Using the general expression for the magnetic moment $\boldsymbol{\mu} = g(q/2m)\mathbf{s}$, this means that the Dirac particle has a g -factor of 2. This solved the then mysterious value of $g = 2$ for the spin magnetic moment of the electron. This fact again defies the classical notion of a spin being the angular momentum of a self-rotating object. From classical electromagnetism, a circulating electric charge should have $g = 1$.

The classical magnetic moment of a circulating charge can be calculated as follows. The Maxwell equation gives the value of the magnetic moment $\mu = IS$ to a closed loop current I with inner area S . Assume the current is made by a particle with charge q circulating at radius r . Taking its velocity v and circulating period T , or the frequency $\nu = 1/T$,

$$\mu = IS = q\nu\pi r^2 = q\frac{v}{2\pi r}\pi r^2 = \frac{q}{2m}rmv = \frac{q}{2m}L \quad (4.79)$$

(L = angular momentum)

Deviation of the value g from 2 is called an anomalous magnetic moment. For instance, the proton and the neutron have magnetic moments of 2.79 and -1.91 (in units of the nuclear magneton $q/2m_N$) and their deviation from the standard value is large, which exposes the nonelementary nature of the particles. Even for the Dirac particle $g - 2$ could be nonzero if we take into account higher order effects (radiative corrections) of the quantum electromagnetic fields (see arguments in Sect. 8.2.2).

4.3.2

Parity

In Sect. 4.1.2, we saw that ϕ_R and ϕ_L are exchanged by parity operation and introduced the β matrix to do the action. Here, we treat the parity operation of the four-component Dirac equation

$$(\gamma^\mu i\partial_\mu - m)\psi(x) = 0 \quad (4.80)$$

By parity operation, the Dirac equation is transformed to

$$(\gamma^\mu i\partial_\mu' - m)\psi'(x') = 0 \quad (4.81a)$$

$$x^\mu = (x^0, \mathbf{x}) \rightarrow x^{\mu'} = (x^0, -\mathbf{x}) = x_\mu \quad (4.81b)$$

$$\psi(x) \rightarrow \psi'(x') = S\psi(x) \quad (4.81c)$$

So if there is a matrix S that satisfies

$$S^{-1}\gamma^0 S = \gamma^0, \quad S^{-1}\gamma^k S = -\gamma^k \quad (k = 1 \text{ to } 3) \quad (4.82)$$

We can claim that the Dirac equation is parity reflection invariant. Obviously

$$S = \eta_P \gamma^0 \quad (4.83)$$

satisfies the condition as expected. Here, η_P is a phase factor with the absolute value of 1. Applying the parity operation to the plane-wave solutions (4.61a) and (4.61b), we obtain

$$S u_r(-\mathbf{p}) = \eta_P u_r(\mathbf{p}) \quad (4.84a)$$

$$S v_r(-\mathbf{p}) = -\eta_P v_r(\mathbf{p}) \quad (4.84b)$$

which means that a particle and its antiparticle have relatively negative parity. Usually $\eta_P = 1$ is assumed and hence the particle has positive parity and the antiparticle negative.

4.3.3

Bilinear Form of the Dirac Spinor

A hermitian bilinear form made of four-component Dirac spinors $T = \bar{\psi} \Gamma \psi$ has 16 independent components. We have seen $V^\mu \equiv \bar{\psi} \gamma^\mu \psi$ behaves as a contravariant vector under proper Lorentz transformation. Under parity operation,

$$V^{0'} = \bar{\psi} S^{-1} \gamma^0 S \psi = \bar{\psi} \gamma^0 \gamma^0 \gamma^0 \psi = \bar{\psi} \gamma^0 \psi = V^0 \quad (4.85a)$$

$$V' = \bar{\psi} \gamma^0 \boldsymbol{\gamma} \gamma^0 \psi = -\bar{\psi} \boldsymbol{\gamma} \psi = -V \quad (4.85b)$$

Therefore, V^μ is a polar vector. Introducing

$$\gamma_\mu = g_{\mu\nu} \gamma^\nu, \quad \sigma_{\mu\nu} = g_{\mu\rho} g_{\nu\sigma} \sigma^{\rho\sigma} \quad (4.86)$$

Equation (4.85) can be expressed as

$$V^\mu \xrightarrow{P} V^{\mu'} = V_\mu \quad (4.87)$$

All the 16 independent bilinear forms are classified according to their Lorentz transformation property and are summarized in Table 4.1.

Problem 4.9

Confirm the parity transformation property of each item in Table 4.1.

Problem 4.10

Using (4.63), prove that the matrix elements $\langle p' | \boldsymbol{\Gamma} | p \rangle = \bar{u}(\mathbf{p}') \boldsymbol{\Gamma} u(\mathbf{p})$ in the non-relativistic limit are given to the order $O(p/m)$ as shown in Table 4.2.

Table 4.1 Sixteen independent bilinear forms $\bar{\psi} \Gamma \psi$.

Name		Bilinear form	# ^a	After parity transf.	Parity
Scalar	S	$\bar{\psi} \psi$	1	$\bar{\psi} \psi$	+
Pseudoscalar	P	$i \bar{\psi} \gamma^5 \psi$ ^b	1	$-i \bar{\psi} \gamma^5 \psi$	−
Vector	V	$\bar{\psi} \gamma^\mu \psi$	4	$\bar{\psi} \gamma_\mu \psi$	−
Axial vector	A	$\bar{\psi} \gamma^5 \gamma^\mu \psi$	4	$-\bar{\psi} \gamma^5 \gamma_\mu \psi$	+
Tensor	T	$\bar{\psi} \sigma^{\mu\nu} \psi$	6	$\bar{\psi} \sigma_{\mu\nu} \psi$	+

a # is the number of independent components.

b $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. See Eq. (4.106). i is multiplied to make the form hermitian.

Table 4.2 Nonrelativistic limit of the γ matrices for plane wave solution.

Γ	$\langle \mathbf{p}' \Gamma \mathbf{p} \rangle / 2m$
1	1
γ^5	$-\langle \boldsymbol{\sigma} \cdot (\mathbf{p}' - \mathbf{p}) \rangle / (2m)$
γ^0	1
$\boldsymbol{\gamma}$	$(\mathbf{p} + \mathbf{p}') / (2m) + [i \boldsymbol{\sigma} \times (\mathbf{p}' - \mathbf{p})] / (2m)$
$\gamma^5 \gamma^0$	$-\langle \boldsymbol{\sigma} \cdot (\mathbf{p}' + \mathbf{p}) \rangle / (2m)$
$\gamma^5 \boldsymbol{\gamma}$	$-\boldsymbol{\sigma}$
$\sigma^{ij} = \frac{i}{2} [\gamma^i, \gamma^j]$	$\epsilon_{ijk} \sigma_k$ ^a
$\sigma^{k0} = \frac{i}{2} [\gamma^k, \gamma^0]$	$i\{(\mathbf{p}' - \mathbf{p}) / (2m) + [i \boldsymbol{\sigma} \times (\mathbf{p}' + \mathbf{p})] / (2m)\}$

a ijk is a cyclic of 123.

Using the fact that in the nonrelativistic limit $\mathbf{p} \approx m\mathbf{v}$, the above quantities can be expressed as

$$\langle 1 \rangle \sim 1, \quad \langle \gamma^0 \rangle \sim 1, \quad \langle \gamma^\mu \rangle \sim (1, \mathbf{v}), \quad \langle \gamma^\mu \gamma^5 \rangle \sim (\boldsymbol{\sigma} \cdot \mathbf{v}, \boldsymbol{\sigma}) \quad (4.88)$$

These expressions will help in understanding the physical interpretation or dynamical properties of $\langle \Gamma \rangle$.

4.3.4

Charge Conjugation

Charge conjugation is the operation of exchanging a particle and its antiparticle. The fact that the signature of an antiparticle appears as the opposite sign in the phase of the wave suggests a connection with complex conjugate operation. The difference between the particle and its antiparticle manifests itself in the interaction with electromagnetic fields. Interaction with the electromagnetic field can be

taken into account by the substitution

$$p^\mu \rightarrow p^\mu - qA^\mu \quad (4.89)$$

and that of the antiparticle by replacing q by $-q$. So if we can transform the field which satisfies the same equation of motion but with opposite charge, the operation is the charge conjugation. Let us start with the simplest example, the Klein–Gordon equation. Substituting Eq. (4.89) in the Klein–Gordon equation,

$$[(\partial_\mu + iqA_\mu)(\partial^\mu + iqA^\mu) + m^2] \phi = 0 \quad (4.90)$$

Taking the hermitian conjugate of the above equation we have

$$[(\partial_\mu - iqA_\mu)(\partial^\mu - iqA^\mu) + m^2] \phi^\dagger = 0 \quad (4.91)$$

Comparing Eq. (4.90) with Eq. (4.91), we see that if ϕ describes a particle, then $\phi^\dagger$ describes its antiparticle, thus confirming our original guess. It also means that if ϕ is a real field, it cannot have antiparticle or it is its own antiparticle. It follows that a real field describes a neutral particle. The converse is not true. If a particle has a quantum number other than charge that differentiates the particle from its antiparticle, for example strangeness, the field is complex.

For the case of the Dirac particle, it is a little more complicated since mixing of the components could result from the charge conjugation operation. Let us start from the four-component Dirac equation including the electromagnetic interaction:

$$[\gamma^\mu(i\partial_\mu - qA_\mu) - m] \psi = 0 \quad (4.92)$$

If there exists ψ^c that satisfies the equation

$$[\gamma^\mu(i\partial_\mu + qA_\mu) - m] \psi^c = 0 \quad (4.93)$$

then ψ^c is the charge conjugate of ψ . Taking the hermitian conjugate of Eq. (4.92), we obtain

$$\psi^\dagger[\gamma^{\mu\dagger}(-i\overleftarrow{\partial}_\mu - qA_\mu) - m] = 0 \quad (4.94)$$

Multiplying by γ^0 from the right, and using the relation $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$, we obtain

$$\overline{\psi}[-\gamma^\mu(i\overleftarrow{\partial}_\mu + qA_\mu) - m] = 0 \quad (4.95)$$

Now, taking the transpose of the above equation

$$[-\gamma^{\mu T}(i\partial_\mu + qA_\mu) - m]\overline{\psi}^T = 0 \quad (4.96)$$

and multiplying by C from the left, we obtain

$$[-C\gamma^{\mu T}C^{-1}(i\overleftarrow{\partial}_\mu + qA_\mu) - m]C\overline{\psi}^T = 0 \quad (4.97)$$

So, if there is a matrix C that satisfies the relation

$$C \gamma^{\mu T} C^{-1} = -\gamma^{\mu} \quad (4.98)$$

we can define a charge conjugate operation and

$$\psi^c = C \bar{\psi}^T \quad (4.99)$$

satisfies the equation (4.93) for the antiparticle.

Noting

$$\gamma^{\mu T} = \gamma^{\mu} (\mu = 0, 2), \quad \gamma^{\mu T} = -\gamma^{\mu} (\mu = 1, 3) \quad (4.100)$$

we can adopt

$$C = i\gamma^2\gamma^0 = \begin{bmatrix} i\sigma_2 & \mathbf{0} \\ \mathbf{0} & -i\sigma_2 \end{bmatrix} \quad (4.101)$$

$i\sigma_2$ flips the spin orientation. Therefore, the charge conjugation operation includes spin flipping, as conjectured in Sect. 4.1.3. This is why we adopted the flipped-spin solution for the negative energy wave function (4.61b). This C matrix has the properties

$$C = C^* = -C^{-1} = -C^T = -C^\dagger \quad (4.102)$$

The charge-conjugated Dirac wave function is given by

$$\psi^c = C \bar{\psi}^T = i\gamma^2\gamma^0\gamma^{0T}\psi^* = i\gamma^2\psi^* \quad (4.103a)$$

$$\bar{\psi}^c = \psi^{c\dagger}\gamma^0 = (C\bar{\psi}^T)^\dagger\gamma^0 = \bar{\psi}^*C^\dagger\gamma^0 = \psi^T\gamma^{0*}C^\dagger\gamma^0 = -\psi^TC^{-1} \quad (4.103b)$$

Naturally, an electromagnetic current should change its sign under the charge-conjugation operation. In quantum mechanics, it is a definition and the sign has to be changed by hand, but in quantum field theory it comes automatically.

Proof:

$$\begin{aligned} j^{\mu c} &= q\bar{\psi}^c\gamma^{\mu}\psi^c = -q\psi^TC^{-1}\gamma^{\mu}C\bar{\psi}^T = q\psi^T\gamma^{\mu T}\bar{\psi}^T \\ &= -q(\bar{\psi}\gamma^{\mu}\psi)^T = -q\bar{\psi}\gamma^{\mu}\psi = -j^{\mu} \end{aligned} \quad (4.104)$$

The minus sign in the first equation of the second line comes from the anticommutative nature of the fermion field, which we will show in Chap. 5. We can remove the transpose in the penultimate equation because the quantity in the parentheses is a 1×1 matrix. $\square$

Problem 4.11

$$\text{Prove } (u_r)^c = C \bar{u}_r^T = v_r, \quad (v_r)^c = C \bar{v}_r^T = u_r \quad (4.105)$$

4.3.5

Chiral Eigenstates

We define a chirality operator γ^5 by

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = -i\alpha_1\alpha_2\alpha_3 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4.106)$$

which satisfies the relation

$$(\gamma^5)^2 = 1, \quad \gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0 \quad (4.107)$$

We also define chiral eigenstates, which play an important role in the weak interaction, by

$$\psi_L(x) = \frac{1}{2}(1 - \gamma^5)\psi = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \phi_L \\ \phi_R \end{bmatrix} = \begin{bmatrix} \phi_L \\ 0 \end{bmatrix} \quad (4.108a)$$

$$\psi_R(x) = \frac{1}{2}(1 + \gamma^5)\psi = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_L \\ \phi_R \end{bmatrix} = \begin{bmatrix} 0 \\ \phi_R \end{bmatrix} \quad (4.108b)$$

It is easily seen that ψ_L and ψ_R are eigenstates of the chirality operator γ^5

$$\gamma^5\psi_L = -\psi_L, \quad \gamma^5\psi_R = +\psi_R \quad (4.109)$$

In the Weyl representation, the two chiral eigenstates are expressed in terms of the two-component spinors ϕ_L and ϕ_R separately. In other representations, they are not necessarily separated, but the following arguments are equally valid. Multiplying Eq. (4.34) by γ^5 and using Eq. (4.107) we can obtain an equation for $\gamma^5\psi$. Then adding and subtracting it from Eq. (4.34), we obtain

$$\gamma^\mu i\partial_\mu\psi_L = m\psi_R \quad (4.110a)$$

$$\gamma^\mu i\partial_\mu\psi_R = m\psi_L \quad (4.110b)$$

which reproduces the four-component version of Eq. (4.53). Since γ^5 anticommutes with all the γ^μ 's ($\mu = 0$ to 3), it commutes with the Lorentz operator (4.41). Therefore the chiral eigenstate is invariant under the proper Lorentz transformation. Under the parity operation

$$\psi(x) \xrightarrow{P} \psi'(x') = \psi'(x^0, -\mathbf{x}) = \gamma^0\psi(x) \quad (4.111)$$

Since γ^0 anticommutes with γ^5 , ψ_L is converted to ψ_R and vice versa. In the Weyl representation, ϕ_L and ϕ_R are the two chiral eigenstates.

Let us take a look at the spin properties of the plane-wave solution and how they are related to the Weyl solutions. We notice that in the high-energy or small-mass limit $p \rightarrow E$

$$\lim_{p \rightarrow E} p \cdot \sigma = 2E \left(\frac{1 - \boldsymbol{\sigma} \cdot \hat{\mathbf{p}}}{2} \right), \quad \lim_{p \rightarrow E} p \cdot \bar{\sigma} = 2E \left(\frac{1 + \boldsymbol{\sigma} \cdot \hat{\mathbf{p}}}{2} \right) \quad (4.112)$$

They are just projection operators to pick up the helicity $\mp$ state. This means ϕ_L and ϕ_R approach the Weyl solution in the high-energy limit ($p \rightarrow E$ or $m \rightarrow 0$). For $p \neq E$, $\sqrt{p \cdot \sigma}$, $\sqrt{p \cdot \bar{\sigma}}$ preferentially pick up helicity $\mp$ components. To decide the helicity expectation value of ϕ_L , ϕ_R , apply them to an unpolarized helicity state, i.e. $\xi_0 = \chi_+ + \chi_-$:

$$\phi_L = \sqrt{p \cdot \sigma} \xi_0 = \sqrt{E + p} \chi_- + \sqrt{E - p} \chi_+ \equiv (a \chi_- + b \chi_+) \quad (4.113)$$

Therefore

$$\begin{aligned} \phi_L &= (\sqrt{E + p} \chi_- + \sqrt{E - p} \chi_+) \\ &\xrightarrow{p \rightarrow E} \sqrt{2p} \left[\chi_- + \left(\frac{m}{2p} \right) \chi_+ + O\left(\frac{m}{p}\right)^2 \right] \\ \phi_R &= (\sqrt{E + p} \chi_+ + \sqrt{E - p} \chi_-) \\ &\xrightarrow{p \rightarrow E} \sqrt{2p} \left[\chi_+ + \left(\frac{m}{2p} \right) \chi_- + O\left(\frac{m}{p}\right)^2 \right] \end{aligned} \quad (4.114)$$

We note that ϕ_L is primarily a negative helicity state, i.e. left-handed, but contains a small $\sim O(m/p)$ component of the opposite helicity state. The expectation value of the helicity is given by

$$\langle h(\phi_L) \rangle = \frac{(\phi_L^\dagger \boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \phi_L)}{(\phi_L^\dagger \phi_L)} = \frac{-|a|^2 + |b|^2}{|a|^2 + |b|^2} = -\frac{p}{E} = -\nu \quad (4.115)$$

Similarly, we obtain $\langle h(\phi_R) \rangle = +\nu$. For antiparticles, whose wave function is represented by $v(\mathbf{p})$, the helicity is inverted. The chirality is a Lorentz-extended version of the helicity operator, with its handedness inverted for antiparticles. Knowledge of the chiral characteristics comes in very handy in understanding weak-interaction phenomenology. Note, however, the chiral states are not the eigenstates of the equation of motion, because γ^5 does not commute with the Hamiltonian. Helicity states are the Hamiltonian eigenstates, but they change their state by interactions. We have also seen in Eq. (4.110) that the two chiral states mix when the mass term exists in the Hamiltonian. But the chirality is Lorentz invariant and is conserved in vector interactions, which include the strong, weak and electromagnetic interactions.

Chirality Conservation in the Vector Interaction

We prove that chirality is conserved in vector-type interactions before and after interaction. The photon, for instance, couples to currents

$$\begin{aligned}
 \bar{\psi}_{L(R)} \gamma^\mu \psi_{R(L)} &= \psi^\dagger \left(\frac{1 \mp \gamma^5}{2} \right)^\dagger \gamma^0 \gamma^\mu \left(\frac{1 \pm \gamma^5}{2} \right) \psi \\
 &= \psi^\dagger \gamma^0 \gamma^\mu \left(\frac{1 \mp \gamma^5}{2} \right) \left(\frac{1 \pm \gamma^5}{2} \right) \psi = 0 \\
 \bar{\psi} \gamma^\mu \psi_{L(R)} &= \bar{\psi}_{L(R)} \gamma^\mu \psi = \bar{\psi}_{L(R)} \gamma^\mu \psi_{L(R)} \neq 0
 \end{aligned} \tag{4.116}$$

In the high-energy limit, where the mass can be neglected, chirality is equivalent to helicity except for the sign, hence helicity is conserved, too.

Note that an interaction that respects the gauge principle is obtained by the replacement $\partial_\mu \rightarrow \partial_\mu + iqA_\mu$. Namely, the interaction Lagrangian has the form $-qj^\mu A_\mu = -q\bar{\psi}\gamma^\mu\psi A_\mu$. It is good to remember that the fundamental interaction in the Standard Model is of vector type and preserves chirality.

Comment 4.1 *In high-energy experiments we are mainly interested in an energy range from 1 GeV to 1 TeV = 1000 GeV. The mass of the u, d, s quarks and that of the leptons (except the τ) are less than ~ 100 MeV (see Table 2.2). They can, for all practical purposes, be regarded as massless particles. For them the chiral eigenstates are essentially equivalent to the helicity eigenstates (flipped for their antiparticles) and hence are eigenstates of the Hamiltonian, too. c, b quarks and the τ lepton have mass in the few gigaelectronvolts range, and in many cases they can be treated similarly. Only for the top quark, whose mass is ~ 171 GeV, is the effect of the mass sizable. This is why we use the Weyl representation, which diagonalizes the chiral operator γ^5 , instead of the Dirac representation, which diagonalizes the mass term β . For historical reasons, the latter has been commonly used in most textbooks. It is convenient only in treating phenomena of relativistic origin in the nonrelativistic energy range.*

4.4

Majorana Particle

Quantum numbers that differentiate particles from antiparticles include electric charge, isospin, strangeness and lepton/baryon number.

Examples	$e^+ \neq e^-$	(electric charge, lepton number)
	$n \neq \bar{n}$	(magnetic moment, baryon number)
	$K^0 \neq \bar{K}^0$	(strangeness)

π^0 and the photon do not have such quantum numbers, and thus they cannot be distinguished from their antiparticles. Now, consider neutrinos. They belong to the lepton family and are considered to have lepton numbers just as charged leptons do. The lepton numbers 1 and -1 can be assigned to the neutrino and antineutrino,

respectively. Experimental observations seem to respect lepton number conservation in neutrino reactions. For instance, ν_μ and $\bar{\nu}_\mu$ are produced through the decay of pions:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu, \quad \pi^- \rightarrow \mu^- + \bar{\nu}_\mu \quad (4.117)$$

and when they collide with nucleons, they produce μ^- (lepton number 1) and μ^+ (lepton number -1)

$$\nu_\mu + n \rightarrow \mu^- + p, \quad \bar{\nu}_\mu + p \rightarrow \mu^+ + n \quad (4.118)$$

in such a way as to conserve the lepton number. A ν_μ does not produce a μ^+ if it collides with a proton. However, the weak interaction acts only on particles with negative chirality.⁹⁾ If the particle's mass is zero, it is a Weyl particle and is in a pure helicity (-) state, while the antiparticle is in pure helicity (+) state. What was considered as lepton number conservation may have been nothing but a test for helicity conservation, although experimentally there is no evidence against lepton number conservation as far as the neutrino is concerned.

Let us assume that the neutrino is a Majorana particle, a name given to a fermion that does not have its own antiparticle. Conditions for the Majorana particle are

1. It satisfies the Dirac equation.
 $(\gamma^\mu i\partial_\mu - m)N(x) = 0$
 2. The particle and its antiparticle are identical.
 $N^c = N$
- (4.119)

In the Weyl representation, we can put $\psi_L = \begin{bmatrix} \xi \\ 0 \end{bmatrix}$ and $\psi_R = \begin{bmatrix} 0 \\ \eta \end{bmatrix}$. Then define N_1 and N_2 as follows:

$$N_1 = \psi_L + (\psi_L)^c = \begin{bmatrix} \xi \\ -i\sigma_2 \xi^* \end{bmatrix} \quad (4.120a)$$

$$N_2 = \psi_R + (\psi_R)^c = \begin{bmatrix} i\sigma_2 \eta^* \\ \eta \end{bmatrix} \quad (4.120b)$$

They satisfy $N_1^c = N_1$, $N_2^c = N_2$. Applying the Dirac equation for N_1

$$[\gamma^\mu i\partial_\mu - m] N_1 = \begin{bmatrix} -m & i\partial_0 + i\boldsymbol{\sigma} \cdot \nabla \\ i\partial_0 - i\boldsymbol{\sigma} \cdot \nabla & -m \end{bmatrix} \begin{bmatrix} \xi \\ -i\sigma_2 \xi^* \end{bmatrix} = 0 \quad (4.121)$$

and rewriting the equation for two-component spinors, we obtain

$$\begin{aligned} (\partial_0 - \boldsymbol{\sigma} \cdot \nabla) \xi &= -m(\sigma_2 \xi^*) \\ (\partial_0 + \boldsymbol{\sigma} \cdot \nabla)(\sigma_2 \xi^*) &= m\xi \end{aligned} \quad (4.122)$$

9) This is true only for charged current interactions, where $W^\pm$ is exchanged. The neutral current interactions act on both chiralities.

The two equations are not independent. The second equation can be derived from the first. It is indeed an equation for a two-component spinor. N_1 and N_2 have only two independent components despite their seemingly four-component structure. Similarly we can obtain the equation for N_2 :

$$\begin{aligned}(\partial_0 + \boldsymbol{\sigma} \cdot \nabla)\eta &= m(\sigma_2 \eta^*) \\ (\partial_0 - \boldsymbol{\sigma} \cdot \nabla)(\sigma_2 \eta^*) &= -m\eta\end{aligned}\tag{4.123}$$

From the above argument, we see ξ and η satisfy two independent equations. In the limit of $m \rightarrow 0$, the equations become identical to the Weyl equation (4.9). If $m \neq 0$, the deviation from the Weyl solution is small for relativistic particles ($m \ll E$). ξ is in a chirality minus state and is predominantly left-handed. Similarly, η is predominantly right-handed. Therefore, we call ξ (η) a left (right)-handed Majorana particle. Since they satisfy independent equations, their mass need not be the same in general ($m_L \neq m_R$).

Note, the Majorana particle is a quantum mechanical concept and there is no equivalent in classical mechanics. If we consider ξ or η as expressed by pure classical numbers, it is straightforward to show that Equation (4.122) has no solution. This can be seen by inserting $\xi^T = (a, b)$ and $\mathbf{p} = 0$ and solving the equation explicitly for a and b . Another way to see this is to note that the Lagrangian to derive the mass term should contain a term like $\xi^\dagger i \sigma_2 \xi^* = i \xi_i^* \{\sigma_2\}_{ij} \xi_j^*$. However, σ_2 is antisymmetric and it would vanish if ξ were an ordinary c-number field. In quantum field theory, ξ is an anticommuting field operator and the contradiction does not appear.

Let us see the effect of charge conjugation on the left- and right-handed Majorana neutrinos. After decomposing the four-component spinors N_1 and N_2 into N_L and N_R using the projection operators $(1 \mp \gamma^5)/2$

$$N_L = \frac{1}{2}(1 - \gamma^5)N, \quad N_R = \frac{1}{2}(1 + \gamma^5)N\tag{4.124}$$

we apply the charge conjugation operation to them:

$$\begin{aligned}(N_L)^c &= C \overline{N}_L^T = i\gamma^2 \gamma^0 \overline{N}_L^T = i\gamma^2 \frac{1}{2}(1 - \gamma^5)N^* \\ &= \frac{1}{2}(1 + \gamma^5)i\gamma^2 \gamma^0 \overline{N}^T = \frac{1}{2}(1 + \gamma^5)N^c = N_R \\ \therefore (N_L)^c &= (N^c)_R = N_R\end{aligned}\tag{4.125a}$$

$$(N_R)^c = (N^c)_L = N_L\tag{4.125b}$$

The Majorana particle changes its handedness by the charge conjugation operation. No distinction exists between a particle and its antiparticle. The handedness can be

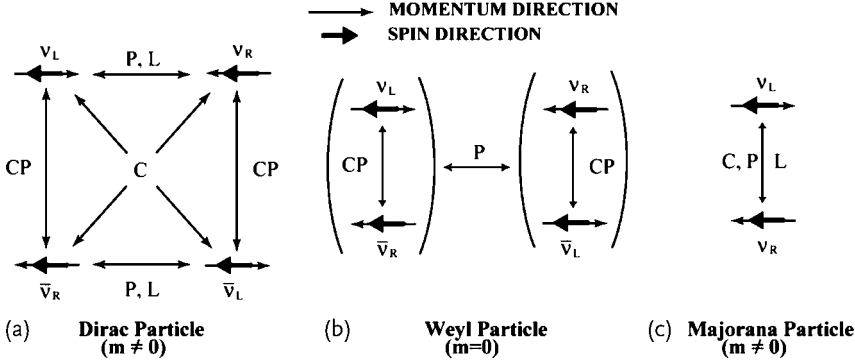


Figure 4.6 Three possibilities for the neutrino:
 (a) four-component Dirac neutrino $m \neq 0$,
 (b) two-component Weyl neutrino ($m = 0$)
 and (c) two-component Majorana neutrino

($m \neq 0$). P and C are the parity and charge-conjugation operations, respectively. L is a Lorentz boost. The neutrino states transform to each other by the designated operation.

exchanged either by the parity or the charge conjugation operation. From now on, we adopt the convention that $\psi_L^c \equiv (\psi^c)_L$, to denote left handed antiparticle for the Dirac particle and left handed particle for the Majorana particle.

Problem 4.12

Since $i\sigma_2\xi^*$ ($-i\sigma_2\eta^*$) satisfies the same equation as η (ξ), it is a right (left)-handed particle. Show that if $m_L = m_R$ then the chirality L(R) component of $\psi = (N_1 + iN_2)/\sqrt{2}$ satisfies Eq. (4.51).

Therefore the Dirac particle is equivalent to two Majorana particles having the same mass. Actually, having the same mass guarantees the Lagrangian to have an extra symmetry, namely invariance under phase transformation ($\psi \rightarrow \psi' = e^{-i\alpha}\psi$, $\psi^\dagger \rightarrow \psi'^\dagger = e^{i\alpha}\psi^\dagger$) and to have a conserved particle number according to Noether's theorem (see Sect. 5.2).

We summarize relative relations of the Dirac, Majorana and Weyl particles in Figure 4.6.

How Can We Tell the Difference Between the Dirac and Majorana Neutrinos?

If the neutrino is a Majorana particle, lepton number is not conserved. The simplest reaction that can test it irrespective of the helicity state is a neutrinoless (0ν) double beta decay, in which a nucleus emits two electrons with no neutrinos:

$$0\nu\beta\beta \text{ decay: } (A, Z) \rightarrow (A, Z + 2) + e^- + e^- \quad (4.126)$$

A double beta decay that emits two neutrinos (2ν) (to conserve the lepton number) appears as a higher order weak interaction and has been observed:

$$2\nu\beta\beta \text{ decay: } (A, Z) \rightarrow (A, Z + 2) + e^- + e^- + \bar{\nu}_e + \bar{\nu}_e$$

Another clue to distinguish the Majorana particle from the Dirac particle is the fact that the former cannot have a magnetic moment. This is because by charge

conjugation the magnetic moment changes its sign but the particle itself remains as it is. However, if there is more than one kind of Majorana particle they can transform into each other through magnetic interaction. This feature is sometimes applied to solar or cosmic magnetic field phenomena to test the neutrino property. A charged Dirac particle can have a magnetic moment, as already described for the case of the electron. A neutral Dirac particle can also have a magnetic moment through high-order intermediate states and also an electric dipole moment if time reversal invariance is violated.¹⁰⁾ According to the Standard Model calculation, the magnetic moment of the Dirac neutrino is [151]

$$\mu_\nu \sim \frac{3eG_F}{8\sqrt{2}\pi^2} m_\nu \sim 3 \times 10^{-19} \left(\frac{m_\nu}{1 \text{ eV}} \right) \mu_{\text{Bohr}}, \quad \mu_{\text{Bohr}} = \frac{e}{2m_e} \quad (4.127)$$

which is far below the observable limit with the present technology. However, if a new interaction exists which induces a large magnetic moment, it may be exposed by improved experiments.

10) We will defer detailed discussion of time reversal to Chap. 9.

5

Field Quantization

Before we quantize fields, we need to understand the Lagrangian formalism and derivation of invariants when the Lagrangian respects a certain symmetry. Important invariants are the energy-momentum and the electric charge, whose underlying symmetries are translational invariance in space-time and gauge invariance.

5.1

Action Principle

5.1.1

Equations of Motion

In solving a particle motion in classical mechanics, we usually start from Newton's equation of motion, where the path of the particle motion is traced by successively connecting infinitesimal changes at each point. There is another approach based on the action principle, where all the possible paths between two points are investigated and the one with the minimum action is selected. The two approaches are equivalent in classical mechanics. In quantum mechanics, however, because of the wave nature of a particle, the path chosen by the least action is the most probable, but other paths are possible, too. The deviation from the most probable path is very small, of the order determined by Planck's constant, but it is quite sizable in the microcosm extending from molecules and atoms down to elementary particles, i.e. in the world where quantum mechanics applies. A good example is the double slit interference phenomenon of an electron beam. Therefore, in our study of elementary particles, a logical choice is to start from the action principle.

In the action principle, a particle's path from a point $q_1 = q(t_1)$ to $q_2 = q(t_2)$ is such that it minimizes the "action S " defined by an integral of the Lagrangian L

$$S = \int_{t_1}^{t_2} L(q, \dot{q}) dt \quad (5.1a)$$

$$L = K - V = \frac{1}{2} m \dot{q}^2 - V(q) \quad (5.1b)$$

where K is the kinetic energy of the particle and V is some potential in which it moves. The action is a functional of both the position q and the velocity $\dot{q}$ of the particle, which, in turn, are functions of time t . Fixing the positions q_1 and q_2 , we change the path $q(t)$ by an infinitesimal amount $\delta q = \delta q(t)$, then the small change in S (variation δS) should vanish if the action is at the minimum:

$$\begin{aligned}\delta S &= \int_{t_1}^{t_2} dt \delta L = \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right] \\ &= \left[\frac{\partial L}{\partial \dot{q}} \delta q \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} \right] \delta q = 0\end{aligned}\quad (5.2)$$

We have used $\delta \dot{q} = \delta(dq/dt) = (d\delta q/dt)$. Since $\delta q(t_1) = \delta q(t_2) = 0$ and δq is arbitrary

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0 \quad (5.3)$$

which is called the Euler–Lagrange equation. One advantage of the Lagrangian formalism is that the equation holds for a general coordinate system. Applying it to a harmonic oscillator where $V(q) = m\omega^2 q^2/2$, the Lagrangian becomes

$$L = \frac{1}{2} m \dot{q}^2 - \frac{1}{2} m \omega^2 q^2 \quad (5.4)$$

The Euler–Lagrange equation gives

$$\ddot{q} = -\omega^2 q \quad (5.5)$$

5.1.2

Hamiltonian Formalism

The Hamiltonian formalism is another way to derive equations of motion equivalent to the action principle and is particularly useful in deriving quantum mechanical equations of motion. In mechanics of a point particle, the Lagrangian uses q and $\dot{q}$, while the Hamiltonian formalism chooses to use q and its canonical conjugate momentum defined by

$$p = \frac{\partial L}{\partial \dot{q}} \quad (5.6)$$

as two independent variables. q and p are called canonical variables. Then

$$\delta L = \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} = \frac{\partial L}{\partial q} \delta q + p \delta \dot{q} = \frac{\partial L}{\partial q} \delta q + \delta(p \dot{q}) - \dot{q} \delta p \quad (5.7)$$

If we define the Hamiltonian by

$$H \equiv p \dot{q} - L \quad (5.8)$$

then

$$\delta H = \delta(p\dot{q} - L) = -\frac{\partial L}{\partial q}\delta q + \dot{q}\delta p = -\dot{p}\delta q + \dot{q}\delta p \quad (5.9)$$

We used the Euler–Lagrange equation in deriving the last equation. From the above relation, it follows that

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q} \quad (5.10)$$

Equations (5.9) and (5.10) mean the Hamiltonian is a constant of motion if all the time dependences are in p and q , namely if it does not have explicit time dependence. For a harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 \quad (5.11)$$

$$\dot{q} = \frac{p}{m}, \quad \dot{p} = -m\omega^2 q \quad \Rightarrow \quad \ddot{q} = -\omega^2 q \quad (5.12)$$

Eq. (5.12) reproduces the equation of motion for the harmonic oscillator Eq. (5.4).

5.1.3

Equation of a Field

Let us extend the action to a system of N^3 particles oscillating independently that are distributed uniformly in a large but finite volume $V = L^3$, each oscillator occupying a volume $\Delta V = (L/N)^3$. This is expressed as

$$S = \int dt \sum_{r=1}^{N^3} L(q_r, \dot{q}_r) = \int dt \sum_r^{N^3} \left(\frac{1}{2} m_r \dot{q}_r^2 - \frac{1}{2} m_r \omega_r^2 q_r^2 \right) \quad (5.13)$$

Taking N^3 independent variations for q_r , we obtain N^3 Euler–Lagrange equations:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_r} \right) - \frac{\partial L}{\partial q_r} = 0 \quad (5.14)$$

In the continuous limit where the spacing L/N goes to zero, the collection of oscillators is called a “field”. Instead of numbering by r , we may use the space coordinate $\mathbf{x} = (x, y, z)$ to number the oscillators. The nomenclature “field” is by definition some quantity defined at all space positions. In this book we restrict its usage to mean an infinite collection of oscillators distributed over all space. To make a connection between the field and discrete oscillators, the field $\phi(t\mathbf{x})$ may be represented by the average value over ΔV at r -th point, and associate q_r with $\phi(\mathbf{x})$ scaled with the volume ΔV :

$$\begin{array}{ll} \text{many-particle system} & \rightarrow \text{field} \\ q_r & \rightarrow \phi(\mathbf{x}) \\ \text{numbering by } r & \rightarrow \text{numbering by } \mathbf{x} \\ \sum_{r=1}^{N^3} \Delta V & \rightarrow \int d^3x \end{array} \quad (5.15)$$

$\phi(\mathbf{x})$ need not necessarily be space oscillators. It may represent any oscillating physical object, such as the electromagnetic field. Therefore the action is represented as

$$S = \int dt \left[\int d^3x \mathcal{L}(\phi(\mathbf{x}), \dot{\phi}(\mathbf{x})) \right] \equiv \int d^4x \mathcal{L}(\phi(x), \partial_t \phi(x)) \quad (5.16)$$

In the last equation x is used to denote four space-time coordinates instead of $\mathbf{x}$ as ϕ is a function of time, too. $\mathcal{L}$ is called Lagrangian density. In field theory, the action appears as a 4-integral of the Lagrangian density, $S = \int dt L = \int d^4x \mathcal{L}$; we may call it simply the Lagrangian where there is no confusion. The field is an infinite number of some physical objects oscillating independently at position $\mathbf{x}$. If we want a field whose oscillation propagates in space, namely a wave, each oscillator has to have some connections with its neighbors. Remember the mechanical model of the field considered in Sect. 2.2.5. Receiving force from or exerting force on the neighboring oscillators was the cause of wave propagation. In the continuous limit, that means the oscillator is a function of derivatives of the space coordinate and Eq. (5.16) should include a space differential:

$$\mathcal{L}(\phi(x), \dot{\phi}(x)) \longrightarrow \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \quad \partial_\mu = (\partial_0, \nabla) \quad (5.17)$$

where ∂_t is replaced by $\partial_0 = (1/c)(\partial/\partial t)$ to make the dimension of the four-differential equal. Taking an infinitesimal variation of the action

$$\delta S = \int d^4x \left[\frac{\delta \mathcal{L}}{\delta \phi} \delta \phi + \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \delta (\partial_\mu \phi) \right] \quad (5.18)$$

and partial integration of the second term gives surface integrals, including $\delta \phi(x)$. In a mechanical system with finite degrees of freedom, the boundary condition was $\delta q_r(t_1) = \delta q_r(t_2) = 0$. For the field we require that $\delta \phi(x) = 0$ also at some boundary point, which is usually chosen at infinity. Then the surface integral vanishes and

$$\begin{aligned} \delta S &= - \int d^4x \left[\partial_\mu \left\{ \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \right\} - \frac{\delta \mathcal{L}}{\delta \phi} \right] \delta \phi(x) = 0 \\ \therefore \quad \partial_\mu \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} - \frac{\delta \mathcal{L}}{\delta \phi} &= 0 \end{aligned} \quad (5.19)$$

This is the Euler–Lagrange equation for the field. An obvious advantage of the Lagrangian formalism is that the Lagrangian density is a Lorentz scalar and easy to handle. The symmetry structure is also easy to see. The Lagrangian density is chosen to give known equations of the field or sometimes given a priori from some symmetry principle. Those to produce the real or complex Klein–Gordon equations, the Maxwell equations of the electromagnetic field, and the Dirac field are given by

$$\mathcal{L}_{\text{KG},\text{r}} = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2) \quad \text{real field} \quad (5.20\text{a})$$

$$\mathcal{L}_{\text{KG},\text{c}} = \partial_\mu \varphi^\dagger \partial^\mu \varphi - m^2 \varphi^\dagger \varphi \quad \text{complex field} \quad (5.20\text{b})$$

$$\mathcal{L}_{\text{em}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2) \quad \text{electromagnetic field} \quad (5.20\text{c})$$

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad \text{Dirac field}^{1)} \quad (5.20\text{d})$$

Problem 5.1

Using the Euler–Lagrange equations, derive the equations of motion from the Lagrangian Eq. (5.20).

Massive Vector Field

For comparison, we give here the Lagrangian of a massive vector field to realize some special features of the massless electromagnetic field. The Lagrangian of a massive vector field is given by analogy to the electromagnetic Lagrangian by

$$\mathcal{L}_{\text{em}} = -\frac{1}{4}f_{\mu\nu}f^{\mu\nu} + \frac{1}{2}m^2 W_\mu W^\mu - g W_\mu j^\mu \quad (5.22\text{a})$$

$$f^{\mu\nu} = \partial^\mu W^\nu - \partial^\nu W^\mu \quad (5.22\text{b})$$

The equation of motion is given by

$$\partial_\mu f^{\mu\nu} + m^2 W^\nu = g j^\nu \quad (5.23)$$

which is called the Proca equation. If j^μ is a conserved current, then by taking the divergence of Eq. (5.23), we can derive an equation

$$\partial_\mu W^\mu = 0 \quad (5.24)$$

Then Eq. (5.23) is equivalent to the following two equations:

$$(\partial_\mu \partial^\mu + m^2) W^\nu = g j^\nu, \quad \partial_\mu W^\mu = 0 \quad (5.25)$$

Equation (5.25) shows that the number of independent degrees of freedom of the Proca field is three, and when quantized, it represents a particle with spin 1. In the case of the electromagnetic field (or $m = 0$ vector field in general), unlike the massive field, the condition Eq. (5.24) does not come out automatically and has to be imposed as a gauge condition.

1) Sometimes a symmetrized version is used:

$$\mathcal{L}_{\text{Dirac}} = \frac{1}{2} [(-i\partial_\mu \bar{\psi} \gamma^\mu \psi - m \bar{\psi} \psi) + \bar{\psi} (\gamma^\mu i\partial_\mu - m) \psi] \quad (5.21)$$

It differs from Eq. (5.20d) by a surface integral.

Hamiltonian Formalism for Fields

The conjugate momentum field is defined just like particle momentum Eq. (5.6):

$$\pi(x) = \frac{\delta \mathcal{L}}{\delta \dot{\phi}} \quad (5.26)$$

Likewise, the Hamiltonian density for the field can be obtained using the same formula as Eq. (5.8)

$$\mathcal{H}(x) = \pi(x)\dot{\phi}(x) - \mathcal{L} \quad (5.27)$$

The Hamiltonian is a constant of motion as is the case for point particle mechanics, i.e.

$$\frac{dH}{dt} = \frac{d}{dt} \int d^3x \mathcal{H} = 0 \quad (5.28)$$

Proof:

$$\dot{\mathcal{H}} = \frac{d}{dt}(\pi\dot{\phi} - \mathcal{L}) = \dot{\pi}\dot{\phi} + \pi\ddot{\phi} - \left(\frac{\partial \mathcal{L}}{\partial t} + \frac{\delta \mathcal{L}}{\delta \phi} \dot{\phi} + \frac{\delta \mathcal{L}}{\delta \dot{\phi}} \ddot{\phi} + \frac{\delta \mathcal{L}}{\delta(\partial_k \phi)} \partial_k \dot{\phi} \right) \quad (5.29a)$$

$$\dot{\pi}\dot{\phi} = \frac{d}{dt} \left(\frac{\delta \mathcal{L}}{\delta \dot{\phi}} \right) \dot{\phi} = \left[-\partial_k \left(\frac{\delta \mathcal{L}}{\delta(\partial_k \phi)} \right) + \frac{\delta \mathcal{L}}{\delta \phi} \right] \dot{\phi} \quad (5.29b)$$

where we have used the Euler–Lagrange equation to obtain the third equality in Eq. (5.29b). Partially integrating the first term gives

$$\dot{\pi}\dot{\phi} = \left[\frac{\delta \mathcal{L}}{\delta(\partial_k \phi)} \partial_k \dot{\phi} + \frac{\delta \mathcal{L}}{\delta \phi} \dot{\phi} \right] \quad (5.29c)$$

Because of this equality, the first term of Eq. (5.29a) is canceled by the second and fourth terms in large parentheses in Eq. (5.29a) and the second term by the third term in large parentheses. Then

$$\dot{\mathcal{H}} = -\frac{\partial \mathcal{L}}{\partial t} \quad (5.29d)$$

Therefore unless the $\mathcal{L}$ has an explicit dependence on t ,²⁾ the time variation of the Hamiltonian density vanishes. $\square$

The Hamiltonian formalism is convenient for quantization. But as we shall see soon, it is the 00 component of a Lorentz tensor and hence more complicated from the Lorentz covariance point of view than a simple scalar Lagrangian.

2) Usually this happens when a time-dependent external force acts on the system. Remember mechanical energy conservation under the effect of a potential.

5.1.4

Noether's Theorem

We have seen that the energy of a field is a conserved quantity. But a question arises immediately: are there any other conserved quantities? There is a convenient and systematic way of deriving conserved quantities known as Noether's theorem. It shows that there is a constant of motion for each continuous symmetry transformation that keeps the Lagrangian invariant. Then observed selection rules can be rephrased in terms of symmetry requirements of the Lagrangian and serve as useful guides for introducing new interaction terms when developing new theories. Let us assume the Lagrangian is a function of the fields ϕ_r and $\partial_\mu \phi_r$, where r represents some internal degree of freedom, and that it has no explicit dependence on the coordinates x^μ . This is the case when the fields do not couple with an external source, i.e. if the system is closed. Then for some infinitesimal variation of the field ϕ_r , the variation of the Lagrangian is given by

$$\delta \mathcal{L} = \frac{\delta \mathcal{L}}{\delta \phi_r} \delta \phi_r + \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi_r)} \delta (\partial_\mu \phi_r) \quad (5.30)$$

As $\delta (\partial_\mu \phi_r) = \partial_\mu \delta \phi_r$, the above equation can be rewritten as

$$\delta \mathcal{L} = \left[\frac{\delta \mathcal{L}}{\delta \phi_r} - \partial_\mu \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi_r)} \right] \delta \phi_r + \partial_\mu \left\{ \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi_r)} \delta \phi_r \right\} \quad (5.31)$$

The first term vanishes because of the Euler–Lagrange equation. So if the Lagrangian is invariant for some continuous transformation of ϕ_r , the second term must vanish, which defines a conserved current J^μ . This is called the Noether current and satisfies the continuity equation

$$J^\mu = \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi_r)} \delta \phi_r, \quad \partial_\mu J^\mu = 0 \quad (5.32)$$

The meaning of the conserved current is that the space integral of the 0th component does not change with time, i.e.

$$\frac{dQ}{dt} = 0, \quad Q = \int d^3x J^0 \quad (5.33)$$

Proof:

$$\frac{dQ}{dt} = \int d^3x \partial_t J^0 = - \int d^3x \nabla \cdot \mathbf{J} = - \int_S d\sigma J_n = 0 \quad (5.34)$$

where in going to the fourth equality we have used Gauss's theorem to change the volume integral to a surface integral, and the last equality follows from the boundary condition that the fields vanish at infinity. $\square$

Conserved Charge

As an example, let us derive a charged current for a complex field. The Lagrangian given by Eq. (5.20b)

$$\mathcal{L}_{\text{KG}} = \partial_\mu \varphi^\dagger \partial^\mu \varphi - m^2 \varphi^\dagger \varphi \quad (5.35)$$

is invariant under continuous phase transformation

$$\varphi \rightarrow e^{-iq\alpha} \varphi, \quad \varphi^\dagger \rightarrow \varphi^\dagger e^{iq\alpha} \quad (5.36)$$

where α is a continuous variable but constant in the sense it does not depend on x . The phase transformation has nothing to do with the space-time coordinates and is an example of internal symmetry, referred to as $U(1)$ symmetry. Then taking

$$\delta \phi_1 = \delta \varphi = -iq \delta \alpha \varphi, \quad \delta \phi_2 = \delta \varphi^\dagger = iq \delta \alpha \varphi^\dagger \quad (5.37)$$

the Noether current is expressed as

$$J^\mu = iq [\varphi^\dagger \partial^\mu \varphi - (\partial^\mu \varphi^\dagger) \varphi] \delta \alpha \quad (5.38)$$

Removing $\delta \alpha$ we have reproduced the conserved current: [Eq. (2.41)]. The conserved charge is expressed as

$$Q = -iq \int d^3x \sum_r [\pi_r(x) \varphi_r(x) - \pi_r^\dagger(x) \varphi_r^\dagger(x)] \quad (5.39a)$$

$$= iq \int d^3x [\varphi^\dagger \dot{\varphi} - \dot{\varphi}^\dagger \varphi] \quad (5.39b)$$

The first expression is a more general form, while the second one is specific to the complex Klein–Gordon field.

Problem 5.2

The Lagrangian density of the Dirac field is made up of combinations of $\bar{\psi}(x)$, $\psi(x)$ and is invariant under the gauge transformation

$$\psi \rightarrow e^{-i\alpha} \psi, \quad \bar{\psi} \rightarrow \bar{\psi} e^{i\alpha} \quad (5.40)$$

Prove that the Noether current corresponding to infinitesimal transformations $\delta \psi = -i \delta \alpha \psi$, $\delta \bar{\psi} = i \delta \alpha \bar{\psi}$ is given by

$$j^\mu = \bar{\psi} \gamma^\mu \psi \quad (5.41)$$

Energy-Momentum Tensor

We consider here translational invariance under an infinitesimal displacement of the coordinate

$$x'^\mu = x^\mu + \delta x^\mu \quad (5.42)$$

Then

$$\delta\phi_r = \phi_r(x^\mu + \delta x^\mu) - \phi_r(x) = \partial_\mu \phi_r \delta x^\mu \quad (5.43a)$$

$$\delta(\partial_\mu \phi_r) = \partial_\mu \phi_r(x^\nu + \delta x^\nu) - \partial_\mu \phi_r(x) = \partial_\mu \partial_\nu \phi_r \delta x^\nu \quad (5.43b)$$

If the Lagrangian does not depend explicitly on x^μ , the variation can be expressed in two ways:

$$\delta\mathcal{L}_1 = \partial_\mu \mathcal{L} \delta x^\mu \quad (5.44a)$$

$$\delta\mathcal{L}_2 = \frac{\delta\mathcal{L}}{\delta\phi_r} \delta\phi_r + \frac{\delta\mathcal{L}}{\delta(\partial_\mu \phi_r)} \delta(\partial_\mu \phi_r) \quad (5.44b)$$

In calculating the action variation, unlike the internal transformation where the variation of the field is evaluated at the same space-time, we have to take into account the volume change in the integral, too, i.e.

$$\delta S = \int d^4 x' \mathcal{L}(x') - \int d^4 x \mathcal{L}(x) = \int d^4 x \delta\mathcal{L} + \int (d^4 x' - d^4 x) \mathcal{L} \quad (5.45)$$

The second term can be calculated using the Jacobian of the transformation, but here we rely on the fact that $\delta S = 0$, which means it is equal to minus the first term. Therefore,

$$\begin{aligned} \delta S &= \int d^4 x (\delta\mathcal{L}_2 - \delta\mathcal{L}_1) \\ &= \int d^4 x \left[\frac{\delta\mathcal{L}}{\delta\phi_r} \delta\phi_r + \frac{\delta\mathcal{L}}{\delta(\partial_\mu \phi_r)} \delta(\partial_\mu \phi_r) - \partial_\mu \mathcal{L} \delta x^\mu \right] \end{aligned} \quad (5.46)$$

Inserting Eq. (5.43) into the first term of Eq. (5.46), the integrand becomes

$$\left[\frac{\delta\mathcal{L}}{\delta\phi_r} - \partial_\mu \frac{\delta\mathcal{L}}{\delta(\partial_\mu \phi_r)} \right] \partial_\nu \phi_r \delta x^\nu + \partial_\mu \left\{ \frac{\delta\mathcal{L}}{\delta(\partial_\mu \phi_r)} \partial_\nu \phi_r - \delta^\mu_\nu \mathcal{L} \right\} \delta x^\nu \quad (5.47a)$$

The first term vanishes because of the Euler–Lagrange equation. Thus we have obtained a conserving tensor called the energy-momentum tensor:

$$\partial_\mu T^{\mu\nu} = 0 \quad (5.47b)$$

$$T^{\mu\nu} = \frac{\delta\mathcal{L}}{\delta(\partial_\mu \phi_r)} \partial^\nu \phi_r - g^{\mu\nu} \mathcal{L} \quad (5.47c)$$

The name is given because

$$T^{00} = \frac{\delta\mathcal{L}}{\delta(\partial_0 \phi_r)} \partial^0 \phi_r - \mathcal{L} \quad (5.48)$$

agrees with the Hamiltonian density given in Eq. (5.27). Its space integral is a constant of motion and reconfirms Eq. (5.28). Then, by Lorentz covariance, $P^i = \int d^3 x T^{0i} \equiv \int d^3 x \mathcal{P}^i$ are momenta of the field. They are also constants of motion.

Problem 5.3

Using the Jacobian prove that

$$\int (d^4 x' - d^4 x) \mathcal{L} = \int d^4 x (-\partial_\mu \mathcal{L} \delta x^\mu) \quad (5.49)$$

Problem 5.4

Calculate the energy momentum tensor of the scalar field whose Lagrangian is given by Eq. (5.20a) and show that it is symmetric with respect to suffixes μ and ν .

We can derive the angular momentum tensor by requiring the invariance under an infinitesimal Lorentz transformation given in Eq. (3.79)

$$\delta x^\mu = \epsilon^\mu_\nu x^\nu \quad (5.50a)$$

$$\delta \phi(x) = \epsilon^\mu_\nu x^\nu \partial_\mu \phi(x) \quad (5.50b)$$

Problem 5.5

(a) Show the corresponding angular momentum tensor is given by

$$M^{\rho, \mu \nu} = T^{\rho \nu} x^\mu - T^{\rho \mu} x^\nu \quad (5.51)$$

(b) Show that

$$\partial_\rho M^{\rho, \mu \nu} = T^{\mu \nu} - T^{\nu \mu} \quad (5.52)$$

Hint: Replace the second term in Eq. (5.47a) by $\delta x^\nu \rightarrow \epsilon^\mu_\nu x^\nu$ and use $\epsilon_{\nu \mu} = -\epsilon_{\mu \nu}$.

Problem 5.5 shows that $T^{\mu \nu} = T^{\nu \mu}$ is a necessary condition for the angular momentum tensor to be conserved. However, the expression for $T^{\mu \nu}$ in Eq. (5.47c) is not manifestly symmetric and indeed it is not always symmetric. We note that the energy momentum tensor itself is not a measurable quantity but its space integral is. So there is the freedom to add a total derivative. We may add a term of the form

$$\partial_\lambda f^{\lambda, \mu \nu}, \quad f^{\lambda, \mu \nu} = -f^{\mu, \lambda \nu} \quad (5.53)$$

so that

$$\partial_\lambda \partial_\mu f^{\lambda, \mu \nu} \equiv 0 \quad (5.54)$$

We consider a new energy momentum tensor

$$T^{\mu \nu} \rightarrow T^{\mu \nu} + \partial_\lambda f^{\lambda, \mu \nu} \quad (5.55)$$

This new energy momentum tensor is conserved and we can always choose this extra term to make it symmetric in μ and ν .

There is another reason to have a symmetric energy momentum tensor. In general relativity, the symmetric gravitational field or curvature tensor $R_{\mu\nu}$ couples to the energy momentum tensor through Einstein's equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -\frac{8\pi G}{c^2}T_{\mu\nu} \quad (5.56)$$

The Ricci tensor $R_{\mu\nu}$ and the metric tensor $g_{\mu\nu}$ are both symmetric and $T_{\mu\nu}$ should be symmetric, too.

The Complex Field

We now apply the energy momentum tensor formula to Eq. (5.20b) and derive the Hamiltonian and the momentum operators expressed in terms of the fields:

$$\begin{aligned} \frac{\delta \mathcal{L}}{\delta(\partial_\mu \varphi)} &= \partial^\mu \varphi^\dagger & \frac{\delta \mathcal{L}}{\delta(\partial_\mu \varphi^\dagger)} &= \partial^\mu \varphi \\ \mathcal{H} = T^{00} &= \sum_r \frac{\delta \mathcal{L}}{\delta(\partial_0 \phi_r)} \partial^0 \phi_r - \mathcal{L} \end{aligned} \quad (5.57a)$$

$$\begin{aligned} &= \partial^0 \varphi^\dagger \dot{\varphi} + \dot{\varphi}^\dagger \partial^0 \varphi - (\partial_\mu \varphi^\dagger \partial^\mu \varphi - m^2 \varphi^\dagger \varphi) \\ &= \dot{\varphi}^\dagger \dot{\varphi} + \nabla \varphi^\dagger \cdot \nabla \varphi + m^2 \varphi^\dagger \varphi \end{aligned}$$

$$\mathcal{P} = \sum_r T^{0i} = \frac{\delta \mathcal{L}}{\delta(\partial_0 \phi_r)} \partial^i \phi_r = -(\dot{\varphi}^\dagger \nabla \varphi + \nabla \varphi^\dagger \dot{\varphi}) \quad (5.57b)$$

$$H = \int d^3x \mathcal{H} = \int d^3x [\dot{\varphi}^\dagger \dot{\varphi} + \nabla \varphi^\dagger \cdot \nabla \varphi + m^2 \varphi^\dagger \varphi] \quad (5.58a)$$

$$P = \int d^3x \mathcal{P} = - \int d^3x [\dot{\varphi}^\dagger \nabla \varphi + \nabla \varphi^\dagger \dot{\varphi}] \quad (5.58b)$$

Problem 5.6

- (1) Prove that H and P are conserved quantities by explicitly differentiating them.
- (2) Prove that the Hamiltonian density $\mathcal{H}$ and the momentum density $\mathcal{P}$ satisfy the continuity equation, namely

$$\frac{\partial \mathcal{H}}{\partial t} + \nabla \cdot \mathcal{P} = 0 \quad (5.59)$$

Problem 5.7

Using the Lagrangian of the electromagnetic field, Eq. (5.20c), show

$$H = \int d^3x \frac{1}{2} (E^2 + B^2) \quad (5.60a)$$

$$P = \int d^3x \mathbf{E} \times \mathbf{B} \quad (5.60b)$$

Problem 5.8

Using the Lagrangian of the Dirac field, Eq. (5.20d), show

$$H = \int d^3x \bar{\psi}(-i\boldsymbol{\gamma} \cdot \nabla + m)\psi = \int d^3x \psi^\dagger i\dot{\psi} \quad (5.61a)$$

$$P = \int d^3x \psi^\dagger(-i\nabla)\psi \quad (5.61b)$$

$$Q = q \int d^3x \psi^\dagger \psi \quad (5.61c)$$

5.2**Quantization Scheme**

There are several ways to make the transition from classical mechanics to quantum mechanics. Probably the most powerful method is that of the path integral, which deals with the action, taking all possible paths into account. Historically, however, the Hamiltonian formalism was used first and has been extensively used since then. The path integral method is a powerful tool for treating the formal aspects of the theory, and indispensable in discussing non-Abelian gauge theories, but if one is interested in quick practice at calculating transition rates, the conventional Hamiltonian formalism is more easily accessible using the accumulated knowledge of quantum mechanics. Since the path integral uses mathematical approaches that for laypeople need some acquaintance before solving realistic problems, we shall defer the study of the path integral method until Chap. 10.

5.2.1**Heisenberg Equation of Motion**

To derive the quantum mechanical representation of the equation of motion that is suitable for the treatment of the harmonic oscillator, we start from the Schrödinger equation

$$i\frac{\partial}{\partial t}|\psi(x)\rangle = H|\psi(x)\rangle \quad (5.62)$$

Here, we have used Dirac's notation to describe state vectors in Hilbert space on which operators in quantum mechanics operate. The formal solution to the equation is expressed as

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle \quad (5.63)$$

where we have suppressed the space coordinate dependence. The Hamiltonian in the Schrödinger equation contains operators as functions of canonical variables q, p in the form $q \rightarrow \mathbf{x}$, $p \rightarrow -i\nabla$ and which are time independent. The time

dependence is in the state vector. This is called the Schrödinger picture. In the following, we attach a subscript S to quantities in the Schrödinger picture and a subscript H to the Heisenberg picture that we describe now. We define state vectors and operators in the Heisenberg picture by

$$|\psi_H\rangle = e^{iH_S t} |\psi_S\rangle = |\psi_S(0)\rangle \quad (5.64a)$$

$$O_H = e^{iH_S t} O_S e^{-iH_S t} \quad (5.64b)$$

In the Heisenberg picture, all the time dependence is in the operators and the state vectors do not change with time. From Eq. (5.64b), we can derive the time dependence of the Heisenberg operators:

$$\frac{dO_H}{dt} = i[H_S, O_H] \quad (5.65)$$

Putting $O_S = H_S$, Eq. (5.64b) shows $H_H = H_S$ and from Eq. (5.65) we see that the Hamiltonian does not change with time, i.e. it is a conserved quantity.³⁾ From now on, we omit the subscript of the Hamiltonian and simply write it as H . Then the equation that the Heisenberg operators satisfy is written as

$$\frac{dO_H}{dt} = i[H, O_H] \quad (5.66)$$

which is called the Heisenberg equation of motion. Its relativistic version can be expressed as

$$\partial_\mu O_H = i[P_\mu, O_H] \quad (5.67)$$

This is still a quantum mechanical equation. We can prove that Eq. (5.67) is also valid in quantum field theory, namely if the operators P_μ and O_H are functions of quantized fields. We shall prove this later in Sect. 9.1.1.

Problem 5.9

Derive the space part of Eq. (5.67).

We can make the transition from classical mechanics to quantum mechanics by the following steps:

- (1) Regard the classical variables
as quantum operators
 - (2) Impose the commutation relation $[q_i, p_j] = i\delta_{ij}$
 - (3) Then calculate $\partial_\mu O_H = i[P_\mu, O_H]$
- (5.68)

The advantage of the Heisenberg picture is that it makes the correspondence with classical mechanics apparent.

3) See Eq. (5.29d) for a caveat.

5.2.2

Quantization of the Harmonic Oscillator

We start with the quantization of the harmonic oscillator, a prescription described in any standard textbook of quantum mechanics. The Hamiltonian is given by Eq. (5.11). Using the Heisenberg equation of motion

$$\dot{q} = i[H, q] = \frac{p}{m}, \quad \dot{p} = i[H, p] = -m\omega^2 q \Rightarrow \ddot{q} = -\omega^2 q \quad (5.69)$$

which is the same equation as Eq. (5.12), but q and p are now operators. Now we introduce operators defined by

$$q = \sqrt{\frac{1}{2m\omega}}(a + a^\dagger), \quad p = -i\sqrt{\frac{m\omega}{2}}(a - a^\dagger) \quad (5.70)$$

The canonical commutation relation (5.68) becomes

$$[a, a^\dagger] = 1 \quad (5.71)$$

The Hamiltonian in terms of $a, a^\dagger$ is given as

$$H = \frac{1}{2}\omega(a^\dagger a + a a^\dagger) = \omega\left(a^\dagger a + \frac{1}{2}\right) \quad (5.72)$$

Inserting the new variables into the Heisenberg equation gives

$$\dot{a} = -i\omega a, \quad \dot{a}^\dagger = i\omega a^\dagger \quad (5.73a)$$

$$\therefore a = a_0 e^{-i\omega t}, \quad a^\dagger = a_0^\dagger e^{i\omega t} \quad (5.73b)$$

We rename a_0 and $a_0^\dagger$ as a and $a^\dagger$ and consider their properties. As will be shown in the following, eigenvalues of the operator $N \equiv a^\dagger a$ are either zero or positive integers. Hence the energy intervals of the harmonic oscillator are all the same ($\Delta E = \hbar\omega$).

Problem 5.10

Using the Hamiltonian (5.72) and the commutator (5.71), derive Eq. (5.73a).

The Number Operator

We prove that the operator $N = a^\dagger a$ has positive integer eigenvalues and can be considered as the number operator. It satisfies the commutation relations

$$[N, a] = -a, \quad [N, a^\dagger] = a^\dagger \quad (5.74)$$

Let us consider an eigenstate of the operator N with eigenvalue n :

$$N|n\rangle = n|n\rangle \quad (5.75)$$

Using the commutator (5.74), we see

$$N a^\dagger |n\rangle = a^\dagger (1 + N) |n\rangle = (n + 1) a^\dagger |n\rangle \quad (5.76a)$$

$$N a |n\rangle = (N - 1) a |n\rangle = (n - 1) a |n\rangle \quad (5.76b)$$

which means that $a^\dagger |n\rangle, a |n\rangle$ are also eigenstates of N with eigenvalues $(n \pm 1)$. Therefore, $a, a^\dagger$ are operators that change the eigenvalue by $+1$ and -1 , respectively. Set the normalization condition as

$$\langle n | n \rangle = 1, \quad a |n\rangle = c_n |n - 1\rangle \quad (5.77a)$$

$$n = \langle n | N | n \rangle = \langle n | a^\dagger a | n \rangle = |c_n|^2 \langle n - 1 | n - 1 \rangle = |c_n|^2 \rightarrow c_n = \sqrt{n} \quad (5.77b)$$

We have chosen the phase of c_n to make it real. Equation (5.76) means that the expectation values of $a^\dagger, a$ are nonzero only between $|n\rangle$ and $|n \pm 1\rangle$:

$$|n\rangle = \frac{a^\dagger |n - 1\rangle}{c_n} = \frac{(a^\dagger)^2 |n - 2\rangle}{c_n c_{n-1}} = \dots = \frac{(a^\dagger)^m |n - m\rangle}{c_n c_{n-1} \dots c_{n-m+1}} \quad (5.78)$$

Repeated operation of a on $|n\rangle$ eventually makes the eigenvalue negative. Because of Eq. (5.77b), the state cannot have negative eigenvalues. To avoid contradiction, there must exist a state $|0\rangle$ that satisfies

$$a |0\rangle = 0, \quad N |0\rangle = 0 \quad (5.79)$$

We call it the ground state or the vacuum state. Equation (5.78) tells us that any state can be constructed from $|0\rangle$ by multiplying by $a^\dagger$:

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle, \quad N |n\rangle = n |n\rangle \quad (5.80a)$$

$$H |n\rangle = \omega \left(a^\dagger a + \frac{1}{2} \right) |n\rangle = \omega \left(n + \frac{1}{2} \right) |n\rangle \quad (5.80b)$$

The value of the energy is given by a multiple of a positive integer times ω plus some fixed constant.

The fact that all the energy intervals of a harmonic oscillator are equal becomes important when we quantize the field. At this point we make an important re-interpretation. We do not consider that the energy is n multiples of ω but that n particles with energy ω are created. Then $N = a^\dagger a$ represents a number operator of the particle and the Hamiltonian $H = \omega a^\dagger a$ represents the energy sum of n particles. a and $a^\dagger$ can be interpreted as the annihilation and creation operators of the particle. The constant term $\omega/2$ in the Hamiltonian is called the zero-point energy. For the moment we ignore it because the reference energy can always be shifted and hence unobservable. We shall come back to it in the last section of this chapter.

Quantization of the Fermion

We have derived the equation for a harmonic oscillator in two ways: A using q, p , and B using $a, a^\dagger$. Method A uses Eq. (5.11) as the Hamiltonian and the commutator Eq. (5.68) while method B uses Eq. (5.72) and Eq. (5.71). They are connected by the relation Eq. (5.70). Whichever we choose as the starting point, we can derive the other. Therefore, both methods are equivalent. However, we notice that if we choose B as the starting point, the commutator Eq. (5.71) is not a unique choice. If, instead of the commutator, we require anticommutator relations, we have

$$[a, a^\dagger]_+ \equiv \{a, a^\dagger\} \equiv aa^\dagger + a^\dagger a = 1 \quad (5.81a)$$

$$\{a, a\} = \{a^\dagger, a^\dagger\} = 0 \quad (5.81b)$$

We can equally obtain the harmonic oscillator using the Heisenberg equation.

Problem 5.11

Using the Hamiltonian Eq. (5.72) and the anticommutator Eq. (5.81), prove that the Heisenberg equation of motion gives the same harmonic oscillator equation Eq. (5.73a) for $a, a^\dagger$.

In this case, however, the eigenvalues of the number operator are limited only to 1, 0:

$$N^2 = a^\dagger a a^\dagger a = a^\dagger (1 - a^\dagger a) a = a^\dagger a = N \quad (5.82a)$$

$$\therefore N = 1, 0 \quad (5.82b)$$

The eigenvectors are $|1\rangle$ and $|0\rangle$. Besides,

$$N a^\dagger |0\rangle = a^\dagger a a^\dagger |0\rangle = a^\dagger (1 - a^\dagger a) |0\rangle = a^\dagger |0\rangle \quad (5.83a)$$

$$\therefore a^\dagger |0\rangle = |1\rangle \quad (5.83b)$$

Similarly,

$$a|1\rangle = |0\rangle, \quad a|0\rangle = 0, \quad a^\dagger|1\rangle = 0 \quad (5.83c)$$

The last equation shows that $a^\dagger$ behaves like an annihilation operator acting on $|1\rangle$, because there is no $n = 2$ state. In summary, $a, a^\dagger$ can be interpreted as the annihilation or creation operator of a particle; only one particle can occupy a state. Such a particle is called a fermion. Expression of the Hamiltonian in terms of q, p using the relation Eq. (5.70) gives

$$H = i\omega qp \quad (5.84)$$

which has no simple interpretation in classical mechanics and is, therefore, a purely quantum mechanical object. Expressed in terms of $a, a^\dagger$

$$H = \omega \left(a^\dagger a - \frac{1}{2} \right) \quad (5.85)$$

which has a logical physical interpretation.

Particles that satisfy the commutation relation are said to obey Bose–Einstein statistics and are called bosons, while those that satisfy the anticommutation relation obey Fermi–Dirac statistics. Any number of bosons can occupy a state, but only one fermion can enter a state.

5.3

Quantization of Fields

How can we quantize fields? In Eq. (5.15), we have seen that the field is a collection of oscillators q_r distributed over all space $\mathbf{x}$, and that $\mathbf{x}$ has been chosen for numbering instead of r . To make the transition from discrete r to continuous $\mathbf{x}$, let us divide the space into rectangular boxes of volume $\Delta V = \Delta x \Delta y \Delta z$ and number them by r . Define q_r as the field averaged over ΔV at point $\mathbf{x} = \mathbf{x}_r$ multiplied by ΔV , i.e. $q_r(t) = \phi(t, \mathbf{x}) \Delta V$. Then the field Lagrangian can be rewritten in terms of q_r :

$$\begin{aligned} L &= \frac{1}{2} \int d^3x [\dot{\phi}^2 - (\nabla \phi)^2 + m^2 \phi^2] = \int d^3x \mathcal{L} = \sum \Delta V \mathcal{L}_r \\ &= \frac{1}{2\Delta V} \sum_r \dot{q}_r^2 - (\text{terms independent of } \dot{q}_r) \end{aligned} \quad (5.86)$$

We define the conjugate momentum by

$$p_r(t) = \frac{\partial L}{\partial \dot{q}_r} = \frac{\dot{q}_r}{\Delta V} = \frac{\partial(\sum_r \Delta V \mathcal{L}_r)}{\partial(\dot{\phi}(\mathbf{x}_r) \Delta V)} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\mathbf{x}_r)} = \dot{\phi}(\mathbf{x}_r) \quad (5.87)$$

Then the Hamiltonian is given by

$$H = \sum p_r \dot{q}_r - L = \sum \pi(\mathbf{x}_r, t) \dot{\phi}(\mathbf{x}_r) \Delta V - L \xrightarrow{\Delta V \rightarrow 0} \int d^3x \mathcal{H} \quad (5.88a)$$

$$\mathcal{H} = \pi(\mathbf{x}, t) \dot{\phi}(t, \mathbf{x}) - \frac{1}{2} [\dot{\phi}^2 - (\nabla \phi)^2 + m^2] \quad (5.88b)$$

which reproduces the field Hamiltonian in the continuous limit. Since we have replaced the field by a collection of discrete oscillators, the quantization can be performed in a quantum mechanical way:

$$[q_r, p_s] = i\delta_{rs}, \quad [q_r, q_s] = [p_r, p_s] = 0 \quad (5.89)$$

The equal time commutator of the field can be obtained by replacing q_r, p_r by $\phi(t, \mathbf{x}_r) \Delta V, \pi(t, \mathbf{x}_r)$

$$[\phi(t, \mathbf{x}_r), \pi(t, \mathbf{x}_s)] = i \frac{\delta_{rs}}{\Delta V} \xrightarrow{\Delta V \rightarrow 0} [\phi(t, \mathbf{x}), \pi(t, \mathbf{y})] = i \delta^3(\mathbf{x} - \mathbf{y}) \quad (5.90)$$

In summary, it is reasonable to assume that the quantization of the fields can be achieved by

1. interpreting the classical fields as quantum field operators in the Heisenberg picture, and
2. imposing equal time commutation relations of the fields and their conjugate momentum fields

$$[\phi_s(x), \pi_t(y)]_{t_x=t_y} = i \delta_{st} \delta^3(\mathbf{x} - \mathbf{y}) \quad (5.91a)$$

$$[\phi_s(x), \phi_t(y)]_{t_x=t_y} = [\pi_s(x), \pi_t(y)]_{t_x=t_y} = 0 \quad (5.91b)$$

where the indices s, t specify additional internal degrees of freedom. We have used x to denote the four-dimensional space-time coordinate, i.e. $x = (t, \mathbf{x})$. The conjugate momentum field is defined by

$$\pi_s(x) = \frac{\delta \mathcal{L}}{\delta \dot{\phi}_s} \quad (5.91c)$$

In this book, however, we shall adopt an alternative method to quantize plane waves with specific wave number $\mathbf{k}$ and frequency ω . Because any wave field can be (Fourier) decomposed in terms of waves that are harmonic oscillators, the method we learned in quantum mechanics is directly applicable. Both methods are equivalent, but the latter probably makes it easier to grasp the physical picture behind the mathematical formula, although once one is used to the method, the former may be easier to handle.

5.3.1

Complex Fields

We start from a spin 0 complex scalar field that obeys the Klein–Gordon equation. It has two internal degrees of freedom and can be expressed in an alternative form:

$$\varphi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), \quad \varphi^\dagger = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2) \quad (5.92)$$

Here, we treat φ and $\varphi^\dagger$ as two independent fields. The Lagrangian is already given by Eq. (5.20b) and the equation of motion can be derived from the Euler–Lagrange equation. The energy, momentum and other conserved physical quantities were

given in [Eq. (5.58)].⁴⁾

$$H = \int d^3x \left[\left(\frac{\partial \varphi^\dagger}{\partial t} \right) \left(\frac{\partial \varphi}{\partial t} \right) + (\nabla \varphi^\dagger \cdot \nabla \varphi) + m^2 \varphi^\dagger \varphi \right] \quad (5.93a)$$

$$\mathbf{P} = - \int d^3x \left[\frac{\partial \varphi^\dagger}{\partial t} \nabla \varphi + \nabla \varphi^\dagger \frac{\partial \varphi}{\partial t} \right] \quad (5.93b)$$

$$Q = iq \int d^3x \left[\varphi^\dagger \frac{\partial \varphi}{\partial t} - \frac{\partial \varphi^\dagger}{\partial t} \varphi \right] \quad (5.93c)$$

Since the Klein–Gordon equation is a linear wave equation, it can be expanded in terms of waves having wave number $\mathbf{k}$

$$\varphi(x) = \sum_{\mathbf{k}} q_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{x}} \xrightarrow{\text{continuous limit}} V \int \frac{d^3p}{(2\pi)^3} q_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{x}} \quad (5.94)$$

The continuous limit is written for later reference, but in this chapter we stick to discrete sum to aid intuitive understanding of the quantization procedure. Applying the Klein–Gordon equation

$$[\partial_\mu \partial^\mu + m^2] \varphi = \sum_{\mathbf{k}} (\ddot{q}_{\mathbf{k}} + \omega^2 q_{\mathbf{k}}) e^{i\mathbf{k} \cdot \mathbf{x}} = 0, \quad \omega = \sqrt{\mathbf{k}^2 + m^2} \quad (5.95)$$

which shows that the Klein–Gordon field is a collection of an infinite number of harmonic oscillators indexed by the wave number $\mathbf{k}$. Those with different $\mathbf{k}$ are in different oscillation modes, and hence independent of each other. We separate $q_{\mathbf{k}}$ into positive and negative frequency parts⁶⁾ and write

$$\varphi(x) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega V}} (a_{\mathbf{k}} e^{-i\omega t} + c_{\mathbf{k}} e^{i\omega t}) e^{i\mathbf{k} \cdot \mathbf{x}} \quad (5.96)$$

where V is the normalization volume for the plane wave. Inserting Eq. (5.96) into Eq. (5.93), and using the orthogonal relation (refer to Eq. 6.73)

$$\frac{1}{V} \int d^3x e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{x}} = \delta_{\mathbf{k}\mathbf{k}'} \quad (5.97)$$

4) The wording “electric charge” is used here for simplicity, but the argument applies to any conserved charge-like quantum numbers, for example, strangeness.

5) Later we will set $V = 1$, but we retain the integration volume V for a while. When $V = L^3$ is finite, we should use a discrete

sum, but even in the continuous limit (i.e. $V \rightarrow \infty$), it is sometimes convenient to think V is large but finite. See Sect. 6.6.2 for detailed discussion.

6) By convention we call $e^{-i\omega t}$ positive frequency and $e^{i\omega t}$ negative frequency.

we obtain

$$H = \sum_k \omega [a_k^\dagger a_k + c_k^\dagger c_k] \quad (5.98a)$$

$$P = \sum_k k [a_k^\dagger a_k - c_k^\dagger c_k] \quad (5.98b)$$

$$Q = \sum_k q [a_k^\dagger a_k - c_k^\dagger c_k] \quad (5.98c)$$

Then we impose

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}, \quad [a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0 \quad (5.99a)$$

$$[c_k, c_{k'}^\dagger] = \delta_{kk'}, \quad [c_k, a_{k'}] = [c_k^\dagger, c_{k'}^\dagger] = 0 \quad (5.99b)$$

$$[a_k, c_{k'}^\dagger] = [a_k^\dagger, c_{k'}] = [a_k, c_{k'}] = [a_k^\dagger, c_{k'}^\dagger] = 0 \quad (5.99c)$$

Here, $(a_k^\dagger, a_k)$ can be interpreted as creation and annihilation operators of a particle with energy-momentum $(E = \hbar\omega, \mathbf{p} = \hbar\mathbf{k})$. However, if we apply the same rule to the negative frequency part, $(c_k^\dagger, c_k)$ creates or annihilates a particle with negative energy. This is not allowed since there should be no particles with negative energy. Let us examine more carefully what effects $c_k^\dagger$ and c_k produce when operating on the state vector. As before, we define the number operator N by

$$N = \sum_k (a_k^\dagger a_k + c_k^\dagger c_k) \quad (5.100)$$

Using the commutation relation Eq. (5.99), we obtain the relations

$$[N, a_k^\dagger] = a_k^\dagger, \quad [N, a_k] = -a_k \quad (5.101a)$$

$$[N, c_k^\dagger] = c_k^\dagger, \quad [N, c_k] = -c_k \quad (5.101b)$$

which shows that $a_k^\dagger, c_k^\dagger$ can be interpreted as creation and a_k, c_k as annihilation operators of particles with momentum $\mathbf{k}$. Now we apply the Heisenberg equation to the field φ and have it operate on a energy eigenstate $|E\rangle$

$$\dot{\varphi}|E\rangle = i[H, \varphi]|E\rangle, \quad H|E\rangle = E|E\rangle \quad (5.102)$$

Compare the coefficients of $e^{-i\omega t}$ and $e^{+i\omega t}$ on the two sides of the first equation. Using the relation (5.101), we obtain

$$\begin{aligned} -i\omega a_k|E\rangle &= i[H, a_k]|E\rangle = i(H - E)a_k|E\rangle \\ i\omega c_k|E\rangle &= i[H, c_k]|E\rangle = i(H - E)c_k|E\rangle \\ \therefore H a_k|E\rangle &= (E - \omega)a_k|E\rangle, \end{aligned} \quad (5.103a)$$

$$H c_k|E\rangle = (E + \omega)c_k|E\rangle \quad (5.103b)$$

a_k , the coefficient of the positive-frequency waves, reduces the number of particles by 1 and also reduces the energy of the state $|E\rangle$ by ω . So it is a legitimate annihilation operator. But c_k , the coefficient of the negative-frequency wave term, actually increases the energy of the state by ω . A logical interpretation is then that c_k creates a particle and should be considered as a creation operator. We then redefine $c_{-k} \rightarrow b_k^\dagger$, $c_{-k}^\dagger \rightarrow b_k$ and the commutation relations as

$$b_k \equiv c_{-k}^\dagger, \quad b_k^\dagger \equiv c_{-k} \quad (5.104a)$$

$$[b_k, b_{k'}^\dagger] = \delta_{kk'}, \quad [b_k, b_{k'}] = [b_k^\dagger, b_{k'}^\dagger] = 0 \quad (5.104b)$$

$$[a_k, b_{k'}^\dagger] = [a_k, b_{k'}] = [a_k^\dagger, b_{k'}^\dagger] = 0 \quad (5.104c)$$

The reason for adopting $-k$ is to make the second term in Eq. (5.98b) positive so that the particle has positive momentum as well as positive energy. We obtain the newly defined field and Hamiltonian, etc. as

$$\begin{aligned} \varphi(x) &= \sum_k \frac{1}{\sqrt{2\omega V}} (a_k e^{-i\omega t} + c_k e^{i\omega t}) e^{ik \cdot x} \\ &= \sum_k \frac{1}{\sqrt{2\omega V}} (a_k e^{-i\omega t + ik \cdot x} + b_k^\dagger e^{i\omega t - ik \cdot x}) \end{aligned}$$

In terms of the new creation/annihilation operators, the energy, momentum and charge operators are rewritten as

$$H = \sum_k \omega [a_k^\dagger a_k + b_k^\dagger b_k + 1] \quad (5.105a)$$

$$P = \sum_k k [a_k^\dagger a_k + b_k^\dagger b_k + 1] \quad (5.105b)$$

$$Q = \sum_k q [a_k^\dagger a_k - b_k^\dagger b_k - 1] \quad (5.105c)$$

The constant term appears because $c_{-k}^\dagger c_{-k} = b_k b_k^\dagger = 1 + b_k^\dagger b_k$. The above expression tells us that b_k is an annihilation operator of a particle having the energy-momentum $(\omega, \mathbf{k})$, but opposite charge, which we call an antiparticle. Note, we often encounter a statement that the antiparticle has negative energy. From the above discussion, the reader should be convinced that there are no particles with negative energy. The antiparticle has positive energy, too.

Summary of Field Quantization Field Expansion

$$\varphi(x) = \sum_k \frac{1}{\sqrt{2\omega V}} (a_k e^{-i\omega t + ik \cdot x} + b_k^\dagger e^{i\omega t - ik \cdot x}) \quad (5.106)$$

Commutation Relations

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}, \quad [b_k, b_{k'}^\dagger] = \delta_{kk'} \quad (5.107a)$$

$$[a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0, \quad [b_k, b_{k'}] = [b_k^\dagger, b_{k'}^\dagger] = 0 \quad (5.107b)$$

$$[a_k, b_{k'}^\dagger] = [a_k^\dagger, b_{k'}] = 0 \quad [a_k, b_{k'}] = [a_k^\dagger, b_{k'}^\dagger] = 0 \quad (5.107c)$$

Energy, Momentum and Charge Operators

$$H = \sum_k \omega [a_k^\dagger a_k + b_k^\dagger b_k + 1] \quad (5.108a)$$

$$P = \sum_k k [a_k^\dagger a_k + b_k^\dagger b_k + 1] \quad (5.108b)$$

$$Q = \sum_k q [a_k^\dagger a_k - b_k^\dagger b_k - 1] \quad (5.108c)$$

Using relations in the box, it is a straightforward calculation to prove

$$[\varphi(x), \pi(y)]_{t_x=t_y} = i\delta^3(\mathbf{x} - \mathbf{y}) \quad (5.109a)$$

$$[\varphi(x), \varphi(y)]_{t_x=t_y} = [\pi(x), \pi(y)]_{t_x=t_y} = 0 \quad (5.109b)$$

where

$$\pi(x) \equiv \frac{\delta \mathcal{L}}{\delta(\dot{\varphi}(x))} = \dot{\varphi}^\dagger(x) \quad (5.109c)$$

Problem 5.12

Decomposing the complex field and its conjugate momentum in terms of annihilation and creation operators, prove Eq. (5.109).

Note: As the conjugate momenta of φ and $\varphi^\dagger$ are given by

$$\pi_\varphi = \frac{\delta \mathcal{L}}{\delta \dot{\varphi}} = \dot{\varphi}^\dagger, \quad \pi_{\varphi^\dagger} = \frac{\delta \mathcal{L}}{\delta \dot{\varphi}^\dagger} = \dot{\varphi} \quad (5.110)$$

the commutators for $[\varphi, \pi_\varphi]$ and $[\varphi^\dagger, \pi_{\varphi^\dagger}]$ are equivalent in the sense that one is the hermitian conjugate of the other. Therefore there are no additional commutators to Eq. (5.109).

Equation (5.106) may be inverted as

$$a_k = i \int d^3x e_k^*(x) \overleftrightarrow{\partial}_t \varphi(x), \quad a_k^\dagger = -i \int d^3x e_k(x) \overleftrightarrow{\partial}_t \varphi^\dagger(x) \quad (5.111a)$$

$$b_k = i \int d^3x e_k^*(x) \overleftrightarrow{\partial}_t \varphi^\dagger(x), \quad b_k^\dagger = -i \int d^3x e_k(x) \overleftrightarrow{\partial}_t \hat{\phi}(x) \quad (5.111b)$$

$$e_k(x) = \frac{1}{\sqrt{2\omega V}} e^{-ik \cdot x}, \quad e_k^*(x) = \frac{1}{\sqrt{2\omega V}} e^{ik \cdot x} \quad (5.111c)$$

where

$$A(t) \overleftrightarrow{\partial}_t B(t) = A(t) \frac{\partial B(t)}{\partial t} - \frac{\partial A(t)}{\partial t} B(t) \quad (5.112)$$

Quantization of the fields can also be achieved by imposing the commutation relations Eq. (5.109) on the fields and their conjugate momentum fields. Then the commutation relations for the creation and annihilation operators Eq. (5.107) can be proved by using Eq. (5.111) and the commutation relations Eq. (5.109). The two methods of quantization are mathematically equivalent and one can be derived from the other.

Problem 5.13

Starting from the field commutators Eq. (5.109) and using Eq. (5.111), prove the commutation relations for the creation and annihilation operators Eq. (5.107).

5.3.2

Real Field

The expression for the real field is obtained by setting $\varphi = \varphi^\dagger = \phi/\sqrt{2}$ and inserting it into Eq. (5.93), giving

$$H = \frac{1}{2} \int d^3x [\dot{\phi}^2 + (\nabla \phi \cdot \nabla \phi) + m^2 \phi^2] \quad (5.113a)$$

$$P = - \int d^3x \left(\frac{\partial \phi}{\partial t} \nabla \phi \right) \quad (5.113b)$$

The complex scalar field is actually two real scalar fields of the same mass added together. Inserting $\varphi = (\phi_1 + i\phi_2)/\sqrt{2}$, $\varphi^\dagger = (\phi_1 - i\phi_2)/\sqrt{2}$, the energy-momentum of φ and ϕ_1, ϕ_2 can be converted into each other. The difference is that no charge operator can be defined for the real field, which is readily seen by looking at Eq. (5.93c). Therefore the real field describes electrically neutral particles, but the converse is not true. For instance, K^0 and $\bar{K}^0$ are electrically neutral, but have the “strangeness” quantum number to distinguish them and hence are described by the complex field.

Fock Space

As can be seen from Eq. (5.106), φ connects states differing in particle number by 1. Especially, a matrix element between the vacuum and one-particle system appears frequently. Defining $|k\rangle = \sqrt{2\omega} a_k^\dagger |0\rangle$ ⁷⁾

$$a_k |0\rangle = \langle 0 | a_k^\dagger = 0, \quad \langle 0 | a_k | k\rangle = \langle k | a_k^\dagger | 0\rangle = \sqrt{2\omega} \quad (5.114a)$$

$$\langle 0 | \varphi(x) | k\rangle = \frac{1}{\sqrt{V}} e^{-ik \cdot x}, \quad \langle k | \varphi(x) | 0\rangle = \frac{1}{\sqrt{V}} e^{ik \cdot x} \quad (5.114b)$$

$$k \cdot x \equiv \omega t - \mathbf{k} \cdot \mathbf{x} \quad (5.114c)$$

Problem 5.14

Show that

$$\begin{aligned} |E\rangle &= |k_1, k_2, \dots, k_m, q_1, q_2, \dots, q_n\rangle \\ &\equiv \sqrt{2\omega_{k_1}} a_{k_1}^\dagger \sqrt{2\omega_{k_2}} a_{k_2}^\dagger \dots \sqrt{2\omega_{k_m}} a_{k_m}^\dagger \sqrt{2\omega_{q_1}} b_{q_1}^\dagger \sqrt{2\omega_{q_2}} b_{q_2}^\dagger \\ &\quad \dots \sqrt{2\omega_{q_n}} b_{q_n}^\dagger |0\rangle \end{aligned} \quad (5.115)$$

is a state with m particles and n antiparticles whose total energy and momentum are given by

$$E = \omega_{k_1} + \dots + \omega_{k_m} + \omega_{q_1} + \dots + \omega_{q_n} \quad (5.116a)$$

$$\mathbf{P} = \mathbf{k}_1 + \dots + \mathbf{k}_m + \mathbf{q}_q + \dots + \mathbf{q}_n \quad (5.116b)$$

A state with any number of particles represents a Fock vector, and all Fock vectors constitute a Fock space.

Summary The relativistic wave function suffered from negative-energy particles. But if it is reinterpreted as the classical field and quantized again, the difficulty disappears. Because of this, the quantization of the field is often called second quantization. The quantized field ϕ is no longer a wave function but an operator that acts on the Fock vectors (states in the Heisenberg picture). It has wave characteristics in its expansion in terms of $e^{-ik \cdot x}$ and also a particle nature, in that the coefficients $a_k, a_k^\dagger$ can be interpreted as the square root of the particle number.

5.3.3

Dirac Field

From the Lagrangian of the Dirac field Eq. (5.20d), using Noether's theorem we can derive expressions for the Hamiltonian etc.:

$$H = \int d^3x \psi^\dagger \left(\boldsymbol{\alpha} \cdot \frac{\nabla}{i} + \beta m \right) \psi = \int d^3x \psi^\dagger i \frac{\partial \psi}{\partial t} \quad (5.117a)$$

7) This is a relativistic normalization that we will adopt. The details will be described in Sect. 6.5.2.

$$P = \int d^3x \psi^\dagger (-i\nabla) \psi \quad (5.117b)$$

$$\begin{aligned} qJ^\mu &= q \int d^3x \bar{\psi} \gamma^\mu \psi = q \int d^3x \psi^\dagger (1, \boldsymbol{\alpha}) \psi \rightarrow Q = qJ^0 \\ &= q \int d^3x \psi^\dagger \psi \end{aligned} \quad (5.117c)$$

We expand the field in terms of plane-wave solutions of the Dirac equation (see Sect. 4.2). Writing down the Dirac fields as

$$\psi = \sum_{p,r} \frac{1}{\sqrt{2E_p V}} \left[a_{p,r} u_r(p) e^{-ip \cdot x} + b_{p,r}^\dagger v_r(p) e^{ip \cdot x} \right] \quad (5.118a)$$

$$\bar{\psi} = \sum_{p,r} \frac{1}{\sqrt{2E_p V}} \left[a_{p,r}^\dagger \bar{u}_r(p) e^{ip \cdot x} + b_{p,r} \bar{v}_r(p) e^{-ip \cdot x} \right] \quad (5.118b)$$

H, P, Q can be rewritten as

$$H = \sum_{p,r} E_p \left(a_{p,r}^\dagger a_{p,r} - b_{p,r} b_{p,r}^\dagger \right) \quad (5.119a)$$

$$P = \sum_{p,r} p \left(a_{p,r}^\dagger a_{p,r} - b_{p,r} b_{p,r}^\dagger \right) \quad (5.119b)$$

$$Q = \sum_{p,r} E_p \left(a_{p,r}^\dagger a_{p,r} + b_{p,r} b_{p,r}^\dagger \right) \quad (5.119c)$$

Notice the difference from and similarity to Eq. (5.98). Here we note that if we apply the same logic as we followed after Eq. (5.98), we cannot use the commutation relations like those of Eq. (5.104), because they produce negative energies, as is shown by the second term of the Hamiltonian. But if we impose the following anticommutators:

$$\begin{aligned} \{a_{p,r}, a_{p',r'}^\dagger\} &= \{b_{p,r}, b_{p',r'}^\dagger\} = \delta_{p,p'} \delta_{r,r'} \\ \{a_{p,r}, a_{p',r'}\} &= \{a_{p,r}^\dagger, a_{p',r'}^\dagger\} = \{b_{p,r}^\dagger, b_{p',r'}^\dagger\} = \{b_{p,r}, b_{p',r'}\} = 0 \\ \{a_{p,r}, b_{p',r'}^\dagger\} &= \{a_{p,r}, b_{p',r'}\} = \{a_{p,r}^\dagger, b_{p',r'}^\dagger\} = 0 \end{aligned} \quad (5.120)$$

the energy becomes positive definite and we obtain physically consistent interpretations:

$$H = \sum_{p,r} E_p \left(a_{p,r}^\dagger a_{p,r} + b_{p,r}^\dagger b_{p,r} - 1 \right) \quad (5.121a)$$

$$P = \sum_{p,r} p \left(a_{p,r}^\dagger a_{p,r} + b_{p,r}^\dagger b_{p,r} - 1 \right) \quad (5.121b)$$

$$Q = \sum_{p,r} E_p \left(a_{p,r}^\dagger a_{p,r} - b_{p,r}^\dagger b_{p,r} + 1 \right) \quad (5.121c)$$

Therefore, the Dirac field must obey Fermi–Dirac statistics.

Problem 5.15

Show that Eq. (5.118) can be inverted to give

$$a_{p,r} = \int d^3x \bar{u}_{p,r}(x) e_p^*(x) \gamma^0 \psi(x), \quad a_{p,r}^\dagger = \int d^3x \bar{\psi}(x) \gamma^0 u_{p,r}(x) e_p(x) \quad (5.122a)$$

$$b_{p,r} = \int d^3x \bar{\psi}(x) \gamma^0 v_{p,r}(x) e_p(x), \quad b_{p,r}^\dagger = \int d^3x \bar{v}_{p,r}(x) e_p^*(x) \gamma^0 \psi(x) \quad (5.122b)$$

$$e_p(x) = \frac{1}{\sqrt{2E_p V}} e^{-ip \cdot x}, \quad e_p^*(x) = \frac{1}{\sqrt{2E_p V}} e^{ip \cdot x} \quad (5.122c)$$

The equal time commutators of the Dirac field are given as follows. The conjugate field is defined by

$$\pi_r = \frac{\delta \mathcal{L}}{\delta \psi_r} = i[\bar{\psi} \gamma^0]_r = i \psi_r^\dagger \quad (5.123)$$

Since the Lagrangian contains no derivatives of $\bar{\psi}$, there is no conjugate momentum to $\psi^\dagger$.⁸⁾ The above equation shows that $i \psi^\dagger$ itself is the conjugate momentum to ψ . Therefore the commutators are

$$\{\psi_r(x), \psi_s^\dagger(y)\}_{t_x=t_y} = \delta_{rs} \delta^3(x - y) \quad (5.124a)$$

$$\{\psi_r(x), \psi_s(y)\}_{t_x=t_y} = \{\psi_r^\dagger(x), \psi_s^\dagger(y)\}_{t_x=t_y} = 0 \quad (5.124b)$$

5.3.4

Electromagnetic Field

The electromagnetic field can be expressed by the potential $A^\mu = (\phi, \mathbf{A})$, where ϕ is the scalar or Coulomb potential and $\mathbf{A}$ the vector potential. It is a real vector potential and, when quantized, represents a spin 1 neutral particle, called the photon. It obeys the Maxwell equation in vacuum:

$$\partial_\mu F^{\mu\nu} = \partial_\mu \partial^\mu A^\nu - \partial^\nu (\partial_\mu A^\mu) = 0 \quad (5.125a)$$

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \quad (5.125b)$$

As it has gauge symmetry, which means the electromagnetic field $F^{\mu\nu}$ does not change by a gauge transformation,

$$A^\mu \rightarrow A'^\mu(x) = A^\mu(x) + \partial^\mu \chi(x) \quad (5.126)$$

8) If one uses the symmetric Lagrangian of Eq. (5.21), one sees that the conjugate momentum to $\psi^\dagger$ is $-i\psi$, which yields the same commutation relation.

The gauge has to be fixed for calculating observable quantities. The usual convention is to adopt the Lorentz gauge

$$\partial_\mu A^\mu = 0 : \quad \text{Lorentz gauge} \quad (5.127)$$

Then the Maxwell equation becomes

$$\partial_\mu \partial^\mu A^\nu = 0 \quad (5.128)$$

which is the Klein–Gordon equation with $m = 0$. Even after adopting the Lorentz gauge, we still have an extra degree of freedom to change the potential using a function that satisfies

$$A^\mu \rightarrow A'^\mu = A^\mu + \partial^\mu \xi, \quad \partial_\mu \partial^\mu \xi = 0 \quad (5.129)$$

We shall show that although the potential has four components, it has only two degrees of freedom because of the gauge symmetry. Writing the polarization vector of the photon having energy-momentum k as $\varepsilon^\mu(k, \lambda)$, the potential can be expanded as

$$A^\mu = \sum_{k, \lambda} \frac{1}{\sqrt{2\omega V}} \left[a_{k, \lambda} \varepsilon^\mu(k, \lambda) e^{-ik \cdot x} + a_{k, \lambda}^\dagger \varepsilon^{*\mu}(k, \lambda) e^{ik \cdot x} \right] \quad (5.130)$$

For the free electromagnetic potential, the following Einstein relation holds:

$$k^2 = k_\mu k^\mu = 0 \quad \text{or} \quad \omega = k^0 = |\mathbf{k}| \quad (5.131)$$

The Lorentz condition Eq. (5.127) means

$$k_\mu \varepsilon^\mu(k, \lambda) = 0 \quad (5.132)$$

Equation (5.129) means ε has an extra degree of freedom to change

$$\varepsilon^\mu \rightarrow \varepsilon'^\mu = \varepsilon^\mu + ik^\mu \xi(k), \quad k^2 = 0 \quad (5.133)$$

We can choose $\xi(k)$ to make $\varepsilon^0 = 0$. Then the Lorentz condition becomes

$$\mathbf{k} \cdot \boldsymbol{\varepsilon} = 0 \quad (5.134)$$

Therefore the free electromagnetic field (i.e. the electromagnetic wave) has only transverse polarization, with two degrees of freedom.

Coulomb Gauge and the Covariant Gauge

Because of the gauge freedom, the quantization of the electromagnetic field is non-trivial. The easiest way to understand it intuitively is to adopt the Coulomb gauge, which minimizes the degrees of freedom. In this gauge, the potential satisfies the condition

$$\nabla \cdot \mathbf{A} = 0 : \quad \text{Coulomb gauge} \quad (5.135)$$

Then the field can be expressed in terms of transverse potential $A_{\perp}$ and instantaneous Coulomb potential (see Appendix B). In this gauge, only the transversely polarized photons are quantized; they are treated as two independent massless Klein–Gordon fields. The instantaneous Coulomb field has no time dependence, hence is not a dynamical object. It is left as a classical field. In this gauge, manifest Lorentz covariance is lost, although a detailed examination shows it is not. Due to absence of manifest covariance, calculations in general become complicated, and the usual convention is to adopt a method we refer to as the covariant gauge. It treats all four components of the field potential A^{μ} on an equal footing and independently. Then we expect to have commutators like Eq. (5.141a), which is in conflict with the Lorentz condition Eq. (5.127), at least as an operator condition. We also encounter another problem. The Lagrangian density does not contain A^0 , which means that the conjugate momentum π^0 does not exist. Therefore A^0 cannot be a dynamical variable. A remedy to the problem is to add an additional term, called the gauge fixing term, to the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2\lambda}(\partial_{\mu}A^{\mu})^2 \quad (5.136)$$

Then the Euler–Lagrange equation becomes

$$\partial_{\mu}\partial^{\mu}A^{\nu} - \left(1 - \frac{1}{\lambda}\right)(\partial_{\mu}A^{\mu}) = 0 \quad (5.137)$$

which reproduces the desired Maxwell equation without imposing the Lorentz gauge condition Eq. (5.127), if we set $\lambda = 1$. The price to pay is that the added term is not gauge invariant and we no longer have the freedom to make another gauge transformation. Namely, the gauge has been fixed and the choice of $\lambda = 1$ is called the Feynman gauge.

In the covariant gauge, the four potential fields are independent and we can assign four independent polarization vectors to each component. In the coordinate system where the z axis is taken along the photon momentum $\mathbf{k}$, they can be expressed as

$$\varepsilon^{\mu}(0) = (1, 0, 0, 0) \quad (5.138a)$$

$$\varepsilon^{\mu}(1) = (0, 1, 0, 0) \quad (5.138b)$$

$$\varepsilon^{\mu}(2) = (0, 0, 1, 0) \quad (5.138c)$$

$$\varepsilon^{\mu}(3) = (0, 0, 0, 1) \quad (5.138d)$$

Circular polarization $\varepsilon(\pm)$ is also used instead of linear polarizations $\varepsilon(1)$, $\varepsilon(2)$

$$\varepsilon^{\mu}(+) = (0, -1/\sqrt{2}, -i/\sqrt{2}, 0) \quad (5.138e)$$

$$\varepsilon^{\mu}(-) = (0, 1/\sqrt{2}, -i/\sqrt{2}, 0) \quad (5.138f)$$

A quantum that has polarization $\varepsilon(3)$ is called a longitudinal photon, and that with $\varepsilon(0)$, a scalar photon. The four polarization vectors satisfy the orthogonal condition

$$\varepsilon_\mu(\lambda)\varepsilon^\mu(\lambda')^* = g_{\lambda\lambda'} \quad (5.139)$$

Notice the polarization of the scalar photon, unlike others, is time-like, i.e. $(\varepsilon_\mu(0)\varepsilon^\mu(0) = 1 > 0)$. Then the completeness condition or sum of the polarization states is written as

$$\sum_\lambda g_{\lambda\lambda'}\varepsilon^\mu(\lambda)\varepsilon^\nu(\lambda')^* = g^{\mu\nu} \quad (5.140)$$

Unphysical Photons

The usual quantization method we have developed so far leads to a commutation relation for the electromagnetic fields

$$[A^\mu(x), \pi^\nu(y)]_{t_x=t_y} = -ig^{\mu\nu}\delta^3(\mathbf{x} - \mathbf{y}) \quad (5.141a)$$

$$\pi^\mu = \frac{\delta \mathcal{L}}{\delta \dot{A}^\mu} = (-\partial_\mu A^\mu, \mathbf{E}) \quad (5.141b)$$

As we can see, the Lorentz gauge condition $(\partial_\mu A^\mu = 0)$ contradicts the above commutation relation and forces π^0 to vanish, thereby invalidating construction of canonical commutation relations for the fields. So we adopt a weaker Lorentz condition valid only for the observable states $|\Phi\rangle, |\phi\rangle$

$$\langle\phi|\partial_\mu A^\mu|\Phi\rangle = 0 \quad (5.142)$$

and eliminate unwanted components from the final expression for the observables. As a rigorous mathematical treatment of the method is beyond the scope of this textbook, we only describe the essential ingredients that are necessary for later discussions. The commutation relation consistent with the field commutator Eq. (5.141a) is

$$[a_{\mathbf{k},\lambda}, a_{\mathbf{k}',\lambda'}^\dagger] = -g_{\lambda\lambda'}\delta_{\mathbf{k}\mathbf{k}'} \quad (5.143)$$

which gives a negative norm $(\langle\lambda = 0|\lambda = 0\rangle = -1)$ for the scalar photon. Obviously a photon having negative probability is unphysical. The fact is that the longitudinal and the scalar photon appear only in intermediate states and never appear as physical photons, which satisfy the Einstein relation $k_\mu k^\mu = \omega^2 - |\mathbf{k}|^2 = 0$. The unusual property of the scalar photon works in such a way as to compensate unphysical contributions of the longitudinal polarization states.

Intermediate photons created by electric current j^μ satisfy the Maxwell equation

$$\partial_\mu \partial^\mu A^\nu = q j^\nu \quad (5.144)$$

Comparing both sides in momentum space, it is clear that the energy-momentum of the intermediate photon does not satisfy the Einstein relation for the particle

$$k_\mu k^\mu \neq 0 \quad (5.145)$$

The photon is said to be off-shell, or in a virtual state. Equation (5.144) is a result of interaction terms in the Lagrangian

$$\mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L} - q j_\mu A^\mu \quad (5.146)$$

Later we shall learn that the interaction term in the above equation appears naturally as a consequence of the “local” gauge invariance for a system including both the matter (Dirac) field and the electromagnetic field, but here we assume it as given. Virtual photons are produced and absorbed by the currents (i.e. by interacting charged particles) and appear as a propagator in the form $\sim 1/k^2$ (to be described soon) in the equation. While in the Coulomb gauge only the transverse polarizations need to be considered, in the covariant gauge all four polarization states have to be included:

$$\sum_{\lambda=\pm} \varepsilon^\mu(\lambda) \varepsilon^\nu(\lambda)^* = \delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2} \quad \text{Coulomb gauge} \quad (5.147a)$$

$$\rightarrow \sum_{\lambda=0}^3 \varepsilon^\mu(\lambda) \varepsilon^\nu(\lambda)^* \rightarrow -g^{\mu\nu} \quad \text{covariant gauge} \quad (5.147b)$$

The arrow in Eq. (5.147b) means it is not an actual equality. Negative norm for the scalar polarization is implicitly assumed as shown by Eq. (5.140). To see the difference between the two methods, we use the fact that in the transition amplitude they appear sandwiched by their sources, the conserved currents. As a typical example, take the matrix element for electron–muon scattering, which has the form [see Eq. (7.12)]

$$\mathcal{M} \sim j_\mu(e) \frac{-g^{\mu\nu}}{k^2} j_\nu(\mu) \quad (5.148)$$

Since the conserved current satisfies

$$\partial_\mu j^\mu = 0 \rightarrow k_\mu j^\mu = k^0 j^0 - |\mathbf{k}| j^3 = 0 \quad (5.149)$$

in the coordinate system where the z -axis is taken along the momentum direction, Eq. (5.148) after summation of polarization states can be rewritten as

$$\begin{aligned} j_\mu \frac{-g^{\mu\nu}}{k^2} j_\nu &= \frac{1}{k^2} \left[\underbrace{-g^{\mu 0} g^{\nu 0}}_{\text{scalar}} + \underbrace{\frac{k^i k^j}{k^2}}_{\text{longitudinal}} + \underbrace{\left(\delta_{ij} - \frac{k^i k^j}{k^2} \right)}_{\text{transverse}} \right] j_\mu j_\nu \\ &= \frac{1}{k^2} \left(\frac{(k^0)^2}{k^2} - 1 \right) (j_0)^2 + \frac{\mathbf{j}_\perp \cdot \mathbf{j}_\perp}{k^2} = \frac{j_0 j_0}{|\mathbf{k}|^2} + \frac{\mathbf{j}_\perp \cdot \mathbf{j}_\perp}{k^2} \end{aligned} \quad (5.150a)$$

$$\mathbf{j}_\perp = \mathbf{j} - \frac{\mathbf{k}}{|\mathbf{k}|^2} (\mathbf{k} \cdot \mathbf{j}), \quad \mathbf{k} \cdot \mathbf{j}_\perp = 0 \quad (5.150b)$$

The first term represents the contribution of an instantaneous Coulomb potential⁹⁾ and the second is the contribution of the current coupled to the transverse photon. In the Coulomb gauge, the two terms are treated separately and only the transverse photons are quantized. In the covariant gauge it is included in the contribution of the longitudinal and the scalar photon. Note, the proof was given in a special coordinate system, but Eq. (5.148) and the current conservation law Eq. (5.149) are written in Lorentz-invariant form. Therefore, a proof valid in one coordinate system is equally valid in any coordinate system connected by the proper Lorentz transformation.

When the photon appears as a real (external) particle, we may still take the total sum instead of only transverse polarizations, as will be shown later in Eq. (7.54).

In summary, longitudinal and scalar photons, which are artifacts of the covariant gauge, have unphysical properties, but they appear only in intermediate states, manifesting themselves as the instantaneous Coulomb potential. We can treat all four components of the electromagnetic potential in a Lorentz-covariant way and may safely take the polarization sum from 0 to 3 to obtain correct results.

5.4

Spin and Statistics

We have seen that the creation and annihilation operators have to satisfy the commutation relation for the scalar and vector fields and anticommutation relation for the Dirac field. This was a requirement to make the energy positive definite. The relation can be generalized and the following well-known spin-statistics relation holds.

Particles having integer spin obey Bose–Einstein statistics and those with half-integer spin obey Fermi–Dirac statistics.

We shall re-derive the spin-statistics from a somewhat different point of view. We allow both commutation and anticommutation relations for the creation–annihilation operators and see what happens. Writing $[A, B]_{\pm} = AB \pm BA$, we set

$$[a_k, a_{k'}^{\dagger}]_{\pm} = \delta_{kk'}, \quad [a_k, a_{k'}]_{\pm} = [a_k^{\dagger}, a_{k'}^{\dagger}]_{\pm} = 0 \quad (5.151)$$

We first calculate the commutators of the scalar fields. Using the above equations we can show

$$[\phi(x), \phi(y)]_{\mp} = \sum_k \frac{1}{2\omega V} [e^{-ik \cdot (x-y)} \mp e^{ik \cdot (x-y)}] = \begin{cases} i\Delta(x-y) \\ i\Delta_1(x-y) \end{cases} \quad (5.152)$$

9) $(j^0)^2/k^2$ is the Fourier transform of a static Coulomb potential $V = \rho(x)\rho(y)/|\mathbf{x} - \mathbf{y}|$.

where

$$\begin{aligned} i\Delta(x-y) &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} [e^{-ik \cdot (x-y)} - e^{ik \cdot (x-y)}]^{10)} \\ &= \int \frac{d^4k}{(2\pi)^3} \varepsilon(k^0) \delta(k^2 - m^2) e^{-ik \cdot (x-y)} \end{aligned} \quad (5.153a)$$

$$\begin{aligned} i\Delta_1(x-y) &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} [e^{-ik \cdot (x-y)} + e^{ik \cdot (x-y)}] \\ &= \int \frac{d^4k}{(2\pi)^3} \delta(k^2 - m^2) e^{-ik \cdot (x-y)} \\ &\sim \frac{\exp\left[-m\sqrt{|x-y|^2 - (x^0 - y^0)^2}\right]}{|x-y|^2 - (x^0 - y^0)^2} \\ &\text{for } -(x-y)^2 \gg \frac{1}{m^2} \end{aligned} \quad (5.153b)$$

where $\varepsilon(k^0) = \pm 1$ for $k^0 \gtrless 0$ and we used

$$\delta(k^2 - m^2) = \frac{1}{2E} [\delta(k^0 - E) + \delta(k^0 + E)], \quad E = \sqrt{\mathbf{k}^2 + m^2} \quad (5.154)$$

The function $\Delta(x)$ has the properties

$$\Delta(x)|_{x^0=0} = 0, \quad \left. \frac{\partial}{\partial t} \Delta(x) \right|_{x^0=0} = -\delta^3(\mathbf{x}) \quad (5.155a)$$

$$[\partial_\mu \partial^\mu + m^2] \Delta(x) = 0 \quad (5.155b)$$

It is Lorentz invariant, as can be seen from Eq. (5.153a), because the factors in the integral are all Lorentz scalars. This applies also to $\varepsilon(k^0)$, since it lies on the light cone ($k^2 - m^2 = 0$) and does not change sign by any proper (orthochronous) Lorentz transformation: time-like ($k^2 > m^2$) momentum vectors with $k^0 > 0$ always lie in the forward light cone and those with $k^0 < 0$ always in the backward light cone. Since $\Delta(x-y) = 0$ for $x^0 = y^0$, the Lorentz invariance dictates that it also vanishes at space-like distance [$(x-y)^2 < 0$]. This is not true for the function Δ_1 . It is also Lorentz invariant but it does not vanish at equal time. In fact, according to Eq. (5.153b) it does not vanish outside the light cone, at least within a distance of the Compton wavelength of the particle. This means that if we adopt the anticommutator for the scalar field it is described by the function $\Delta_1(x-y)$, and causality is violated.¹¹⁾ Remember that in quantum mechanics, uncommutative operators obey the uncertainty principle and their values cannot be determined independently. On the other hand, if we adopt the commutator then it becomes $\Delta(x-y)$, the two measurements at space-like distance are not related and causality is respected.

10) For the relation between discrete sum and continuous integral, see Eq. (5.94).

11) Strictly speaking we should talk about commutators of observables composed of ϕ 's, but they are related by simple algebra.

We can do similar calculations for the Dirac field:

$$\begin{aligned}
 [\psi(x), \bar{\psi}(y)]_{\pm} &= \sum_{p,r} \frac{1}{2E_p V} [u_r(p) \bar{u}_r(p) e^{-ip \cdot (x-y)} \pm v_r(p) \bar{v}_r(p) e^{ip \cdot (x-y)}] \\
 &= \sum_p \frac{1}{2E_p V} [(\not{p} + m) e^{-ip \cdot (x-y)} \pm (\not{p} - m) e^{ip \cdot (x-y)}] \\
 &= (i\not{\nabla} + m) \int \frac{d^3 p}{(2\pi)^3 2E_p} [e^{-ip \cdot (x-y)} \mp e^{ip \cdot (x-y)}] \\
 &= \begin{cases} (i\not{\nabla} + m) i\Delta(x-y) \\ (i\not{\nabla} + m) i\Delta_1(x-y) \end{cases} \quad (5.156)
 \end{aligned}$$

where $\not{p} = \gamma^\mu p_\mu$ is the Feynman slash. This shows that unless we choose the anticommutation relation for the Dirac field, the causality condition is not satisfied. Since fields of spin higher than 1/2 can be considered as being built up of products of spin-1/2 fields, any (half) integer spin field has to satisfy (anti)commutation relations if causality is respected.

5.5

Vacuum Fluctuation

We have seen that the Hamiltonian of the fields [see Eq. (5.105a) and Eq. (5.121a)] has a constant term, i.e. the zero point energy. In the case of the harmonic oscillator, it represents the uncertainty of the position in confinement of the potential, and the resultant uncertainty of the momentum, in turn, produces the minimum energy even in the ground state. Namely, it is a quantum fluctuation associated with the uncertainty principle. In field theory it becomes a diverging vacuum energy. Usually, observed energy is measured from a certain reference point and only the difference matters in ordinary situations. As long as the infinite energy stays constant, it does no harm, unless we are treating gravity. Experimentally, however, there is evidence that the zero-point energy is real. It is the energy of the vacuum, where “nothing!” exists. By nothing we mean every form of matter that is known to us and energy fields such as the electromagnetic field. It permeates space, so how do we know of its existence? The fact is that lacking a gravitational tool we cannot know its absolute value. But we can change the boundary condition surrounding the vacuum we want to measure and by doing so its variation can be measured. Other known examples of vacuum fluctuation are the Lamb shift, which is described in Sect. 8.2.1, and van der Waals forces, which acts on molecules.

We emphasize the importance of understanding the vacuum energy for two reasons. First, the Standard Model of elementary particles has its foundation in the Higgs mechanism, which considers the vacuum to be a dynamical object that can metamorphose and is indeed in a sort of superconducting phase as a way of generating mass for the particles. Second, the recently established concordance model of cosmology acknowledges that 70% of the cosmic energy is the vacuum energy,

alias dark energy which causes accelerated expansion of the universe. Pursuit of the vacuum energy could well be the theme of the 21st century.

5.5.1

The Casimir Effect*

Let us investigate the existence of the zero-point energy. As an example, consider a limited space between two large parallel metal plates.

We idealize the plate by assuming it to be perfectly conducting and to have a square shape of size L^2 at a gap distance $d \ll L$, as depicted in Fig. 5.1; we ignore edge effects. The vacuum energy of the electromagnetic field is given by

$$E_0 = \frac{1}{2} \sum_{\mathbf{k}, \lambda} \omega_{\mathbf{k}}, \quad \omega_{\mathbf{k}} = \sqrt{k_x^2 + k_y^2 + k_z^2} \quad (5.157)$$

When the gap distance is sufficiently small, the wave modes that can exist inside the gap volume are limited. Only standing waves are allowed. We expect the volume to be filled with less energy inside the boundary than outside the boundary, and there will be an attractive force between the plates, which is called the Casimir force [92].¹²⁾ Although the zero-point energy in each case is divergent, the difference turns out to be finite. The calculation goes as follows. The wave number, k_z , that is perpendicular to the plates can have only discrete values $k_z = n\pi/d$ ($n = 0, 1, 2, \dots$), whereas it can be continuous if the plates do not exist. Only waves with transverse polarization contribute to the energy density:

$$\begin{aligned} E_{\text{free}} &= \frac{1}{2} \sum_{\mathbf{k}, \lambda=1,2} \omega_{\mathbf{k}} = \frac{1}{2} \frac{2L^2 d}{(2\pi)^3} \int d^3 \mathbf{k} \sqrt{k_x^2 + k_y^2 + k_z^2} \\ &= \frac{1}{2} \frac{L^2}{(2\pi)^2} \int dk_x dk_y \int_{-\infty}^{\infty} \frac{dk_z}{2\pi} 2 \sqrt{k_x^2 + k_y^2 + k_z^2} \\ &= \frac{1}{2} \frac{L^2}{(2\pi)^2} \int dk_x dk_y \int_0^{\infty} dn 2 \sqrt{k_x^2 + k_y^2 + \left(\frac{n\pi}{d}\right)^2} \end{aligned} \quad (5.158a)$$

$$E_{\text{gap}} = \frac{1}{2} \frac{L^2}{(2\pi)^2} \int dk_x dk_y \left\{ \sqrt{k_x^2 + k_y^2} + 2 \sum_{n=1}^{\infty} \sqrt{k_x^2 + k_y^2 + \left(\frac{n\pi}{d}\right)^2} \right\} \quad (5.158b)$$

In writing Eq. (5.158b), we have taken into account that when $k_z = 0$, the propagating waves in the gap can have only one polarization mode. Then the difference

12) The sign of the Casimir force depends on the geometry. For instance, the geometry of two hemispheres causes a repulsive force [74].

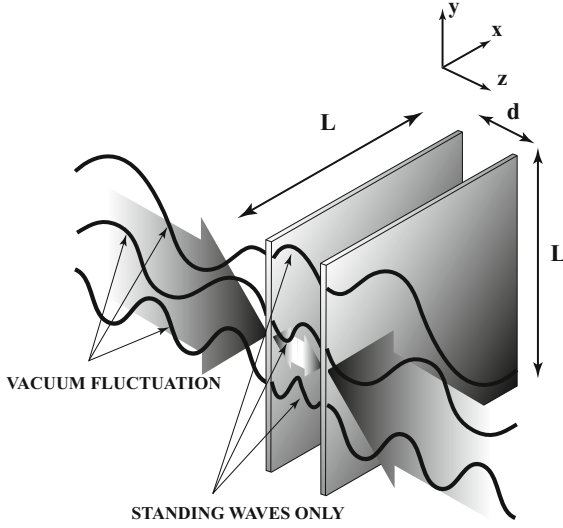


Figure 5.1 Casimir effect illustrated. Between two metal plates, an attractive force is generated through vacuum fluctuations.

in the energy density is

$$\begin{aligned}
 \mathcal{E} &= \frac{E_{\text{gap}} - E_{\text{free}}}{L^2 d} \\
 &= \frac{1}{2\pi d} \int_0^\infty k dk \left\{ \frac{k}{2} + \sum_{n=1}^\infty \sqrt{k^2 + \left(\frac{n\pi}{d}\right)^2} - \int_0^\infty dn \sqrt{k^2 + \left(\frac{n\pi}{d}\right)^2} \right\} \\
 &= \frac{\pi^2}{8d^4} \int_0^\infty d\tau \left[\sum_{n=-\infty}^\infty \sqrt{\tau + n^2} - \int_{-\infty}^\infty dz \sqrt{\tau + z^2} \right] = -\frac{\pi^2}{720d^4}
 \end{aligned} \tag{5.159}$$

Proof: Equation (5.159) is apparently not well defined because it diverges as $\tau \rightarrow \infty$. Therefore we regularize the equation by putting¹³⁾

$$f_a(\tau, z) = \sqrt{\tau + z^2} e^{-a\sqrt{\tau + z^2}} \tag{5.160}$$

Then we can use the Abel–Plana formula

$$\sum_{n=0}^\infty f(n) - \int_0^\infty f(z) dz = \frac{1}{2} f(0) + i \int_0^\infty \frac{f(it) - f(-it)}{e^{2\pi t} - 1} dt \tag{5.161}$$

13) The regularization treatment is not quite an ad hoc patch up. For wavelengths shorter than the atomic size, it is unrealistic to use a perfect conductor approximation. So it is natural to introduce a smooth cutoff, which attenuates the wave when $k \gg 1/a$, where $a \simeq$ atomic distance. In fact, it is known that above the plasma frequency $\omega = \sqrt{ne^2/m\epsilon_0}$ a metal becomes transparent to light.

Conditions for the formula to be valid are:

- (1) $f(z)$ is analytic (differentiable) in the region $\text{Re}(z) \geq 0$.
- (2) $\lim_{y \rightarrow \pm\infty} e^{-2\pi|y|} f(x + iy) = 0$ uniformly for all $x \geq 0$.
- (3) $\lim_{x \rightarrow \infty} \int_{-\infty}^{\infty} e^{-2\pi|y|} f(x + iy) dy = 0$.

Obviously the regularized function $f(z) = \int_0^\infty d\tau f_\alpha(\tau, z)$ Eq. (5.160) satisfies the conditions (2) and (3) for $\alpha > 0$. We note we can safely put $\alpha = 0$ in the r.h.s. of the Abel–Plana formula. When $\tau + t^2 < 0$, we choose the branch of the square root function as follows to ensure condition (1):

$$f(it) = f(-it) = \sqrt{\tau - t^2} \quad \text{for } t^2 < \tau \quad (5.162a)$$

$$f(it) = -f(-it) = i\sqrt{t^2 - \tau} \quad \text{for } t^2 > \tau \quad (5.162b)$$

Then Eq. (5.159) becomes

$$\begin{aligned} \mathcal{E} &= \frac{\pi^2}{8d^4} \int_0^\infty d\tau 2 \left[\sum_{n=0}^\infty \sqrt{\tau + n^2} - \int_0^\infty dz \sqrt{\tau + z^2} - \frac{1}{2} \sqrt{\tau} \right] \\ &= \frac{\pi^2}{8d^4} \int_0^\infty d\tau 2i \int_0^\infty dt \frac{f(it) - f(-it)}{e^{2\pi t} - 1} \\ &= -\frac{\pi^2}{2d^4} \int_0^\infty d\tau \int_{\sqrt{\tau}}^\infty dt \frac{\sqrt{t^2 - \tau}}{e^{2\pi t} - 1} \\ &= -\frac{\pi^2}{2d^4} \int_0^\infty dt \int_0^{t^2} d\tau \frac{\sqrt{t^2 - \tau}}{e^{2\pi t} - 1} = -\frac{\pi^2}{3d^4} \int_0^\infty dt \frac{t^3}{e^{2\pi t} - 1} \\ &= -\frac{\pi^2}{3d^4} \int_0^\infty dt \sum_{n=1}^\infty t^3 e^{-2\pi n t} = -\frac{1}{8\pi^2} \zeta(4) = -\frac{\pi^2}{720d^4} \end{aligned} \quad (5.163)$$

where the Zeta (ζ) function is defined in the region $\text{Re}(s) > 1$ as

$$\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s} \quad (5.164)$$

which is known to be holomorphic (an analytic function) in all the complex s plane except at $s = 1$. $\square$

Now that we have obtained the extra energy density in the gap, we can calculate the Casimir force per unit area that acts on the parallel plates:

$$F_{\text{Casimir}} = -\frac{\partial}{\partial d} \left(\frac{\mathcal{E} V}{L^2} \right) = -\frac{\pi^2 \hbar c}{240d^4} = \frac{1.3 \times 10^{-3}}{\{d[\mu\text{m}]\}^4} \text{ N/m}^2 \quad (5.165)$$

Practically, the above value has to be corrected for nonperfect conductivity and for the finite temperature environment. When an electric potential of 17 mV is applied between the plates at the distance of $d = 1 \mu\text{m}$, the Coulomb force becomes stronger, and therefore it is a very difficult experiment to prove the formula. Experimentally, it is difficult to configure two parallel plates uniformly separated by

a distances of less than a micrometer, a minimum requirement for the Casimir force to become dominant over gravity between the plates. Therefore, one plate is replaced by a sphere of radius R . For $R \gg d$, where d is the minimum distance between the sphere and the plate, the Casimir force acting on the sphere is modified to

$$F_{\text{sphere}} = 2\pi R\mathcal{E} = -\frac{\pi^3}{360} R \frac{\hbar c}{d^3} \quad (5.166)$$

The first successful experiment [249] proved the existence of the Casimir force to an accuracy of better than 5%. Recent experiments (Fig. 5.2) have reached an accuracy on the 1% level.

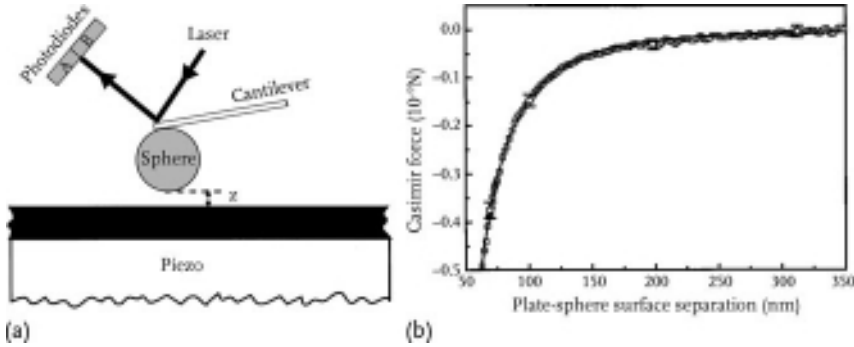


Figure 5.2 (a) The experimental apparatus of Harris et al. [197]. A metal sphere is supported by a cantilever. When the Casimir force acts on the sphere, the cantilever bends and its motion is measured by the displacement

of the laser reflection at the photodiodes. (b) Experimental data on the Casimir force as a function of the plate–sphere distance. The data exactly fit the theoretical curve.

Strange Properties of the Vacuum

Negative Pressure A vacuum filled with zero-point energy has the same density everywhere. This means that the vacuum has negative pressure. Compressed air confined in a balloon has outward pressure. Does the negative pressure of the vacuum produce inward forces? If we confine a small volume of vacuum in a cylinder and consider the work the vacuum does outside as depicted in Fig. 5.3, $P = -\rho$ can be shown. Another way is to use the first law of thermodynamics. For an isolated system, the internal energy decreases if work is done to the outside world, namely $dU = -PdV$. Since $U = \rho V$, we obtain

$$\frac{d\rho}{dV} = -\frac{\rho + P}{V} = 0, \quad \therefore \quad P = -\rho \quad (5.167)$$

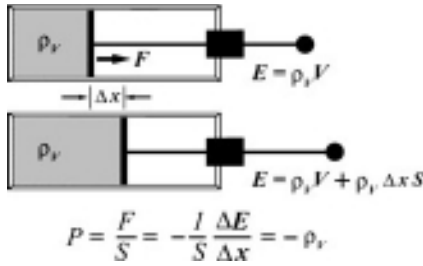


Figure 5.3 Negative pressure of vacuum. When confined vacuum in a box with volume V expands, exerting force on the moving wall, the pressure is given by

$$P = F/S = -\partial E / \partial V = -(1/S)(\partial E / \partial x).$$

Since the total vacuum energy in a volume is $\rho_v V$ ($\rho_v = \text{constant}$), $P = -\rho_v$.

Antigravity We know our universe is expanding, and until only a few years ago its rate was thought to be decreasing because the gravitational force of matter in the universe counteracts it. Intuitively one might think the space filled with vacuum energy would shrink even more. On the contrary, it accelerates the expansion according to general relativity. Namely, the vacuum energy has repulsive gravitational force. Whether the universe keeps expanding or eventually collapses depends critically on the balance between matter and vacuum energy in the universe. It is known that, in cosmic history, there were two epochs in which the vacuum energy dominated: the inflation epoch and now. The vacuum energy of the former could be of dynamic origin, where the field (dubbed inflaton) breaks a symmetry spontaneously and liberates the vacuum energy in the form of latent heat that has disappeared as time went by. But the vacuum energy (also known as dark energy), which occupies $\sim 70\%$ of the present cosmic energy budget, seems static, i.e. has constant energy density (see Section 19.2.7). This means it could also have existed in the inflation epoch. However, the observed value differs from that of the inflation by some 120 orders of magnitude! Remember, for instance, that a tiny quark ($< 10^{-18}$ cm) could be larger than the whole universe (> 130 billion light years) if magnified by a mere 50 orders of magnitude. How could nature provide two kinds of energy of such magnitude? This is a big number. This is the notorious cosmological constant problem. The constant vacuum energy density means also that the total cosmic energy increases as the universe expands. Is the law of energy conservation violated?

6

Scattering Matrix

Clues to clarify the inner structure of particles can be obtained, in most cases, by investigating their scattering or decay patterns. In nonrelativistic quantum mechanics, the transition rates of these processes are usually calculated using perturbation expansions. The underlying assumption is that the interaction energy, i.e. the expectation value of the interaction Hamiltonian, is small compared to the system energy, which is represented by the total Hamiltonian H . The transition rate in quantum mechanics is given by the so-called Fermi's golden rule:

$$W_{fi} = 2\pi |\langle f|T|i\rangle|^2 \rho \quad (6.1a)$$

$$\langle f|T|i\rangle = \langle f|H_I|i\rangle + \sum_{n \neq i} \frac{\langle f|H_I|n\rangle \langle n|H_I|i\rangle}{E_i - E_n} \quad (6.1b)$$

Here, $\langle f|H_I|i\rangle$ is the matrix element of the interaction Hamiltonian between the initial state and the final state. ρ is the final state density, and in the case of a one-body problem with a scattering center it is expressed as

$$\rho = \frac{dn}{dE} = \delta(E_i - E_f) \frac{V d^3 p}{(2\pi)^3} \quad (6.1c)$$

where V is the normalization volume. The purpose of this chapter is to derive relativistic version of Eqs. (6.1a)–(6.1c) and prepare materials. Namely, we calculate the relativistic version of the vertex functions ($\langle n|H_I|i\rangle$) of the electromagnetic interaction and the propagators [$1/(E_i - E_n)$] in Eq. (6.1b). QED calculation of the various processes, which amounts to assembling these materials in the form of scattering matrices, will be dealt with in the next chapter.

6.1

Interaction Picture

In quantized field theory, the field operators in the Heisenberg picture obey the Heisenberg equation of motion:

$$\frac{dO_H}{dt} = i[H, O_H], \quad O_H = O(\phi_H, \pi_H) \quad (6.2)$$

where “ O ” means any operator composed of the field ϕ and its conjugate π . States on which the operators act are the Fock vectors $|\alpha_H\rangle$, which are constant in time. So far we have expressed every physical object in terms of the free fields. From now on we want to handle interacting fields. The Heisenberg picture naturally takes into account interactions when the Hamiltonian contains the interaction. However, generally we do not know how to express the exact solutions in analytical form, if they can be solved formally. The standard procedure is that we start from the free field, for which the exact solutions are known, and treat the interaction as small perturbation. This is the most practical way to approach realistic problems. It is convenient to introduce the interaction picture to handle the perturbative expansion. To formulate it we start from the Schrödinger picture. In the Schrödinger picture, operators are constants of time and all the time dependence is contained in the state vectors. In the following, we attach the subscripts S, H and I for operators and vectors in the Schrödinger, Heisenberg and interaction pictures, respectively. The Schrödinger equation is expressed as

$$i \frac{\partial}{\partial t} |\alpha_S(t)\rangle = H_S |\alpha_S(t)\rangle \quad (6.3)$$

A formal solution can be written as

$$|\alpha_S(t)\rangle = e^{-iH_S t} |\alpha_S(0)\rangle \quad (6.4)$$

where $|\alpha_S(0)\rangle$ is to be determined by initial conditions. The Heisenberg picture can be obtained by a unitary transformation

$$|\alpha_H(t)\rangle = e^{iH_S t} |\alpha_S(t)\rangle = |\alpha_S(0)\rangle \quad (6.5a)$$

$$\phi_H = e^{iH_S t} \phi_S e^{-iH_S t}, \quad \pi_H = e^{iH_S t} \pi_S e^{-iH_S t} \quad (6.5b)$$

$$O_H = e^{iH_S t} O_S(\phi_S, \pi_S) e^{-iH_S t} = O(\phi_H, \pi_H) \quad (6.5c)$$

$$H_H = e^{iH_S t} H_S e^{-iH_S t} = H_S = H(\phi_H, \pi_H) \quad (6.5d)$$

Note, the last equation tells us that the Hamiltonian in the Heisenberg picture expressed in terms of the Heisenberg operators is the same as that in the Schrödinger picture. So we shall simply use H to denote the total Hamiltonian in both pictures. Despite the time independence of the Heisenberg state vector, the time variable is retained. It does not mean the state changes with time, but rather it should be interpreted as reflecting the dynamical development of the operator at time t . The Heisenberg equation (6.2) can easily be derived from Eq. (6.5c).

The interaction picture is obtained by first separating the total Hamiltonian into the free and interacting parts:

$$H = H_0 + H_{IS} \quad (6.6a)$$

$$|\alpha_I(t)\rangle = e^{iH_0t} |\alpha_S(t)\rangle \quad (6.6b)$$

$$\phi_I(t) = e^{iH_0t} \phi_S e^{-iH_0t}, \quad \pi_I(t) = e^{iH_0t} \pi_S e^{-iH_0t} \quad (6.6c)$$

The equation of motion that $|\alpha_I(t)\rangle$ satisfies is given by

$$i \frac{\partial}{\partial t} |\alpha_I(t)\rangle = e^{iH_0t} \left(-H_0 + i \frac{\partial}{\partial t} \right) |\alpha_S(t)\rangle = H_I(t) |\alpha_I(t)\rangle \quad (6.6d)$$

$$H_I(t) = e^{iH_0t} H_{IS} e^{-iH_0t} = H_I[\phi_I(t), \pi_I(t)] \quad (6.6e)$$

The operators in the interaction picture satisfy the free field equation

$$\frac{d\phi_I(t)}{dt} = i[H_0, \phi_I(t)] \quad (6.6f)$$

where $H_0 = H_0\{\phi_I(t), \pi_I(t)\}$. State vectors and operators in the Heisenberg and the interaction picture are related by a unitary transformation:

$$|\alpha_H(t)\rangle = e^{iHt} |\alpha_S(t)\rangle = e^{iHt} e^{-iH_0t} |\alpha_I(t)\rangle \equiv U^\dagger(t) |\alpha_I(t)\rangle \quad (6.7a)$$

$$O_H = U^\dagger(t) O_I U(t) \quad (6.7b)$$

The time dependence of the operators in the interaction picture is dictated by H_0 , and that of the state vector by H_I . Therefore, if H_I vanishes, $|\alpha_I(t)\rangle$ does not depend on time and the interaction picture coincides with the Heisenberg picture. The equation of motion for the state vector $|\alpha_I(t)\rangle$ resembles the Schrödinger equation but the Hamiltonian H_I includes the free field operators that are time dependent.

A convenient feature of the interaction picture is that the field operators always satisfy the equation of free fields and that the field expansion developed in the previous section can be directly applied. That is no longer true for fields in the Heisenberg picture. The interacting Klein–Gordon field, for instance, can be expanded just like the free field

$$\phi_H(x) = \sum_k \frac{1}{\sqrt{2\omega V}} \left[a_H(k) e^{-ik \cdot x} + b_H^\dagger(k) e^{ik \cdot x} \right] \quad (6.8)$$

However, the coefficients of the plane waves are related with those of free fields by

$$a_H(k) = U^\dagger(t) a_{\text{free}}(k) U(t), \quad b_H^\dagger(k) = U^\dagger(t) b_{\text{free}}^\dagger(k) U(t) \quad (6.9)$$

and no longer retain their simple physical interpretation as annihilation and creation operators for single quantum of the field. Note, the interacting fields satisfy

the same commutation relation as the free fields, because the commutation relation is invariant under unitary transformation.

What we shall need in the following is the transition amplitude from, say, a state $|\alpha(t')\rangle$ at time t' , which develops to $|\alpha(t)\rangle$ at time t , then transforms to another state $|\beta(t)\rangle$:

$$\langle f|i\rangle = \langle \beta_H(t)|\alpha_H(t)\rangle = \langle \beta_S(t)|e^{-iH(t-t')}|\alpha_S(t')\rangle \quad (6.10a)$$

where we have used Eqs. (6.5a) and (6.4). The transition amplitude does not depend on which picture is used because they are related by unitary transformations. Now we express it in the interaction picture. Using Eq. (6.7a), they can be expressed as

$$\langle f|i\rangle = \langle \beta_I(t)|\alpha_I(t)\rangle = \langle \beta_I(t)|e^{iH_0t}e^{-iH(t-t')}e^{-iH_0t'}|\alpha_I(t')\rangle \quad (6.10b)$$

$$|\alpha_I(t)\rangle = U(t, t')|\alpha_I(t')\rangle \quad (6.10c)$$

$$U(t, t') = e^{iH_0t}e^{-iH(t-t')}e^{-iH_0t'} = U(t)U^\dagger(t') \quad (6.10d)$$

$U(t, t')$ is referred to as the time evolution operator. $U(t, t')$ is manifestly unitary and satisfies the following identity (for $t_1 \geq t_2 \geq t_3$):

$$\begin{aligned} U(t_1, t_2)U(t_2, t_3) &= U(t_1, t_3) \\ U(t_1, t_3)U^\dagger(t_2, t_3) &= U(t_1, t_2) \\ U(t, t) &= 1 \end{aligned} \quad (6.11)$$

If the interaction works only in a limited volume, an assumption which is generally justified, it vanishes for $t \rightarrow \infty$, $t' \rightarrow -\infty$ in scattering processes.¹⁾ Then, $\lim_{t \rightarrow -\infty} |\alpha_I(t)\rangle$, $\lim_{t \rightarrow \infty} |\beta_I(t)\rangle$ can be taken as free Hamiltonian eigenstates. We may call the states prepared before scattering “in” states and those after scattering “out” states. As they are both eigenstates of the free Hamiltonian and are considered to make a complete orthonormal basis, they are connected by a unitary transformation, which we call the scattering matrix and denote as S . Then Eq. (6.10b) and Eq. (6.10c) give

$$S = U(\infty, -\infty) \quad (6.12a)$$

$$\langle f|i\rangle = \langle \beta_{\text{out}}|\alpha_{\text{in}}\rangle = \langle \beta_{\text{in}}|S|\alpha_{\text{in}}\rangle = \langle \beta_{\text{out}}|S|\alpha_{\text{out}}\rangle \quad (6.12b)$$

$$\langle \beta_{\text{out}}| = \langle \beta_{\text{in}}|S, \quad \text{or} \quad |\alpha_{\text{in}}\rangle = S|\alpha_{\text{out}}\rangle \quad (6.12c)$$

The mean transition probability per unit time is given by

$$W_{fi} = \lim_{\substack{t \rightarrow \infty \\ t_0 \rightarrow -\infty}} \frac{|\langle \beta_I(t)|U(t, t_0)|\alpha_I(t_0)\rangle - \delta_{fi}|^2}{t - t_0} \quad (6.13)$$

1) The interaction takes place typically within the scale of hadron size $\sim 10^{-15}$ m. If a particle traverses the region at the velocity of light, the typical time scale of scattering is $\sim 10^{-24}$ s. Comparatively speaking, the human time scale, for instance “a second”, is infinite. For a long-range Coulomb-like force, the interaction range is, in principle, infinite but a screening effect due to the dielectricity of matter effectively damps the force beyond a few atomic distances.

where by subtracting δ_{fi} , we have eliminated the process where the incoming particles pass without interaction. If the asymptotic states are represented by plane waves of definite momentum, the wave function extends over all space and can never be free. For correct treatment we need to use a wave packet formalism, but here we simply assume an adiabatic condition is implicit, in which the interaction Hamiltonian is slowly turned on and off at $|t| \rightarrow \infty$ so that there is no energy transfer to the system. A mathematical technique to warrant the assumption is to modify

$$H_I(t) \rightarrow H_I^\varepsilon, \quad H_I^\varepsilon = H_I e^{-|\varepsilon|t} \quad (6.14)$$

and let ε go to zero after all the calculations. This procedure is called the adiabatic approximation.

6.2

Asymptotic Field Condition

We consider an explicit representation for the interacting fields. In the relativistic action approach, the interaction Lagrangian is just minus the free Hamiltonian in the absence of derivative couplings. The fundamental interactions in the real world are governed by gauge theories. Other types of interactions appear generally as approximations in phenomenological approaches. In gauge theories, replacement of derivatives by covariant derivatives produces the interaction term and gives the source of the electromagnetic (or more generally gauge) field:

$$\bar{\psi} \gamma^\mu i \partial_\mu \psi(x) \rightarrow \bar{\psi} \gamma^\mu i D_\mu \psi(x) = \bar{\psi} \gamma^\mu i \partial_\mu \psi(x) - q \bar{\psi} \gamma^\mu \psi A_\mu \quad (6.15a)$$

$$\partial_\mu F^{\mu\nu} = q J^\nu = q \bar{\psi} \gamma^\nu \psi \quad (6.15b)$$

The interaction of quarks is governed by QCD, which is also a gauge theory with the color degree of freedom obeying $SU(3)$ internal symmetry. However, the pion, a composite of quarks, can phenomenologically be treated as a (pseudo)scalar field, which obeys the Klein–Gordon equation. Its interaction with the nucleon field ψ_N is described by the Yukawa interaction, whose interaction Lagrangian is expressed as

$$\mathcal{L}_{\text{INT}} = -g \bar{\psi}_N(x) \gamma^5 \psi_N(x) \pi(x) \quad (6.16)$$

The equation of motion can be expressed in a similar form as for the electromagnetic field. Setting $g \bar{\psi}_N(x) \gamma^5 \psi_N(x) = J$

$$K_x \phi(x) \equiv (\partial_\mu \partial^\mu + m^2)_x \phi(x) = J(x) \quad (6.17)$$

For the following discussions, we generically write the source term as J or J^μ , depending on whether it is a scalar or a vector, and do not concern ourselves with its detailed structure.

Let us solve Eq. (6.17) formally. Denoting the Green's function by $G(x - y)$, we have, by definition

$$(\partial_\mu \partial^\mu + m^2)_x G(x - y) = -\delta^4(x - y) \quad (6.18)$$

The general solution to Eq. (6.17) is the sum of the homogeneous solutions (free field) to the equation and the integral of the Green's function times the inhomogeneous term:

$$\phi(x) = \phi_{\text{free}}(x) - \int d^4 y G(x - y) J(y) \quad (6.19)$$

We have to specify the boundary conditions to determine ϕ_{free} . The Green's function itself is not unique, either. Here we choose to use the advanced and retarded Green's function

$$\Delta_{\text{ret}}(x - y) = 0 \quad \text{for} \quad x^0 - y^0 < 0 \quad (6.20a)$$

$$\Delta_{\text{adv}}(x - y) = 0 \quad \text{for} \quad x^0 - y^0 > 0 \quad (6.20b)$$

Then the second term in Eq. (6.19) vanishes as $t \rightarrow \mp\infty$, depending on whether we use the retarded or advanced functions. Therefore, we may identify the free field as $\phi_{\text{in}}^{\text{out}}$, which are generators of the “in” and “out” states:

$$\phi(x) = \phi_{\text{in}}(x) + \int d^4 y \Delta_{\text{ret}}(x - y) J(y) \quad (6.21a)$$

$$\phi(x) = \phi_{\text{out}}(x) + \int d^4 y \Delta_{\text{adv}}(x - y) J(y) \quad (6.21b)$$

The “in” and “out” fields are related by

$$\phi_{\text{in}} = S \phi_{\text{out}} S^{-1} \quad (6.22)$$

This can be easily seen as

$$a_{\text{in}}^\dagger(p) = S a_{\text{out}}^\dagger(p) S^{-1} \quad a_{\text{out}}(p') = S^{-1} a_{\text{in}}(p') S \quad (6.23)$$

and

$$|p_{\text{in}}\rangle \sim a_{\text{in}}^\dagger(p)|0\rangle = S a_{\text{out}}^\dagger(p) S^{-1}|0\rangle = S a_{\text{out}}^\dagger(p)|0\rangle \sim S|p_{\text{out}}\rangle \quad (6.24)$$

where $\sim$ means “equal” except for the normalization. We also assumed the stability of the vacuum, namely $S|0\rangle = |0\rangle$ apart from the phase, which can be set as 1.

Since $\phi(x)$ approaches $\phi_{\text{in}}^{\text{out}}(x)$ asymptotically in the limit $t \rightarrow \pm\infty$, we are tempted to think

$$\phi(x) \xrightarrow{t \rightarrow \pm\infty} \sqrt{Z} \phi_{\text{in}}^{\text{out}}(x) \quad (6.25)$$

which is called the “strong” asymptotic condition. But it has been shown that if it is imposed as an operator condition, no scattering occurs. Therefore, we impose a “weak” asymptotic condition

$$\lim_{t \rightarrow \pm\infty} \langle f | \phi(x) | i \rangle = \sqrt{Z} \langle f | \phi_{\text{out}}(x) | i \rangle \quad (6.26)$$

The renormalization constant Z appears because, ϕ has ability to produce extra virtual states and reduce the probability amplitude of remaining in a state that ϕ_{free} can create. However, for most of the following discussions we shall assume $Z = 1$. We need to include Z explicitly when considering the higher order corrections, which we will treat in Chap. 8.

6.3

Explicit Form of the S-Matrix

We now want to calculate the time evolution operator. Inserting Eq. (6.10c) into Eq. (6.6d), we obtain an equation for U :

$$i \frac{\partial}{\partial t} U(t, t_0) = H_I U(t, t_0) \quad (6.27)$$

The initial condition is

$$U(t_0, t_0) = 1 \quad (6.28)$$

Using Eq. (6.28), Eq. (6.27) can be converted to an integral equation

$$U(t, t_0) = 1 - i \int_{t_0}^t dt H_I(t) U(t, t_0) \quad (6.29)$$

Expanding successively,

$$U(t, t_0) = 1 - i \int_{t_0}^t dt H_I(t) + (-i)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1) H_I(t_2) + \cdots \quad (6.30a)$$

$$\therefore S = 1 - i \int_{-\infty}^{\infty} dt H_I(t) + (-i)^2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1) H_I(t_2) + \cdots \quad (6.30b)$$

Interchanging the variables $t_1 \leftrightarrow t_2$ in the third term and changing the integration order gives

$$\begin{aligned} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1) H_I(t_2) &= \int_{-\infty}^{\infty} dt_1 \int_{t_1}^{\infty} dt_2 H_I(t_2) H_I(t_1) \\ &= \frac{1}{2} \int_{-\infty}^{\infty} dt_1 \left[\int_{-\infty}^{t_1} dt_2 H_I(t_1) H_I(t_2) + \int_{t_1}^{\infty} dt_2 H_I(t_2) H_I(t_1) \right] \\ &= \frac{1}{2} \int_{-\infty}^{\infty} dt_1 \left[\int_{-\infty}^{\infty} dt_2 T[H_I(t_1) H_I(t_2)] \right] \end{aligned} \quad (6.31)$$

where $T[\dots]$ is a time-ordered product defined as

$$T[H_I(t_1)H_I(t_2)] = \begin{cases} H_I(t_1)H_I(t_2) & t_1 > t_2 \\ H_I(t_2)H_I(t_1) & t_1 < t_2 \end{cases} \quad (6.32)$$

Expressing H_I in terms of the Hamiltonian density

$$H_I(t) = \int d^3x \mathcal{H}(t, \mathbf{x}) \equiv \int d^3x \mathcal{H}(x) \quad (6.33)$$

The S matrix can be expressed as

$$\begin{aligned} S &= 1 - i \int d^4x \mathcal{H}(x) + \frac{(-i)^2}{2!} \iint d^4x_1 d^4x_2 T[\mathcal{H}_I(x_1)\mathcal{H}_I(x_2)] + \dots \\ &\quad + \frac{(-i)^n}{n!} \iint \dots \int d^4x_1 d^4x_2 \dots d^4x_n T[\mathcal{H}_I(x_1)\mathcal{H}_I(x_2)\dots\mathcal{H}_I(x_n)] + \dots \\ &\equiv T \exp \left[-i \int d^4x \mathcal{H}_I(x) \right] \end{aligned} \quad (6.34)$$

Note that the last equation is a formal expression. Its real meaning is given by the preceding equation. As we can see from Eq. (6.34), the integration is carried out over all space-time and $\mathcal{H}_I = -\mathcal{L}_I$ if the interaction does not contain time derivatives, and hence relativistically invariant. The effect of multiplying by T is a bit uncertain, but it does not destroy Lorentz invariance either, if the operator commutator vanishes in the space-like distance.

Next we define the transition operator T and the Lorentz-invariant matrix element $\mathcal{M}_{fi}$ by

$$\begin{aligned} \langle f|S|i\rangle &= \delta_{fi} - i2\pi\delta(E_i - E_f)\langle f|T|i\rangle \\ &= \delta_{fi} - i(2\pi)^4\delta^4(P_i - P_f)\langle f|\mathcal{M}|i\rangle \end{aligned} \quad (6.35)$$

where P_i and P_f represent the sum of the initial- and final-state energy-momentum, respectively. Although we obtained a formal solution Eq. (6.34) for the scattering matrix, expansion up to second order is enough for most of our later treatments. In the nonrelativistic treatment, we may extract and separate time dependence explicitly.

$$\langle f|H_I(t)|i\rangle = \langle f|e^{iH_0t}H_{IS}e^{-iH_0t}|i\rangle = \langle f|H_{IS}|i\rangle e^{-i(E_i-E_f)t} \quad (6.36)$$

Inserting the above equation into Eq. (6.30), we obtain an equivalent formula to Eq. (6.35) as

$$\langle f|S|i\rangle = \delta_{fi} - 2\pi i\delta(E_i - E_f)\langle f|T(E = E_i)|i\rangle \quad (6.37a)$$

$$\langle f|T(E_i)|i\rangle = \langle f|H_{IS}|i\rangle + \sum_n \frac{\langle f|H_{IS}|n\rangle\langle n|H_{IS}|i\rangle}{E_i - E_n + i\epsilon} + \dots \quad (6.37b)$$

where the adiabatic approximation $H_I^\varepsilon = H_I e^{-\varepsilon|t|}$ was used in the calculation. Thus we have re-derived Eq. (6.1b). We see the adiabatic approximation gives a prescription for handling the singularity at $E_i = E_n$. $T(E)$ can be expressed in operator form as

$$T(E) = H_I + H_I \frac{1}{E - H_0 + i\varepsilon} H_I + H_I \frac{1}{E - H_0 + i\varepsilon} H_I \frac{1}{E - H_0 + i\varepsilon} H_I + \cdots \quad (6.38)$$

where we have used H_I for $H_I(0) = H_{IS}$. This equation is the basis for the non-relativistic treatment of perturbation theory, but sometimes it is more useful than completely relativistic equations.

Problem 6.1

Replacing H_I with H_I^ε of Eq. (6.14), derive Eq. (6.37) from Eq. (6.30).

6.3.1

Rutherford Scattering

Before using the relativistic scattering amplitude, let us reproduce the Rutherford scattering formula for later comparison. This is scattering of a charged particle having spin 0 by the Coulomb potential created by a scattering center. The matrix element in Eq. (6.37) is expressed as

$$\langle f | H_{IS} | i \rangle = \int d^3x \psi_f^*(x) H_{IS} \psi_i(x) \quad (6.39a)$$

$$H_{IS} = \frac{Ze^2}{4\pi} \frac{1}{r} \quad (6.39b)$$

If the 3-momentum of the incoming and outgoing particles are assumed to be $\mathbf{p}_i$ and $\mathbf{p}_f$, respectively, the wave functions are given by

$$\psi_i = N e^{i\mathbf{p}_i \cdot \mathbf{x}}, \quad \psi_f = N e^{i\mathbf{p}_f \cdot \mathbf{x}} \quad (6.40)$$

where $N = 1/\sqrt{V}$ is a normalization factor, normalized as one particle per unit volume. Inserting Eqs. (6.39b) and (6.40) into Eq. (6.39a)

$$\langle f | H | i \rangle = \frac{Ze^2}{4\pi V} \int d^3x \frac{e^{i(\mathbf{p}_i - \mathbf{p}_f) \cdot \mathbf{x}}}{r}, \quad r = |\mathbf{x}| \quad (6.41)$$

Putting $\mathbf{q} = \mathbf{p}_i - \mathbf{p}_f$ and carrying out the integration we find

$$\begin{aligned} I \equiv \int d^3x \frac{e^{i\mathbf{q} \cdot \mathbf{x}}}{r} &= 2\pi \int_0^\infty dr \int_{-1}^1 dz \frac{e^{iqrz}}{r} r^2 \\ &= \frac{2\pi}{iq} \int_0^\infty dr (e^{iqr} - e^{-iqr}) \end{aligned} \quad (6.42a)$$

The integral diverges formally, but we may consider that the potential has a damping factor $e^{-\mu r}$ due to the screening effect of the surrounding matter. Then

$$\begin{aligned} I &= \frac{2\pi}{iq} \int_0^\infty dr \left(e^{-\mu r + iqr} - e^{-\mu r - iqr} \right) = \frac{2\pi}{iq} \left(\frac{1}{\mu - iq} - \frac{1}{\mu + iq} \right) \\ &= \frac{4\pi}{q^2 + \mu^2} \rightarrow \frac{4\pi}{q^2} \end{aligned} \quad (6.42b)$$

Inserting Eq. (6.42b) into Eq. (6.41)

$$\langle f | H | i \rangle = \left(\frac{Ze^2}{4\pi V} \right) \frac{4\pi}{q^2} = \frac{Z\alpha}{V} \frac{4\pi}{q^2}, \quad \alpha = \frac{e^2}{4\pi} \simeq \frac{1}{137} \quad (6.43a)$$

$$q^2 = |\mathbf{p}_i - \mathbf{p}_f|^2 = 4p^2 \sin^2 \frac{\theta}{2} \quad (6.43b)$$

where θ is the angle between $\mathbf{p}_i$ and $\mathbf{p}_f$, and $|\mathbf{p}_i| = |\mathbf{p}_f| = p$ by energy conservation. The final state density is calculated to be

$$\rho = \delta(E_i - E_f) \frac{V d^3 p}{(2\pi)^3} = V \left[\frac{dE}{dp} \right]^{-1} \frac{p^2 d\Omega}{8\pi^3} = V \frac{m p d\Omega}{8\pi^3} \quad (6.44)$$

where we used $E = p^2/2m$ and integrated over E . Inserting Eqs. (6.43) and (6.44) into Eq. (6.1a)

$$W_{fi} = 2\pi \left(\frac{Z\alpha}{V} \right)^2 \left(\frac{4\pi}{q^2} \right)^2 V \frac{m p d\Omega}{8\pi^3} = \frac{p}{V} \frac{4Z^2 \alpha^2 m}{q^4} d\Omega \quad (6.45)$$

To convert the transition probability into a cross section, it has to be divided by incoming flux. The flux is the number of particles passing through unit area in unit time. When the wave function is normalized to one particle per unit volume, the incoming flux can be given as v/V , where $v = p/m$ is the velocity of the particle,

$$\therefore \frac{d\sigma}{d\Omega} = \frac{4Z^2 \alpha^2 m^2}{q^4} = \frac{Z^2 \alpha^2 m^2}{4p^4 \sin^4(\theta/2)} \quad (6.46)$$

The normalization volume disappears from the final expression as it should. This is the famous Rutherford scattering formula and coincides with that derived classically.

6.4

Relativistic Kinematics

6.4.1

Center of Mass Frame and Laboratory Frame

In dealing with reactions of elementary particles, there are two commonly used reference frames. For theoretical treatments and for collider accelerator experiments,

the center of mass (CM) frame, in which total momentum is zero, is used. Before the advent of colliders, most experiments were of the type that a beam extracted from an accelerator hits a target at rest. This is referred to as the laboratory (LAB for short) frame, and most nuclear and some particle experiments today are of this type. They are connected by Lorentz transformation and relations that will become useful later are derived here. Let us consider two-body reactions $1 + 2 \rightarrow 3 + 4$. We assign numbers to particles as follows: incoming (1), target (2), scattered (3) and recoiled (4), and denote 4-momenta and masses as $p_1, p_2, p_3, p_4; m_1, m_2, m_3, m_4$. We indicate variables in the CM frame by an asterisk. In the laboratory frame, they are expressed as

$$p_1 = (E_1, \mathbf{p}_1), \quad p_2 = (m_2, \mathbf{0}), \quad p_3 = (E_3, \mathbf{p}_3), \quad p_4 = (E_4, \mathbf{p}_4) \quad (6.47)$$

and in the CM frame

$$\begin{aligned} p_1^* &= (E_1^*, \mathbf{p}^*), & p_2^* &= (E_2^*, -\mathbf{p}^*), & p_3^* &= (E_3^*, \mathbf{p}'^*), \\ p_4^* &= (E_4^*, -\mathbf{p}'^*) \end{aligned} \quad (6.48)$$

Graphically they are expressed as in Fig. 6.1. By energy-momentum conservation

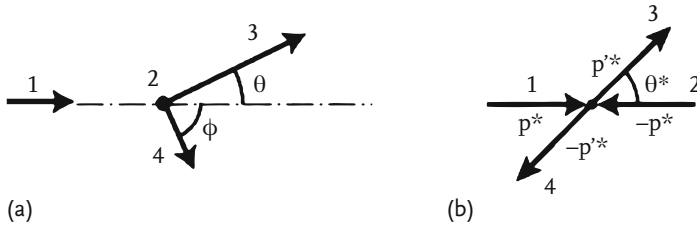


Figure 6.1 Two-body scattering in (a) the laboratory frame and (b) the center of mass frame.

$$p_1 + p_2 = p_3 + p_4 \quad (6.49)$$

Consider the $1 + 2$ particle system as composing a single particle having $E_{\text{tot}} = E_1 + E_2$ and $\mathbf{p}_{\text{tot}} = \mathbf{p}_1 + \mathbf{p}_2$, then in the CM frame the energy $W = E_{\text{tot}}^*$ can be considered as the mass of the two-body system because $\mathbf{p}_{\text{tot}}^* = \mathbf{0}$ by definition. The Lorentz factors β and γ to transform the LAB frame to the CM frame can be obtained by applying one-body formulae ($\beta = p/E$, $\gamma = E/m$) to the system, i.e.

$$\beta = \frac{p_{\text{tot}}}{E_{\text{tot}}} = \frac{p_1}{E_1 + m_2}, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}} = \frac{E_1 + m_2}{W} \quad (6.50a)$$

$$W = (m_1^2 + 2E_1m_2 + m_2^2)^{1/2} \quad (6.50b)$$

Now that we know the Lorentz factors to connect variables in the two frames, we can apply the standard Lorentz transformation formula to each particle:

$$E_{\text{LAB}} = \gamma (E^* + \beta p_{\parallel}^*) \quad (6.51a)$$

$$p_{\parallel \text{ LAB}} = \gamma (\beta E^* + p_{\parallel}^*) \quad (6.51b)$$

$$p_{\perp \text{ LAB}} = p_{\perp}^* \quad (6.51c)$$

where $p_{\parallel}$, $p_{\perp}$ are components of the momentum along and perpendicular to the z-axis (direction of CM frame motion):

$$p_{\parallel \text{ LAB}} = p_{\text{LAB}} \cos \theta, \quad p_{\perp \text{ LAB}} = p_{\text{LAB}} \sin \theta \quad (6.52a)$$

$$p_{\parallel}^* = p^* \cos \theta^*, \quad p_{\perp}^* = p^* \sin \theta^* \quad (6.52b)$$

An important point here is that the transverse components are not affected by the Lorentz transformation. This fact is extensively utilized in analyzing particles, for example sorting b-quarks out of many-particle jets.²⁾

Problem 6.2

Prove that particles that are distributed isotropically in the CM frame ($dN/d\Omega^* = \text{constant}$), are distributed uniformly in energy in the LAB frame ($dN/dE = \text{constant}$).

Although Eqs. (6.52) are the bases for connecting variables in the two frames, it is more convenient in actual calculations to use the Lorentz-invariant Mandelstam variables, defined as follows:

$$s \equiv (p_1 + p_2)^2 = (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 \quad (6.53a)$$

$$t \equiv (p_1 - p_3)^2 = (E_1 - E_3)^2 - (\mathbf{p}_1 - \mathbf{p}_3)^2 \quad (6.53b)$$

$$u \equiv (p_1 - p_4)^2 = (E_1 - E_4)^2 - (\mathbf{p}_1 - \mathbf{p}_4)^2 \quad (6.53c)$$

As they are Lorentz scalars, their values do not change between frames. For instance

$$s = \begin{cases} (E_1^* + E_2^*)^2 & : \text{CM} \\ (E_1 + m)^2 - \mathbf{p}_1^2 = m_1^2 + 2E_1 m_2 + m_2^2 & : \text{LAB} \end{cases} \quad (6.54)$$

As $s = W^2$, s can be considered as the invariant mass (squared) of the two-body system. Writing s in terms of CM variables

$$\sqrt{s} = W = E_1^* + E_2^* = \sqrt{\mathbf{p}^{*2} + m_1^2} + \sqrt{\mathbf{p}^{*2} + m_2^2} \quad (6.55)$$

- 2) B hadrons, which contain a b (or $\bar{b}$) quark, have mass around $m_B \sim 5 \text{ GeV}$. When they are produced in high-energy collisions, they typically appear as jets in which many hadrons fly more or less in the same direction having high longitudinal momentum but limited transverse momentum $p_T < m_B/2$ relative to the jet axis.

we can solve the CM energy in terms of s

$$E_1^* = \frac{s + m_1^2 - m_2^2}{2\sqrt{s}}, \quad E_2^* = \frac{s - m_1^2 + m_2^2}{2\sqrt{s}} \quad (6.56a)$$

$$p^* = \sqrt{E_1^{*2} - m_1^2} = \frac{\lambda(s, m_1^2, m_2^2)}{2\sqrt{s}} = \frac{\sqrt{s}}{2} \lambda(1, x_1, x_2) \quad (6.56b)$$

$$\lambda(x, y, z) = (x^2 + y^2 + z^2 - 2xy - 2yz - 2zx)^{1/2} \quad (6.56c)$$

$$x_1 = m_1^2/s, \quad x_2 = m_2^2/s \quad (6.56d)$$

$|t|$ is referred to as the momentum transfer (squared) and can be written down similarly

$$t = (E_1 - E_3)^2 - (\mathbf{p}_1 - \mathbf{p}_3)^2 = m_1^2 + m_3^2 - 2(E_1 E_3 - |\mathbf{p}_1||\mathbf{p}_3| \cos \theta) \quad : \text{LAB} \quad (6.57)$$

θ is the scattering angle and is defined by $\cos \theta = (\mathbf{p}_2 \cdot \mathbf{p}_3)/|\mathbf{p}_2||\mathbf{p}_3|$. In most high-energy experiments we are interested in $E \gg m$, $E \simeq p = |\mathbf{p}|$, then

$$t \simeq -2p_1 p_3 (1 - \cos \theta) \simeq -4E_1 E_3 \sin^2 \frac{\theta}{2} \quad (6.58)$$

t has the same form either in LAB or in CM. The role of u is similar to t , except particle 3 is replaced by particle 4. In CM, t and u are expressed in the relativistic approximation ($E \gg m$),

$$t \simeq -2p_1^* p_3^* (1 - \cos \theta) \quad (6.59a)$$

$$u \simeq -2p_1^* p_3^* (1 + \cos \theta) \quad (6.59b)$$

From the above formula, we see that s gives the total energy, while t, u give the measure of the scattering angle. When the masses can be neglected, i.e. $s \gg m_i^2$, the ranges of the variables are $s \geq 0$, $t \leq 0$, $u \leq 0$.

Problem 6.3

Show that s, t, u are not independent but satisfy the relation

$$s + t + u = \sum_{i=1}^4 m_i^2 \quad (6.60)$$

6.4.2

Crossing Symmetry

The physical meanings of the three variables are apparently different, but considering the fact that mathematically antiparticles appear as having negative energy-

momentum and as backward-traveling particle, we can show that the roles of s , t and u are equal and symmetric. Take the example of π -proton elastic scattering (Fig. 6.2a)

$$\pi^+(p_1) + p(p_2) \rightarrow \pi^+(p_3) + p(p_4) \quad (6.61)$$

By changing the variables

$$p_1 \rightarrow -p_1, \quad p_4 \rightarrow -p_4 \quad (6.62)$$

a process we call “crossing”, the physical meaning of the scattering is modified to that of Fig. 6.2b

$$\bar{p}(p_4) + p(p_2) \rightarrow \pi^-(p_1) + \pi^+(p_3) \quad (6.63)$$

Because by the change in Eq. (6.62) the Mandelstam variables are changed to

$$s = (p_1 + p_2)^2 = (-p_4 - p_3)^2 \rightarrow u_t \equiv (p_4 - p_3)^2 \sim -2p_t^2(1 + \cos \theta_t) \quad (6.64a)$$

$$t = (p_1 - p_3)^2 = (-p_4 + p_2)^2 \rightarrow s_t \equiv (p_4 + p_2)^2 \sim 4p_t^2 \quad (6.64b)$$

$$u = (p_1 - p_4)^2 \rightarrow t_t \equiv (p_4 - p_1)^2 \sim -2p_t^2(1 - \cos \theta_t) \quad (6.64c)$$

p_t represents the momentum in the $\bar{p}p$ CM frame, θ_t is the angle that the outgoing π^- makes with the incoming antiproton $\bar{p}$. Here, $t \geq 0$, $s, u \leq 0$ and $t \equiv s_t$, $u \equiv t_t$, $s \equiv u_t$ play the role of s, t, u in reaction (6.61). The reactions (6.61) and (6.63) are referred to as s-channel and t-channel reactions, respectively. Similarly, the u-channel reaction (Fig. 6.2c) can be defined. Thus, the three variables s, t, u can be negative as well as positive and, depending on their value, represent three different reactions. This means that if the scattering amplitudes are written in terms of s, t, u , other crossed-channel reactions are obtained immediately by crossing. Figure 6.3 shows the three physical regions in an oblique coordinate frame, for the case where the pion mass is μ and the proton mass M .

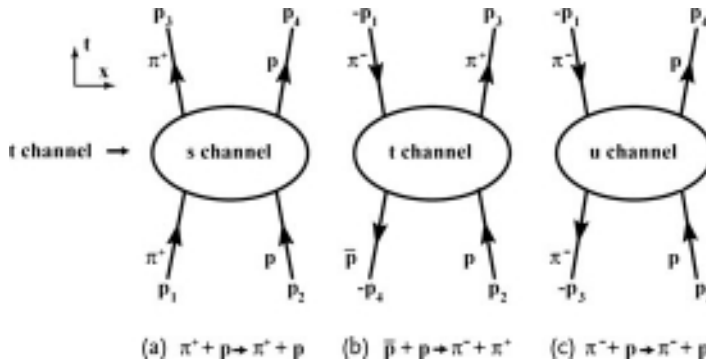


Figure 6.2 Crossing symmetry. Mutual relations of three different scatterings.

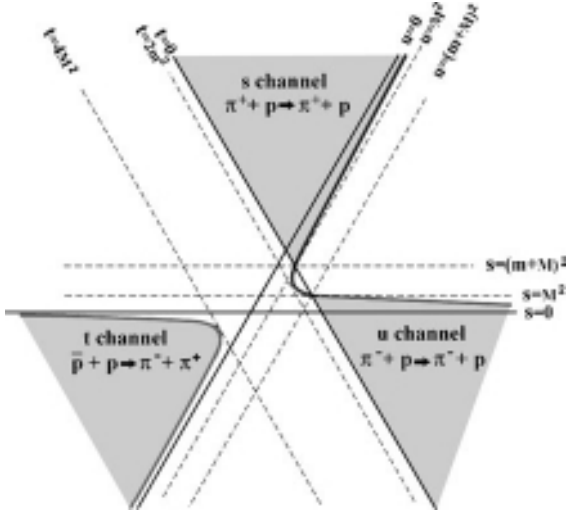


Figure 6.3 Three physical regions of s, t, u in an oblique coordinate frame. A point in the plane represents (s, t, u) by its signed perpendicular length to the $s = 0, t = 0, u = 0$ lines.

6.5

Relativistic Cross Section

6.5.1

Transition Rate

We first derive the relativistic version of Eq. (6.1a) ($W_{fi} = 2\pi |T_{fi}|^2 \rho$). The mean transition probability per unit time is given by Eq. (6.13), so referring to Eq. (6.35)

$$\begin{aligned}
 |S_{fi} - \delta_{fi}|^2 &= |-(2\pi)^4 i \delta^4(p_i - p_f) \mathcal{M}_{fi}|^2 \\
 &= (2\pi)^4 \delta^4(p_i - p_f) \left[\int d^4x e^{i(p_i - p_f) \cdot x} \right] |\mathcal{M}_{fi}|^2 \\
 &= (2\pi)^4 \delta^4(p_i - p_f) \left[\int d^4x \right] |\mathcal{M}_{fi}|^2 \\
 &= (2\pi)^4 \delta^4(p_i - p_f) VT |\mathcal{M}_{fi}|^2
 \end{aligned} \tag{6.65}$$

Here, the δ -function was first expressed as an integral and we used the fact that the total integrated volume VT is large on microscopic scale but finite on macroscopic scale. The transition rate per unit time per unit volume is obtained by dividing W_{fi} by VT and multiplying by the final state density ρ_f :

$$W_{fi} = (2\pi)^4 \delta^4(p_i - p_f) |\langle f | \mathcal{M} | i \rangle|^2 \rho_f \tag{6.66}$$

Then the cross section is obtained by dividing W_{fi} by the incoming flux and target particle density.

6.5.2

Relativistic Normalization

In the relativistic treatment, the usual normalization condition

$$\int \rho(x) dV = 1 \quad (6.67)$$

is unsatisfactory. For flying particles the volume element dV is Lorentz contracted, and the space integral in the rest frame overestimates the volume by $\gamma = E/m$. It is better to adopt a normalization that is proportional to γ . Therefore, starting from this section, we adopt relativistic normalization, which we choose to set

$$\int \rho dV = 2E \quad (6.68)$$

Writing plane-wave solutions obeying the Klein–Gordon equation as $\varphi = N e^{-ip \cdot x}$ and those obeying the Dirac equation as $\psi = N u_r(p) e^{-ip \cdot x}$, the normalization conditions for both of them are

$$\int \rho d^3x \stackrel{\text{KG}}{=} \int d^3x i [\varphi^* \dot{\varphi} - \dot{\varphi}^* \varphi] = 2E N^2 V \quad \text{boson} \quad (6.69a)$$

$$\stackrel{\text{Dirac}}{=} \int d^3x \psi^* \psi = N^2 V u_r^\dagger(p) u_r(p) = 2E N^2 V \quad \text{fermion} \quad (6.69b)$$

In both cases, they give $N = 1/\sqrt{V}$. The condition Eq. (6.68) means the wave functions of the Dirac and Klein–Gordon fields are normalized as

$$f_p(x) = \frac{1}{\sqrt{V}} e^{-ip \cdot x} \quad \text{boson} \quad (6.70a)$$

$$U_{pr} = u_r(p) f_p(x) \quad \text{fermion} \quad (6.70b)$$

In the above treatment we integrated the density over a large but finite volume $V = L^3$, which in turn means that the momentum is discrete. If we impose periodic boundary conditions for the plane wave,

$$e^{ip_x(x+L)} = e^{ip_x x}, \rightarrow p_x L = 2\pi n_x, \quad n_x = 0, \pm 1, \pm 2, \dots, \quad (6.71)$$

$$\therefore \Delta n_x = \frac{L}{2\pi} \Delta p_x, \quad \Delta n_y = \frac{L}{2\pi} \Delta p_y, \quad \Delta n_z = \frac{L}{2\pi} \Delta p_z$$

then the number of wave states in the volume is given by

$$\mathbf{p} = (p_x, p_y, p_z), \quad p_i L = 2\pi n_i \quad n_i = 0, \pm 1, \pm 2, \dots \quad (6.72a)$$

$$\Delta \mathbf{n} = \left(\frac{L}{2\pi} \right)^3 \Delta \mathbf{p} \quad (6.72b)$$

The discrete sum becomes an integral in the large volume limit:

$$\sum_{\mathbf{p}} = \sum_{\mathbf{n}} \Delta \mathbf{n} = \left(\frac{L}{2\pi} \right)^3 \sum_{\mathbf{p}} \Delta \mathbf{p} \rightarrow V \int \frac{d^3 p}{(2\pi)^3} \quad (6.72c)$$

The integral of the plane waves in discrete or continuous format is

$$\frac{1}{V} \int d^3x e^{i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} = \delta_{\mathbf{p},\mathbf{p}'} \xrightarrow{V \rightarrow \infty} \frac{(2\pi)^3}{V} \delta^3(\mathbf{p} - \mathbf{p}') \quad (6.73)$$

Proof:

$$\begin{aligned} \frac{1}{L} \int_{-L/2}^{L/2} e^{i(p_x - p'_x)x} dx &= \begin{cases} 1; & (p_x = p'_x) \\ \frac{2}{L(p_x - p'_x)} \sin \frac{L}{2}(p_x - p'_x) = 0; & (p_x \neq p'_x) \end{cases} \\ \therefore p_x L = 2\pi n, \quad (n = 0, \pm 1, \pm 2 \dots) &\quad (6.74) \end{aligned}$$

The one-particle state orthogonality and completeness conditions are now

$$\begin{aligned} \langle \mathbf{p}' | \mathbf{p} \rangle &= \frac{(2\pi)^3}{V} 2E_p \delta^3(\mathbf{p}' - \mathbf{p}) \\ 1 &= V \int \frac{d^3p}{(2\pi)^3 2E_p} |\mathbf{p}\rangle \langle \mathbf{p}| \end{aligned} \quad (6.75)$$

The commutator of the discrete creation/annihilation operators becomes in the continuous limit

$$[a_{\mathbf{p}r}, a_{\mathbf{p}'s}^\dagger]_{\pm} = \frac{(2\pi)^3}{V} \delta_{rs} \delta^3(\mathbf{p} - \mathbf{p}') \quad (6.76)$$

Then the one-particle state is expressed in terms of the creation operator as

$$|\mathbf{p}\rangle = \sqrt{2E} a^\dagger(\mathbf{p}) |0\rangle \quad (6.77)$$

Having defined normalizations with finite V , we set the volume as 1 from now on unless otherwise noted, because they always disappear from the final expression for measurable quantities.³⁾

With the above definitions, the expansion formulae for the fields now become

$$\begin{aligned} \varphi(x) &= \sum_k \frac{1}{\sqrt{2\omega}} [a_k e^{-ik\cdot x} + b_k^\dagger e^{ik\cdot x}] \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^3 \sqrt{2\omega}} [a_k e^{-ik\cdot x} + b_k^\dagger e^{ik\cdot x}] \quad : \text{scalar} \end{aligned} \quad (6.78a)$$

$$\psi(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3 \sqrt{2E_p}} \sum_r [a_p u_r(p) e^{-ip\cdot x} + b_p^\dagger v(p)_r e^{ip\cdot x}] \quad : \text{Dirac} \quad (6.78b)$$

$$\begin{aligned} A^\mu(x) &= \int \frac{d^3\mathbf{k}}{(2\pi)^3 \sqrt{2\omega}} \\ &\quad \sum_\lambda [a_{k\lambda} \varepsilon^\mu(k, \lambda) e^{-ik\cdot x} + a_{k\lambda}^\dagger \varepsilon^\mu(k, \lambda)^* e^{ik\cdot x}] \quad : \text{photon} \end{aligned} \quad (6.78c)$$

3) We shall prove this for the case of the cross section in the following section.

The matrix elements of the field between the vacuum and one-particle state appear frequently and we give them here:

$$\langle 0|\varphi(x)|k\rangle = e^{-ik\cdot x} \quad \langle k|\varphi^\dagger(x)|0\rangle = e^{ik\cdot x} \quad (6.79a)$$

$$\langle 0|\psi(x)|p, r\rangle = u_r(p)e^{-ip\cdot x} \quad \langle p|\bar{\psi}(x)|0\rangle = \bar{u}_r(p)e^{ip\cdot x} \quad (6.79b)$$

$$\langle 0|A^\mu(x)|k\lambda\rangle = \varepsilon^\mu(k, \lambda)e^{-ik\cdot x} \quad \langle k|A^\mu(x)|0\rangle = \varepsilon^\mu(k, \lambda)^*e^{ik\cdot x} \quad (6.79c)$$

Note As we said, taking the continuous limit does not necessarily mean V is infinite. Therefore $\delta(0)$ should be interpreted as

$$\delta^3(0) = \frac{1}{(2\pi)^3} \int_V d^3x = \frac{V}{(2\pi)^3} \quad (6.80)$$

Then

$$\langle p|p\rangle = \frac{(2\pi)^3}{V} 2E\delta^3(0) = 2E \quad (6.81)$$

6.5.3

Incoming Flux and Final State Density

In the relativistic normalization, the final state density of one particle is given by

$$d\rho_f = \frac{V}{(2\pi)^3} \frac{d^3p}{2E} \quad (6.82)$$

for each particle in the final state. Let us name the incoming particle 1, the target particle 2 and assume N_f particles are produced in the final state. The incoming flux is given by $2E_1v_{12}/V$, where v_{12} is the incoming velocity relative to the target, and the target particle density is given by $2E_2/V$. Then the cross section for the reaction becomes

$$d\sigma = \frac{V^2}{2E_1 2E_2 v_{12}} |\mathcal{M}_{fi}|^2 (2\pi)^4 \delta^4 \left(p_1 + p_2 - \sum_{f=1}^{N_f} p_f \right) \prod_{f=1}^{N_f} \left(\frac{V d^3p_f}{(2\pi)^3 2E_f} \right) \quad (6.83)$$

We define Lorentz-invariant flux F by

$$\begin{aligned} F &\equiv 4E_1 E_2 v_{12} = 4E_1 E_2 \left(\frac{p_1}{E_1} - \frac{p_2}{E_2} \right) = \begin{cases} 4p^* W & \text{CM} \\ 4p_1 m_2 & \text{LAB} \end{cases} \\ &= 4[(p_1 \cdot p_2)^2 - m_1^2 m_2^2]^{1/2} = 2s\lambda(1, x_1, x_2) \\ \lambda(x, y, z) &= (x^2 + y^2 + z^2 - 2xy - 2yz - 2zx)^{1/2} \\ x_1 &= m_1^2/s, \quad x_2 = m_2^2/s \end{aligned} \quad (6.84)$$

As the state vector contains a factor $1/\sqrt{V}$ for each particle, $\mathcal{M}_{fi}$ produces a total of $(N_f + 2) 1/\sqrt{V}$. Then naturally, all the volume V 's are canceled and disappear from the expression of $d\sigma$. As V always disappears from the final expression for measurable quantities, there is no harm in setting $V = 1$.

6.5.4

Lorentz-Invariant Phase Space

We introduce Lorentz-invariant phase space by

$$dLIPS \equiv (2\pi)^4 \delta^4 \left(\sum_i p_i - \sum_f p_f \right) \frac{1}{m!} \prod_f \frac{d^3 p_f}{(2\pi)^3 2E_f} \quad (6.85)$$

where $m!$ is a correction factor if identical particles are included in the state. Inserting Eqs. (6.84) and (6.85) into Eq. (6.83), we obtain the cross-section formula

$$d\sigma = \frac{|\mathcal{M}_{fi}|^2}{2s\lambda(1, x_1, x_2)} dLIPS \quad (6.86)$$

The decay rate for particle 2 into many particles $d\Gamma(2 \rightarrow 3 + 4 + \dots)$ is obtained by dividing the transition rate $W_{fi} = W(2 \rightarrow 3 + 4; \dots)$ by the flux of particle 2 in its rest frame

$$d\Gamma = \frac{|\mathcal{M}_{fi}|^2}{2m_2} dLIPS \quad (6.87)$$

Equations (6.86) and (6.87) are the relativistic version of Fermi's golden rule Eq. (6.1a).

6.5.5

Cross Section in the Center of Mass Frame

Let us calculate the cross section for two-body scattering

$$1 + 2 \rightarrow 3 + 4$$

in the CM frame. Using Eq. (6.83), the cross section for two-body scattering is given by

$$d\sigma = \frac{1}{4E_1 E_2 v_{12}} |\mathcal{M}_{fi}|^2 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \frac{d^3 p_3}{(2\pi)^3 2E_3} \frac{d^3 p_4}{(2\pi)^3 2E_4} \quad (6.88)$$

In CM, $\mathbf{p}_1 + \mathbf{p}_2 = 0$, $E_1 + E_2 = W$. Writing $p_i = |\mathbf{p}_1| = |\mathbf{p}_2|$, $p_f = |\mathbf{p}_3| = |\mathbf{p}_4|$,

$$v_{12} = \left| \frac{\mathbf{p}_1}{E_1} - \frac{\mathbf{p}_2}{E_2} \right| = \frac{p_i(E_1 + E_2)}{E_1 E_2} = \frac{p_i W}{E_1 E_2} \quad (6.89a)$$

$$\therefore F = 4E_1 E_2 v_{12} = 4p_i W \quad (6.89b)$$

$$dLIPS_{\text{CM}} = \frac{1}{16\pi^2} \delta(W - E_3 - E_4) \delta^3(\mathbf{p}_3 + \mathbf{p}_4) \frac{d^3 p_3}{E_3} \frac{d^3 p_4}{E_4} \quad (6.89c)$$

$\int d^3 p_4$ gives 1. Then to carry out p_3 integration, we use the relation $\delta(G) dp_f = (\partial G / \partial p_f)^{-1} \delta(G) dG$:

$$G = E_3 + E_4 - W = \sqrt{p_f^2 + m_3^2} + \sqrt{p_f^2 + m_4^2} - W \quad (6.89d)$$

$$\frac{\partial G}{\partial p_f} = \frac{p_f}{E_3} + \frac{p_f}{E_4} = \frac{p_f W}{E_3 E_4} \quad (6.89e)$$

$$\therefore \frac{dLIPS}{F} = \frac{1}{4p_i W} \left(\frac{\partial G}{\partial p_f} \right)^{-1} \frac{p_f^2 d\Omega_{\text{CM}}}{16\pi^2 E_3 E_4} = \frac{1}{64\pi^2 s} \left(\frac{p_f}{p_i} \right) d\Omega_{\text{CM}} \quad (6.89f)$$

We obtain

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} = \left(\frac{p_f}{p_i} \right) \frac{|\mathcal{M}_{fi}|^2}{64\pi^2 s} \quad (6.90)$$

Problem 6.4

Prove that the cross section in LAB for the case $m_1 = m_3 = 0$, $m_2 = m_4 = M$ is given by

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{LAB}} = \left(\frac{p_3}{p_1} \right)^2 \frac{|\mathcal{M}_{fi}|^2}{64\pi^2 M^2} \quad (6.91)$$

Problem 6.5

Sometimes it is useful to express the cross section in terms of the Lorentz-invariant variable t rather than the scattering angle $\cos \theta$. Assuming $m_1 = m_3 = m$, $m_2 = m_4 = M$, show that

$$dt = 2p_1 p_3 d\cos \theta = \frac{p_1 p_3}{\pi} d\Omega \simeq \frac{(s - M^2)^2}{4\pi s} d\Omega_{\text{CM}} \quad (6.92a)$$

$$\frac{d\sigma}{dt} = \frac{\pi}{p_1 p_3} \left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} = \frac{|\mathcal{M}_{fi}|^2}{16\pi \lambda(s, m^2, M^2)^2} \simeq \frac{|\mathcal{M}_{fi}|^2}{16\pi (s - M^2)^2} \quad (6.92b)$$

where the last equation in both lines holds when $s \gg m^2$.

Hint: Use

$$t = q^2 = (p_1 - p_3)^2 = 2m^2 - 2(E_1 E_3 - p_1 p_3 \cos \theta) \quad \text{in any frame} \quad (6.93)$$

and in CM

$$p_1 = p_3 = \frac{\sqrt{s}}{2} \lambda \left(1, \frac{m^2}{s}, \frac{M^2}{s} \right) \simeq \frac{s - M^2}{2\sqrt{s}} \quad (6.94)$$

6.6

Vertex Functions and the Feynman Propagator

In Sect. 6.3 we derived an explicit formula for the scattering matrix that is the relativistic version of Eq. (6.1b):

$$\langle f|T|i\rangle = \langle f|H_I|i\rangle + \sum_n \frac{\langle f|H_I|n\rangle\langle n|H_I|i\rangle}{E_i - E_n} \quad (6.95)$$

$\langle f|H_I|i\rangle$ is the first-order matrix element of the interaction Hamiltonian, but in Feynman diagrams, which will appear in abundance in the following, it is referred to as the vertex function and $1/(E_i - E_n)$ as the propagator, the relativistic version of which is referred to as the Feynman propagator. It is necessary as well as elucidating to learn beforehand how elements $\langle f|H_I|i\rangle$ and $1/(E_i - E_n)$ that appear in the above formula are expressed in quantized field theory. To calculate the vertex function, we need an explicit form of the interaction Hamiltonian. In the following, we mainly discuss the electromagnetic interaction.

6.6.1

$e e \gamma$ Vertex Function

As was described in Chap. 4, the interaction of electrons ($q = -e$) with the electromagnetic field is given by the gauge principle, which is to replace derivatives by covariant derivatives $\partial_\mu \rightarrow D_\mu = \partial_\mu - ieA_\mu$. This replacement produces the interaction Hamiltonian density $\mathcal{H}_I$ for a Dirac particle

$$\mathcal{H}_I = -\mathcal{L}_{\text{int}} = q\bar{\psi}\gamma^\mu\psi A_\mu = qj^\mu A_\mu \quad (6.96)$$

Similarly, the interaction Hamiltonian for the complex scalar field is given by

$$\begin{aligned} \mathcal{L} &= \partial_\mu\varphi^\dagger\partial^\mu\varphi - m^2\varphi^\dagger\varphi \Rightarrow \mathcal{L} + \mathcal{L}_I \\ \mathcal{H}_I &= qA_\mu [i(\varphi^\dagger\partial^\mu\varphi - \partial^\mu\varphi^\dagger\varphi)] - q^2A_\mu A^\mu\varphi^\dagger\varphi \end{aligned} \quad (6.97)$$

Consider a transition from an initial state with an electron having the energy-momentum p and spin state r to a final state with an electron having p' , s and a photon with k and polarization λ . The expectation value or the matrix element of the Hamiltonian is given by

$$\langle p's; k\lambda | \mathcal{H}_I(x) | p r \rangle = -e \langle p's | \bar{\psi}(x) \gamma^\mu \psi(x) | p r \rangle \langle k\lambda | A_\mu(x) | 0 \rangle \quad (6.98)$$

Here, we have separated the electron and photon, because their Fock spaces are independent of each other. We have also expressed the sign of the charge explicitly ($q = -e$). The expectation value of the electromagnetic field is readily obtained using Eq. (6.78c)

$$\langle k\lambda | A_\mu(x) | 0 \rangle = \varepsilon_\mu(k, \lambda)^* e^{ik \cdot x} \quad (6.99)$$

We use Eq. (6.78b) for the expression of $\psi, \bar{\psi}$. Considering $|pr\rangle \sim a_{pr}^\dagger|0\rangle$, $\langle p's| \sim \langle 0|a_{p's}$, the expectation value of the current $\sim \bar{\psi}\psi$ contains a term

$$\langle 0|a_{p'}a_s^\dagger a_t a_p^\dagger|0\rangle \quad (6.100)$$

where we have used the abbreviated notation p', p, t, s to denote both momentum and spin states:

$$\begin{aligned} a_t a_p^\dagger &= [\{a_t, a_p^\dagger\} - a_p^\dagger a_t]|0\rangle = (2\pi)^3 \delta^3(t-p)|0\rangle \\ \langle 0|a_{p'}a_s^\dagger &= \langle 0|[\{a_{p'}a_s^\dagger\} - a_s^\dagger a_{p'}] = \langle 0|(2\pi)^3 \delta^3(s-p') \end{aligned} \quad (6.101)$$

In deriving the last equation of both lines, we used the fact that the annihilation (creation) operator acting on the vacuum state from the left (right) annihilates the states. Therefore

$$\langle 0|a_{p'}a_s^\dagger a_t a_p^\dagger|0\rangle = (2\pi)^6 \delta^3(t-p)\delta^3(s-p') \quad (6.102)$$

Taking into account the normalization factor ($\sqrt{2E_p} \dots$) for the states and integrating over t, s (note we are also taking the sum over spin)

$$\langle p's|\bar{\psi}(x)\gamma^\mu\psi(x)|pr\rangle = \bar{u}_s(p')\gamma^\mu u_r(p)e^{-i(p-p')\cdot x} \quad (6.103)$$

Therefore

$$\begin{aligned} (a) \quad e &\rightarrow e + \gamma \\ \langle p's; k\lambda|\mathcal{H}_1|p\rangle &= -e\bar{u}_s(p')\gamma^\mu u_r(p)\varepsilon_\mu(k, \lambda)^* e^{-i(p-p'-k)\cdot x} \end{aligned} \quad (6.104a)$$

This matrix element corresponds to diagram (a) of Fig. 6.4. Notice the sign of the energy momentum that appears in the exponent. The sign of the energy-momentum p of an incoming and annihilated particle in the exponent is negative and that of outgoing or created particles is positive. As the interaction can happen anywhere in space-time, the matrix element will be integrated over space-time, producing δ -functions. Note, the action is an integral of the Lagrangian over all space-time. Namely, the energy and the momentum are conserved at each interaction point or at the “vertex” of the diagram:

$$E = E' + \omega, \quad p = p' + k \quad (6.104b)$$

In a similar way, the diagram in Fig. 6.4b corresponds to

$$\begin{aligned} (b) \quad \gamma + e &\rightarrow e \\ \langle p's|\mathcal{H}_1|pr; k\lambda\rangle &= -e\bar{u}_s(p')\gamma^\mu u_r(p)\varepsilon_\mu(k\lambda)e^{-i(p-p'+k)\cdot x} \end{aligned} \quad (6.104c)$$

Next, we consider the interaction of an antiparticle. Corresponding to the diagram in Fig. 6.4c,

$$\begin{aligned} (c) \quad e^+ &\rightarrow e^+ + \gamma \\ \langle \bar{p}'s; k\lambda|\mathcal{H}_1|\bar{p}r\rangle &= +e\bar{v}_r(p)\gamma^\mu v_s(p')\varepsilon_\mu(k\lambda)^* e^{-i(p-p'-k)\cdot x} \end{aligned} \quad (6.105a)$$

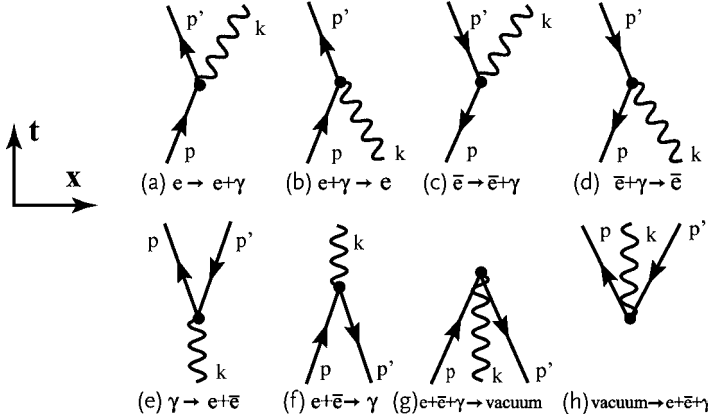


Figure 6.4 Various processes of the electron–photon interaction in the lowest order that is given by $\mathcal{L}_{\text{INT}} = -qA_\mu j^\mu$. Time goes upward; a line with an upward-going arrow designates a particle and one with a downward arrow, an antiparticle.

Here again, the sign of the energy-momentum in the exponent gives (–) for incoming and (+) for outgoing particles and there is no distinction between the particle and antiparticle. The remnant of the statement that the antiparticle goes backward in time is on the order of $\bar{v}_r(p)\gamma^\mu v_s(p')$. Notice also that the sign of the electric charge is correctly given, i.e. opposite to the electron, as it must be. The contribution to the matrix element is a product of $b_s\bar{v}(s)$, which is in the positive-frequency part of $\bar{\psi}$, and $b_t^\dagger v(t)$, which is in the negative-frequency part of ψ

$$\langle \bar{p}' | b_s b_t^\dagger | \bar{p} \rangle = \langle \bar{p}' | \{ b_s b_t^\dagger \} - b_t^\dagger b_s | \bar{p} \rangle = \langle \bar{p}' | (2\pi)^3 \delta^3(s - t) - b_t^\dagger b_s | \bar{p} \rangle$$

The second term gives the above matrix element with the correct negative sign. The negative sign results from the anticommutation relation, but it also produces the first term, an unwanted constant contribution, which gives a diverging integral. This first term can be treated in a similar manner as the constant term in the Hamiltonian, namely such an unphysical term is to be subtracted at the very beginning of the calculation. The technique is to assume that terms like $b_s b_t^\dagger$ in the bilinear expression $\bar{\psi}\psi$, in which the annihilation operator comes to the left of the creation operator, are already ordered like $-b_t^\dagger b_s$ from the beginning. Operators arranged in such a manner are called “normal ordered” or “N product” in quantized field theory. We shall assume that the interaction Hamiltonian is normal ordered.⁴⁾ The diagram in Fig. 6.4d is quite similar to Fig. 6.4b with appropriate changes ($u_r(p) \rightarrow v_s(p')$, $\bar{u}_s(p') \rightarrow \bar{v}_r(p)$, $-e \rightarrow +e$)

$$(d) \gamma + e^+ \rightarrow e^+$$

$$\langle \bar{p}' s | \mathcal{H}_1 | \bar{p} r; k \lambda \rangle = +e \bar{v}_r(p) \gamma^\mu v_s(p') \varepsilon_\mu(k\lambda) e^{-i(p-p'+k)\cdot x} \quad (6.105b)$$

4) The normal ordering is a formal manipulation to avoid unwanted constants, which are often infinite. Whether this is physically allowed is a matter to be treated carefully. For instance, the vacuum energy that was inherent in quantized harmonic oscillators can be eliminated, but we have seen that its existence is experimentally established by the Casimir effect [92].

Figures 6.4e,f, which describe pair creation or annihilation by the photon, are expressed as

$$(e) \gamma \rightarrow e^- + e^+ \quad \langle p r, \bar{p}' s | \mathcal{H}_1 | k \lambda \rangle = -e \bar{u}_r(p) \gamma^\mu v_s(p') \varepsilon_\mu(k \lambda) e^{-i(k-p-p') \cdot x} \quad (6.106a)$$

$$(f) e^- + e^+ \rightarrow \gamma \quad \langle k \lambda | \mathcal{H}_1 | p r, \bar{p}' s \rangle = -e \bar{v}_s(p') \gamma^\mu u_r(p) \varepsilon_\mu(k \lambda)^* e^{-i(p+p'-k) \cdot x} \quad (6.106b)$$

Finally Figs. 6.4g,h, which describe total transition to vacuum and vice versa, are expressed as

$$(g) \gamma + e^- + e^+ \rightarrow \text{vacuum} \quad \langle 0 | \mathcal{H}_1 | p r, \bar{p}' s; k \lambda \rangle = -e \bar{v}_s(p') \gamma^\mu u_r(p) \varepsilon_\mu(k \lambda) e^{-i(p+p'+k) \cdot x} \quad (6.107a)$$

$$(h) \text{vacuum} \rightarrow e^- + e^+ + \gamma \quad \langle p r, \bar{p}' s; k \lambda | \mathcal{H}_1 | 0 \rangle = -e \bar{u}_r(p) \gamma^\mu v_s(p') \varepsilon_\mu(k \lambda)^* e^{i(p+p'+k) \cdot x} \quad (6.107b)$$

Virtual Particles

In quantum field theory, we describe most physical phenomena in terms of particles. However, in the case of interacting electromagnetic fields, for instance a static Coulomb field between two charged particles, no particles are emitted, yet we talk of the force as the result of particle exchange. What is the difference between a real photon, which appears as a quantum of freely propagating waves, and the quantum of the particles exchanged between two interacting objects? The freely propagating wave satisfies the Klein–Gordon equation in vacuum. Then Einstein’s relation

$$p^2 = E_p^2 - |\mathbf{p}|^2 = m^2 \quad \text{massive particle} \quad (6.108a)$$

$$k^2 = \omega^2 - |\mathbf{k}|^2 = 0 \quad \text{photon} \quad (6.108b)$$

should hold for the field quanta. A particle satisfying Eq. (6.108) is often said to be “on the mass shell” or simply “on shell”. The electric field that is produced by a point charged particle satisfies

$$\partial_\mu \partial^\mu A^\nu(x) = e \delta^3(r) \rightarrow k^2 \neq 0 \quad (6.109)$$

The quantum of the interacting field, therefore, does not satisfy Einstein’s relation. It is said to be “off shell” or a virtual particle. It is not a physical particle and cannot be observed as such. The reader is reminded that for the virtual particle, the ordinary physical picture does not apply. For instance, if the force is created by exchange of a real particle, it can only give a repulsive force; no attractive force can result because emission or absorption of a particle having momentum always causes the party involved to recoil. Another example: the real photon has only two degrees of freedom in its polarization. Since the polarization is a vector in space-time, it needs four components in a relativistic treatment. Artificially (mathematically) introduced extra quanta (longitudinal and scalar photons) appear only in the

intermediate state of a process. They show peculiar properties such as negative probability, superluminal velocity⁵⁾ and nonzero mass.

The vertices described above generally connect initial states to intermediate and final states. Initial and final states that are prepared by experimental set ups are composed of real (on mass shell) particles that satisfy the Einstein relation $E = \sqrt{p^2 + m^2}$, ω (photon) = $|k|$, or in relativistic notation $p^2 = m^2$, $k^2 = 0$. This means that once the momentum p of a particle is specified, its energy is fixed, too. At the vertex, both the energy and the momentum are conserved in a relativistic treatment, which in many cases does not allow all the particles to be “on shell”; one or two or all of the states are off mass shell. An exception is an initial massive particle (say a π) converting to two less-massive particles ($\mu + \nu$), where all the particles could be “on shell” if energy-momentum conservation is satisfied.

6.6.2

Feynman Propagator

We have so far studied only free fields, which satisfy a certain equation of motion, a homogeneous differential equation. An interacting field, which is created by a source, obeys an inhomogeneous differential equation. A formal solution to the interacting field is given using a Green’s function. In quantized field theories this is the role played by the Feynman propagator. It appears everywhere and is an indispensable tool for the calculation of scattering matrices. The definition and the properties of the Feynman propagator are given below.

Feynman Propagator of the Photon

The Feynman propagator of the photon is defined as

$$\begin{aligned} i D_{F\mu\nu}(x - y)^{(6)} &\equiv \langle 0 | T[A_\mu(x) A_\nu(y)] | 0 \rangle \\ &= \begin{cases} \langle 0 | A_\mu(x) A_\nu(y) | 0 \rangle & t_x > t_y \\ \langle 0 | A_\nu(y) A_\mu(x) | 0 \rangle & t_x < t_y \end{cases} \end{aligned} \quad (6.110)$$

$T(\dots)$ is the “T product”, defined in Eq. (6.32). Inserting a unit operator $1 = \sum |n\rangle \langle n|$ and considering that A_μ has matrix elements only between states differing in particle number by 1, only one-particle intermediate states differing in momentum or spin contribute. Inserting Eq. (6.75) and considering

$$\langle 0 | A_\mu(x) | q\lambda \rangle = \varepsilon_\mu(\lambda) e^{-iq \cdot x} \quad (6.111a)$$

$$\langle q\lambda | A_\nu(y) | 0 \rangle = \varepsilon_\nu(\lambda)^* e^{iq \cdot y} \quad (6.111b)$$

$$\sum_\lambda \varepsilon_\mu(\lambda) \varepsilon_\nu(\lambda)^* \rightarrow -g_{\mu\nu} \text{ see Eq. (5.147b)} \quad (6.111c)$$

- 5) The space-like momentum $k^2 < 0$ means its (formal) velocity is larger than the velocity of light.
- 6) It is customary to write D_F for particles with $m = 0$ and Δ_F when $m^2 \neq 0$. This applies to other functions, e.g. Δ , $\Delta_{\text{adv}}^{\text{ret}}$, too.

we can express the Feynman propagator as

$$\begin{aligned} & \langle 0 | T[A_\mu(x) A_\nu(y)] | 0 \rangle \\ &= -g_{\mu\nu} \int \frac{d^3 q}{(2\pi)^3 2\omega} \{ \theta(x-y) e^{-iq \cdot (x-y)} + \theta(y-x) e^{iq \cdot (x-y)} \} \end{aligned} \quad (6.112a)$$

$$\theta(x) = \begin{cases} 1 & t > 0 \\ 0 & t < 0 \end{cases} \quad (6.112b)$$

The step function can be expressed by a contour integral in the complex plane:

$$\theta(t) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} dz \frac{e^{izt}}{z - i\varepsilon} \quad (6.113)$$

where $i\varepsilon$ is a small imaginary part prescribed to handle the singularity. Then

$$\begin{aligned} & \langle 0 | T[A_\mu(x) A_\nu(y)] | 0 \rangle \\ &= -g_{\mu\nu} \int \frac{d^3 q}{(2\pi)^3 2\omega} \frac{1}{2\pi i} \\ & \quad \times \int dz \left[\frac{e^{-i(\omega-z)t' + iq \cdot x'}}{z - i\varepsilon} + \frac{e^{-i(z-\omega)t' - iq \cdot x'}}{z - i\varepsilon} \right]_{x'=x-y} \\ &= \frac{-ig_{\mu\nu}}{(2\pi)^4} \int \frac{d^3 q}{2\omega} \int dq^0 \left[\frac{1}{q^0 - \omega + i\varepsilon} - \frac{1}{q^0 + \omega - i\varepsilon} \right] e^{-iq^0 t' + iq \cdot x'} \end{aligned} \quad (6.114a)$$

$$= \frac{-ig_{\mu\nu}}{(2\pi)^4} \int d^4 q \frac{e^{-iq \cdot (x-y)}}{q^2 + i\varepsilon} \quad (6.114b)$$

In summary, the Feynman photon propagator is expressed as an integral

$$\langle 0 | T[A_\mu(x) A_\nu(y)] | 0 \rangle \equiv iD_{F\mu\nu}(x-y) = \frac{-ig_{\mu\nu}}{(2\pi)^4} \int d^4 q \frac{e^{-iq \cdot (x-y)}}{q^2 + i\varepsilon} \quad (6.115)$$

It is a vacuum expectation value of two field operators that create a photon at y and annihilate it at x , and is referred to as a two-point function of the field. In the language of classical field theory, it is the effect of the electromagnetic field created by a source (a charged particle) at a point y , acting on a charged particle at point x . That the time ordering is considered suggests the role of a Green function. We shall see that the Feynman propagator D_F is indeed a Green's function of the electromagnetic field that satisfies the equation

$$\partial_\mu \partial^\mu A^\nu(x) = q j^\nu(x) \quad (6.116)$$

Green's Function

Let us start from the simplest relativistic field $\phi(x)$ that satisfies a spin-less Klein-Gordon equation. The field produced by a source obeys the equation given by

$$[\partial_\mu \partial^\mu + m^2] \phi = \rho(x) \quad (6.117)$$

where ρ is a source to produce the field. A formal solution to the equation is expressed using a Green's function $G(x, y)$, which satisfies

$$\phi(x) = \phi_0(x) - \int d^4 y G(x, y) \rho(y) \quad (6.118a)$$

$$(\partial_\rho \partial^\rho + m^2) G(x, y) = -\delta^4(x - y) \quad (6.118b)$$

$$(\partial_\rho \partial^\rho + m^2) \phi_0(x) = 0 \quad (6.118c)$$

The physical meaning of the Green's function is that the field at space-time point x is given by the action of the field created by a source at space-time point y that then propagates to point x .

The boundary conditions imposed on G is most frequently translationally invariant, so G can be regarded as a function of $x - y$. Then it can be solved using Fourier transformation:

$$G(x, y) = \frac{1}{(2\pi)^4} \int d^4 q \tilde{G}(q) e^{-iq \cdot (x - y)} \quad (6.119a)$$

$$\delta^4(x - y) = \frac{1}{(2\pi)^4} \int d^4 q e^{-iq \cdot (x - y)} \quad (6.119b)$$

Inserting Eq. (6.119) into Eq. (6.118b) we obtain

$$\tilde{G}(q) = \frac{1}{q^2 - m^2}, \quad q^2 = (q^0)^2 - \mathbf{q}^2, \quad \omega = \sqrt{q^2 + m^2} \quad (6.120)$$

The Green's function is expressed as

$$\begin{aligned} G(x, y) &= \frac{1}{(2\pi)^4} \int d^4 q \frac{e^{-iq \cdot (x - y)}}{q^2 - m^2} \\ &= \frac{1}{(2\pi)^4} \int d^3 q e^{iq \cdot (x - y)} \int dq^0 \frac{e^{-iq^0(t_x - t_y)}}{2\omega} \left[\frac{1}{q^0 - \omega} - \frac{1}{q^0 + \omega} \right] \end{aligned} \quad (6.121)$$

The integration on q^0 can be interpreted as a contour integral in the complex plane, but a rule has to be established to handle the singularity. The retarded and advanced Green's function can be defined by requiring $G(x - y) = 0$ for $t_x < t_y$ and $t_x > t_y$. For $t_x > t_y$ ($t_x < t_y$), the contour can be closed in the lower (upper) semi-circle, therefore, adding a small imaginary part, $q^0 \rightarrow q^0 \pm i\epsilon$, gives the desired effect:

$$\begin{aligned} G_{\text{adv}}^{\text{ret}}(x - y) &= \frac{1}{(2\pi)^4} \int d^3 q e^{iq \cdot (x - y)} \\ &\quad \int dq^0 \frac{e^{-iq^0(t_x - t_y)}}{2\omega} \left[\frac{1}{q^0 \pm i\epsilon - \omega} - \frac{1}{q^0 \pm i\epsilon + \omega} \right] \\ &= \mp i\theta[\pm(x^0 - y^0)] \int \frac{d^3 q}{(2\pi)^3 2\omega} [e^{-iq \cdot (x - y)} - e^{iq \cdot (x - y)}] \\ &= \begin{cases} \theta(x^0 - y^0) \Delta(x - y) \\ -\theta(y^0 - x^0) \Delta(x - y) \end{cases} \end{aligned} \quad (6.122a)$$

$$i\Delta(x) \equiv \frac{1}{(2\pi)^3} \int \frac{d^3q}{2\omega} [e^{-iq \cdot x} - e^{iq \cdot x}] = \int \frac{d^4q}{(2\pi)^3} \epsilon(k^0) \delta(k^2 - m^2) e^{-ik \cdot x} \quad (6.122b)$$

where $\epsilon(k^0) = \pm 1$ for $k_0 \gtrless 0$. In relativistic quantum theory, we have learned that the negative energy contribution should be interpreted as a particle proceeding backward in time. Then the contribution of the negative energy should vanish for $t_x > t_y$.

To make $e^{-iq \cdot x}$ an ordinary propagating wave when $q^0 = \omega > 0$ and a backward-traveling wave when $q^0 = -\omega < 0$, we have to close the integral path in such a way as to include the pole at $q^0 = \omega$ when $t > 0$ and $q^0 = -\omega$ when $t < 0$. This can be achieved by adding or subtracting a small imaginary number $i\epsilon$ to the denominator. Using

$$\begin{aligned} \int_{-\infty}^{+\infty} dq^0 \frac{e^{-iq^0 t}}{(q^0)^2 - \omega^2 + i\epsilon} &= \int_{-\infty}^{+\infty} dq^0 \frac{e^{-iq^0 t}}{2\omega} \left(\frac{1}{q^0 - \omega + i\epsilon} - \frac{1}{q^0 + \omega - i\epsilon} \right) \\ &= \begin{cases} -2\pi i \frac{e^{-i\omega t}}{2\omega} & t > 0 \\ -2\pi i \frac{e^{+i\omega t}}{2\omega} & t < 0 \end{cases} \end{aligned} \quad (6.123a)$$

the Green's function for the quantized Klein-Gordon field can be expressed as

$$G_F(x - y) = -i \int \frac{d^3q}{(2\pi)^3 2\omega_q} [\theta(x^0 - y^0) e^{-iq \cdot (x-y)} + \theta(y^0 - x^0) e^{iq \cdot (x-y)}] \quad (6.123b)$$

which is nothing but the Feynman propagator Eq. (6.112) except for the polarization sum factor $-g_{\mu\nu}$ and i . The Feynman propagator for a spinless scalar field is defined as

$$\begin{aligned} \langle 0 | T[\varphi^\dagger(x) \varphi(y)] | 0 \rangle &= i\Delta_F(x - y) = \frac{i}{(2\pi)^4} \int d^4q \frac{e^{-iq \cdot (x-y)}}{q^2 - m^2 + i\epsilon} \\ (\partial_\mu \partial^\mu + m^2 - i\epsilon) \Delta_F(x - y) &= -\delta^4(x - y) \end{aligned} \quad (6.124)$$

Electron Propagator

The propagator for the Dirac particle can be defined similarly:

$$\begin{aligned} iS_F(x - y) &\equiv \langle 0 | T[\psi(x) \bar{\psi}(y)] | 0 \rangle \\ &= \begin{cases} \langle 0 | \psi(x) \bar{\psi}(y) | 0 \rangle & t_x > t_y \\ -\langle 0 | \bar{\psi}(y) \psi(x) | 0 \rangle & t_x < t_y \end{cases} \end{aligned} \quad (6.125)$$

The minus sign in the second line comes from the anticommutativity of the fermion field. Taking the sum over intermediate states, we obtain a similar expression as for the photon propagator except for the spin sum, which has to be done

separately for particle and antiparticle. Using the spin sum expression Eq. (4.69) in Sect. 4.2,

$$\begin{aligned}
 i S_F(x, y) &= \int \frac{d^3 p}{(2\pi)^3 2E} \left[\theta(x - y) \sum_r u_r(p) \bar{u}_r(p) e^{-i p \cdot (x - y)} \right. \\
 &\quad \left. - \theta(y - x) \sum_r v_r(p) \bar{v}_r(p) e^{i p \cdot (x - y)} \right] \\
 &= \int \frac{d^3 p}{(2\pi)^3 2E} [\theta(x - y) (\not{p} + m) e^{-i p \cdot (x - y)} \\
 &\quad + \theta(y - x) (-\not{p} + m) e^{i p \cdot (x - y)}] \\
 &= (i \not{\nabla} + m) \int \frac{d^3 p}{(2\pi)^3 2E} [\theta(x - y) e^{-i p \cdot (x - y)} \\
 &\quad + \theta(y - x) e^{i p \cdot (x - y)}] \\
 &= (i \not{\nabla} + m) \Delta_F(x - y) = \frac{i}{(2\pi)^4} \int d^4 p e^{-i p \cdot x} \frac{(\not{p} + m)}{p^2 - m^2 + i\epsilon}
 \end{aligned} \tag{6.126}$$

The spin sum $(\not{p} \pm m)$ was replaced with derivatives in the penultimate line. This is permissible because the two time derivatives of the step functions compensate each other and vanish. The electron propagator Eq. (6.126) has exactly the same form as the photon propagator Eq. (6.115) except for the numerator, where the photon polarization sum is replaced by the electron spin sum when $t_x > t_y$ and minus the positron spin sum when $t_x < t_y$:

$$\tilde{S}_F(p) = \frac{\not{p} + m}{p^2 - m^2 + i\epsilon} \equiv \frac{1}{\not{p} - m + i\epsilon} \tag{6.127}$$

The last equality is often used when one wants a simple expression, but its real meaning is defined by the preceding equation. Since it is easily shown that the electron propagator satisfies the equation

$$(\gamma^\mu i \partial_\mu - m) S_F(x, y) = \delta^4(x - y) \tag{6.128}$$

it, too, is the Green's function for the Dirac equation.

Relation with the Nonrelativistic Propagator

Apart from the spin sum in the numerator, the Feynman propagator consists of two energy denominators:

$$\tilde{D}_F(q) \sim \frac{1}{q^2 + i\epsilon} = \frac{1}{2\omega} \left[\frac{1}{q^0 - \omega + i\epsilon} + \frac{1}{-q^0 - \omega + i\epsilon} \right] \tag{6.129}$$

It is elucidating to consider how they are related to the nonrelativistic propagator

$$\frac{1}{E_i - E_n} \tag{6.130}$$

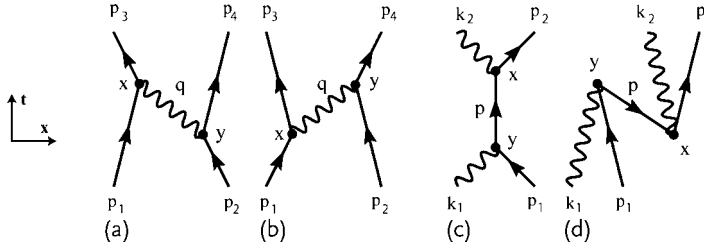


Figure 6.5 The two processes of $e-\mu$ scattering (a) $t_x > t_y$, (b) $t_x < t_y$, and Compton scattering (c) $t_x > t_y$, (d) $t_x < t_y$. The Feynman propagator contains both of them together.

Let us consider the example of $e-\mu$ scattering via hypothetical spinless photon exchange. In the nonrelativistic treatment, the time sequence of the process has to be considered. There are two intermediate states, depending on which particle emits a photon first (see Fig. 6.5a,b). As the energy of the two intermediate states after emission of the photon is given by

$$E_n(a) = E_1 + E_4 + \omega, \quad E_n(b) = E_2 + E_3 + \omega \quad (6.131a)$$

the virtual photon energy that the electron at point x receives from the muon at point y is given by

$$q^0 = \begin{cases} E_3 - E_1 = E_2 - E_4 > 0 & \text{case (a)} \\ E_3 - E_1 = E_2 - E_4 < 0 & \text{case (b)} \end{cases} \quad (6.131b)$$

The two denominators of the Feynman propagator can be rewritten as

$$q^0 - \omega = (E_2 - E_4) - \{E_n(a) - E_1 - E_4\} = E_1 + E_2 - E_n(a) = E_i - E_n(a) \quad (6.131c)$$

$$-(q^0 + \omega) = -(E_3 - E_1) - \{E_n(b) - E_2 - E_3\} = E_1 + E_2 - E_n(b) = E_i - E_n(b) \quad (6.131d)$$

which exactly reproduces the two energy denominators of the nonrelativistic treatment. The factor $1/2\omega$ arises from the relativistic normalization. Namely, the Feynman propagator contains both intermediate states by allowing the particle to take negative energy and interpreting it as propagating backward in time. Actually, it means the reaction of the electron back to the muon, and really there is no particle traveling backward in time.

Negative energy is meant to represent an antiparticle, too. In the case of the photon, there is no distinction between particle and antiparticle and the spin sum gives $(-g_{\mu\nu})$ in both cases. For the case of Compton scattering (see Fig. 6.5c,d), the Feynman propagator in (c) represents an electron created at y proceeding to x with positive energy, while that in (d) represents a particle proceeding with negative energy going backward in time. The latter is an antiparticle, i.e. positron, emitted at x that proceeds normally to y . The expression for the energy denominator is the

same as that of the photon propagator in $e\text{-}\mu$ scattering, but now the spin sum has to be done separately for the particle and antiparticle:

$$\begin{aligned}\tilde{S}_F(p) &\sim \frac{1}{2E} \left[\frac{\sum_{r=1,2} u(p)\bar{u}(p)}{p^0 - E + i\varepsilon} + \frac{-\sum_{r=1,2} v(-p)\bar{v}(-p)}{-p^0 - E + i\varepsilon} \right] \\ &\Rightarrow \frac{\not{p} + m}{p^2 - m^2 + i\varepsilon}\end{aligned}\quad (6.132)$$

where the minus sign in front of the $v(-p)\bar{v}(-p)$ comes from the Feynman prescription for the change of time order and the negative momentum comes from the change of the integration variable $\mathbf{p} \rightarrow -\mathbf{p}$. Recall that $u(p)\bar{u}(p)$ and $v(p)\bar{v}(p)$ appeared as the coefficient of $e^{-ip\cdot(x-y)}$ and $e^{ip\cdot(x-y)}$ originally.

Note: The intermediate photon in the above example has the energy-momentum $q^\mu = (q^0, \mathbf{q})$. Since $q = p_3 - p_1 = p_2 - p_4$, $q^2 \neq 0$, the photon is not on shell. On the other hand, intermediate states inserted to calculate the Feynman propagator are particles created by the creation operator of the free field and therefore are on shell in the sense that their energy ω is defined by the relation $\omega = |\mathbf{q}|$. Where did the difference come from? The fact is that q^0 that appears in the Feynman propagator is an integration variable obtained by setting $q^0 = \pm(z + \omega)$ in Eq. (6.114a). It is this q^0 that satisfies energy conservation and has nothing to do with ω , the energy of the photon. Some people say that in the nonrelativistic treatment particles in the intermediate states are on shell but the energy is not conserved in the transition, and in the relativistic treatment the energy and momenta are conserved but the particles are not on shell. The reader should realize that the latter half of this statement is not correct, as is clear from the above argument.

6.7

Mott Scattering

6.7.1

Cross Section

A process in which an electron is scattered by an unmovable scattering center, typically a heavy nucleus, is called Mott scattering. The target should be heavy enough for its recoil to be neglected. However, the electron should be energetic enough to be treated relativistically. The interaction Hamiltonian in this case is

$$\mathcal{H}_I = q j^\mu A_\mu, \quad j^\mu = \bar{\psi} \gamma^\mu \psi \quad (6.133)$$

As long as the recoil of the target can be neglected, there is no effect that the scattering electron can give to the target. Then the field created by the target is not disturbed and can be treated as an external source, namely as a classical c-number potential. Such external fields are, for example, a static Coulomb force [$A^\mu = (Zq'/4\pi r, 0, 0, 0)$] or an oscillating magnetic field [$A^\mu = (0, -\gamma B \sin \omega t,$

$x B \sin \omega t, 0]$. The scattering matrix element is given by

$$S_{fi} - \delta_{fi} = -i \int d^4x \langle f | \mathcal{H}_I | i \rangle = -i \int d^4x \langle p_f s | q j^\mu | p_i r \rangle A_\mu \quad (6.134a)$$

$$A_\mu(x) = (Z q' / 4\pi r, 0, 0, 0) \quad (6.134b)$$

where $p_i r$ and $p_f s$ denote the momentum and the spin state of the incoming and outgoing electrons, respectively. The matrix element of the Hamiltonian was calculated in Sect. 6.6.1:

$$\begin{aligned} \langle p_f s | q j^\mu(x) | p_i r \rangle &= q \langle p_f s | \bar{\psi}(x) \gamma^\mu \psi(x) | p_i r \rangle \\ &= q \bar{u}_s(p_f) \gamma^\mu u_r(p_i) e^{-i(p_i - p_f) \cdot x} \end{aligned} \quad (6.135)$$

Inserting Eq. (6.135) and Eq. (6.134b) into Eq. (6.134a), we obtain

$$\begin{aligned} S_{fi} - \delta_{fi} &= -2\pi i \delta(E_i - E_f) q \bar{u}_s(p_f) \gamma^0 u_r(p_i) \left[\frac{Z q'}{4\pi} \int d^3x \frac{e^{i(p_i - p_f) \cdot x}}{r} \right] \\ &= -2\pi i \delta(E_i - E_f) q \bar{u}_s(p_f) \gamma^0 u_r(p_i) [\tilde{V}(\mathbf{q})] \\ &= -2\pi i \delta(E_i - E_f) \frac{Z q q'}{|\mathbf{q}|^2} \bar{u}_s(p_f) \gamma^0 u_r(p_i) \\ \tilde{V}(\mathbf{q}) &= \int d^3x e^{i\mathbf{q} \cdot \mathbf{x}} V(\mathbf{x}) = \frac{Z q'}{|\mathbf{q}|^2}, \quad \mathbf{q} = \mathbf{p}_i - \mathbf{p}_f \end{aligned} \quad (6.136)$$

Then

$$d\sigma = \frac{1}{2E_i v} \left(\frac{Ze^2}{4\pi} \right)^2 \left(\frac{4\pi}{|\mathbf{q}|^2} \right)^2 |\bar{u}_s(p_f) \gamma^0 u_r(p_i)|^2 2\pi \delta(E_i - E_f) \frac{d^3p_f}{(2\pi)^3 2E_f} \quad (6.137)$$

This expression, like that for Rutherford scattering, does not contain the δ -function $\delta^3(\mathbf{p}_i - \mathbf{p}_f)$, i.e. the momentum is not conserved. This originates from the external field approximation that the target nucleus does not recoil.

In order to obtain the cross section, we need to calculate $|\bar{u}_s(p_f) \gamma^0 u_r(p_i)|^2$. Given a special example, where only the 0th component of the spinor contributes, it may be easier and faster to use the explicit expression for $u_r(p)$ given in Sect. 4.2, Eq. (4.63). But for later applications, we adopt a more general approach here. Under usual experimental conditions, the incoming particles are not polarized, and the final polarization is not measured either. In that case, we take the average (take the sum and divide by the degrees of freedom) of the initial-state polarization and the sum of the final-state polarization. Write the quantity $L^{\mu\nu}$ as

$$L^{\mu\nu} \equiv \frac{1}{2} \sum_{r,s=1}^2 [\bar{u}_s(p_f) \gamma^\mu u_r(p_i)] [\bar{u}_s(p_f) \gamma^\nu u_r(p_i)]^* \quad (6.138a)$$

$$[\bar{u}_s(p_f)\gamma^\nu u_r(p_i)]^* = \bar{u}_r(p_i)\bar{\Gamma}u_s(p_f) \quad (6.138b)$$

$$\bar{\Gamma} = \gamma^0 \Gamma^\dagger \gamma^0 \quad (6.138c)$$

When $\Gamma = \gamma^\mu$, $\bar{\Gamma} = \gamma^\mu$,

$$L^{\mu\nu} = \frac{1}{2} \sum_{r,s=1}^2 [\bar{u}_s(p_f)\gamma^\mu u_r(p_i)] [\bar{u}_r(p_i)\gamma^\nu u_s(p_f)] \quad (6.139)$$

Now we use the properties of the Dirac spinors

$$\sum_r u_r(p)\bar{u}_r(p) = \not{p} + m, \quad \sum_r v_r(p)\bar{v}_r(p) = \not{p} - m \quad (6.140)$$

and insert them in Eq. (6.139) to obtain

$$L^{\mu\nu} = \frac{1}{2} \text{Tr}[\gamma^\mu(\not{p}_i + m)\gamma^\nu(\not{p}_f + m)] \quad (6.141)$$

The above trace of the γ matrix can be calculated by making use of the properties of γ matrices (see Appendix C),

$$\text{Tr}[\gamma^{\mu_1}\gamma^{\mu_2}\dots\gamma^{\mu_{2n+1}}] = 0 \quad (6.142a)$$

$$\text{Tr}[\gamma^{\mu_1}\gamma^{\mu_2}\dots\gamma^{\mu_n}] = (-1)^n \text{Tr}[\gamma^{\mu_n}\dots\gamma^{\mu_2}\gamma^{\mu_1}] \quad (6.142b)$$

$$\text{Tr}[\gamma^\mu\gamma^\nu] = 4g^{\mu\nu} \quad (6.142c)$$

$$\text{Tr}[\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma] = 4(g^{\mu\nu}g^{\rho\sigma} + g^{\mu\sigma}g^{\nu\rho} - g^{\mu\rho}g^{\nu\sigma}) \quad (6.142d)$$

Then the final expression for $L^{\mu\nu}$ is

$$\begin{aligned} L^{\mu\nu} &= 2 \left[\left\{ p_f^\mu p_i^\nu + p_f^\nu p_i^\mu - (p_f \cdot p_i)g^{\mu\nu} \right\} + m^2 g^{\mu\nu} \right] \\ &= 2 \left[p_f^\mu p_i^\nu + p_f^\nu p_i^\mu + \frac{q^2}{2} g^{\mu\nu} \right] \end{aligned} \quad (6.143a)$$

$$\begin{aligned} q^2 &= (p_i - p_f)^2 = 2m^2 - 2(p_f \cdot p_i) \\ &\simeq -2p_f p_i (1 - \cos \theta) \quad (p_i, p_f \gg m) \end{aligned} \quad (6.143b)$$

Problem 6.6

Using the properties of γ matrices Eq. (4.36), derive Eq. (6.142).

Hint: To derive the first equation, use $(\gamma^5)^2 = 1$, $\gamma^5\gamma^\mu = -\gamma^\mu\gamma^5$.

Now, going back to the Mott scattering cross section, only $\mu = \nu = 0$ components contribute. Since the energy is conserved by virtue of the δ -function, $E_i = E_f = E$, $|\mathbf{p}_i| = |\mathbf{p}_f| = p$, $v = p/E$. Then

$$L^{00} = 2[2E_f E_i + m^2 - E_f E_i + p_f p_i \cos \theta] = 4E^2 \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) \quad (6.144)$$

The incoming flux and the final state density are given by

$$2E_i v = 2p_i = 2p \quad (6.145a)$$

$$\delta(E_i - E_f) \frac{d^3 p}{(2\pi)^3 2E} = \left(\frac{dE}{dp} \right)^{-1} \frac{p^2 d\Omega}{16\pi^3 E} = \frac{p d\Omega}{16\pi^3} \quad (6.145b)$$

Inserting Eqs. (6.144) and (6.145) into Eq. (6.137), we obtain

$$\frac{d\sigma}{d\Omega} = \frac{4Z^2 \alpha^2}{q^4} E^2 \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) \quad (6.146)$$

which is called the Mott scattering formula. Comparing this with the Rutherford scattering formula (6.46), we see that the relativistic effect makes $m \rightarrow E$ or ($v = p/m \rightarrow p/E$), and the spin effect of the electron gives

$$1 \rightarrow \frac{1}{2} \frac{\sum |\bar{u}_s(p_f) \gamma^0 u_r(p_i)|^2}{4E_1 E_2} = \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) \quad (6.147)$$

Problem 6.7

Using the explicit solution Eq. (4.63)

$$u_r(\mathbf{p}) = \begin{bmatrix} \sqrt{p \cdot \vec{\sigma}} \xi_r \\ \sqrt{p \cdot \vec{\sigma}} \xi_r \end{bmatrix} = \sqrt{\frac{E+m}{2}} \begin{bmatrix} \left(1 - \frac{\sigma \cdot \mathbf{p}}{E+m} \right) \xi_r \\ \left(1 + \frac{\sigma \cdot \mathbf{p}}{E+m} \right) \xi_r \end{bmatrix} \quad (6.148)$$

show that spin nonflip and spin flip cross sections are proportional to

$$d\sigma(\text{spin nonflip}) \propto \left[1 + \frac{p^2}{(E+m)^2} \cos \theta \right]^2 \quad (6.149a)$$

$$d\sigma(\text{spin flip}) \propto \left[\frac{p^2}{(E+m)^2} \sin \theta \right]^2 \quad (6.149b)$$

and that their sum is equal to Eq. (6.147) apart from the normalization factor.

Problem 6.8

Using the current of the charged scalar particle

$$q j^\mu = -ie(\varphi^\dagger \partial^\mu \varphi - \partial^\mu \varphi^\dagger \varphi) \quad (6.150)$$

derive the scattering formula for a given external Coulomb potential and show that it agrees with the first term of Eq. (6.146).

6.7.2

Coulomb Scattering and Magnetic Scattering

Scattering by the electromagnetic field can be divided into that by the electric and that by the magnetic field. The Hamiltonian is written as

$$H_I = q\phi - \boldsymbol{\mu} \cdot \mathbf{B}, \quad \boldsymbol{\mu} = -\frac{e}{2m} g \mathbf{s} \quad (6.151)$$

where $\mathbf{s}$ is the electron spin and $g = 2$ is the g -factor. Therefore, we see scattering by the electric field is spin nonflip and that by the magnetic field is a spin-flip processes. Mott scattering by a static Coulomb potential contains the spin-flip term. This is because, viewed from the electron's rest frame, the target is moving and hence a magnetic field is induced. As is clear from Eq. (6.146), the magnetic contribution is proportional to v^2 , and it is small for nonrelativistic particles but becomes important and comparable to the electric contribution for relativistic particles.

6.7.3

Helicity Conservation

In the high-energy limit, where the electron mass can be neglected, $v = 1$ and the Mott scattering cross section becomes

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{Mott}} = \frac{4Z^2\alpha^2}{q^4} E^2 \cos^2 \frac{\theta}{2} \quad (6.152)$$

which differs from that of spinless particles by $\cos^2(\theta/2)$. This is a helicity rotation effect. The reason is

1. the massless particle is a Weyl particle, and only polarizations along the direction of its momentum are allowed (i.e. helicity eigenstates) and
2. a vector-type interaction, like that of the electromagnetic force, conserves the chirality (see arguments in Sect. 4.3.5), i.e. helicity is conserved in the high-energy limit.

The initial helicity is along the z -axis (parallel to the incoming momentum), and the final helicity along the scattered particle at angle θ . The matrix element of the rotation operator is given by $d_{f,i}^j = \langle j, j_z' | \exp(-i\theta J_y) | j, j_z \rangle$ and for the Mott case $j = 1/2$, $j_z = \pm 1/2$, $d_{1/2,1/2}^{1/2} = \cos(\theta/2)$. (See Appendix E for the rotation matrix.)

6.7.4

A Method to Rotate Spin

The above argument inspires a method to rotate spin using low-energy Coulomb scattering. Place heavy nuclei around a circle and have a low-energy charged particle with longitudinal polarization start from one end of a diameter. Let it be scattered by a nucleus and detected at the other end of the diameter (see Fig. 6.6). As

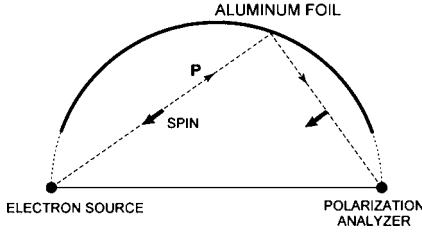


Figure 6.6 Converting longitudinal polarization to transverse polarization using Mott scattering.

the scattering angle is 90° and the spin does not flip, the particle's polarization is rotated by 90° , and now the particle has transverse polarization, which is suitable for measurement. The method was used to decide the helicity of the decay electron from β decay.

6.8

Yukawa Interaction

Yukawa interaction was first proposed to explain the nuclear force and it was claimed it was due to the exchange of a spin 0 massive meson. The interaction of a spin $1/2$ particle with a scalar field is important both historically and for future applications. Historically, it offered the first viable explanation of low-energy nuclear interactions, i.e. strong interactions. For future applications, it is important for interaction with the Higgs field, which generates the mass of all otherwise massless quarks and leptons. Here, we investigate a low-energy nuclear scattering mediated by a scalar or a pseudoscalar meson. The relevant Lorentz-invariant interaction is of the type

$$\mathcal{L}_{\text{INT}} = -g\bar{\psi}\Gamma\psi\phi \quad (6.153)$$

where $\Gamma = 1$ if ϕ is a scalar and $\Gamma = i\gamma^5$ if ϕ is a pseudoscalar. We consider a process

$$N_1(p_1) + N_2(p_2) \rightarrow N_1(p_3) + N_2(p_4) \quad (6.154)$$

and assume N_1 and N_2 are distinguishable. The mass of the nucleon and the meson are taken to be m and μ , respectively. The second-order scattering amplitude corresponding to the exchange of the ϕ meson is (Fig. 6.7)

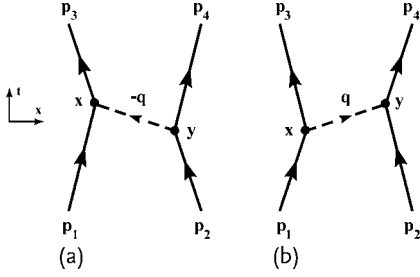


Figure 6.7 Nucleon–nucleon interaction via a scalar meson exchange.

$$\begin{aligned}
 S_{fi} - \delta_{fi} &= -iT_{fi} \\
 &= \langle p_3 p_4 | T \left[\frac{(-i)^2}{2!} g \int d^4x \bar{\psi}(x) \Gamma \psi(x) \phi(x) \right. \\
 &\quad \times g \int d^4y \bar{\psi}(y) \Gamma \psi(y) \phi(y) \left. \right] | p_1 p_2 \rangle \\
 &= \frac{(-ig)^2}{2!} \left[\langle p_3 | \int d^4x \bar{\psi}(x) \Gamma \psi(x) | p_1 \rangle \langle 0 | T[\phi(x) \phi(y)] | 0 \rangle \right. \\
 &\quad \times \langle p_4 | \int d^4y \bar{\psi}(y) \Gamma \psi(y) | p_2 \rangle + (x \leftrightarrow y) \left. \right]
 \end{aligned} \tag{6.155}$$

Interactions at x and y can be separated because particles 1 and 2 are two distinguishable particles and the interactions take place at separate space-time points. The second term, where x and y are exchanged, gives an identical contribution to the first, it merely compensates the factor $2!$. Inserting

$$\langle p_3 | \int d^4x \bar{\psi}(x) \Gamma \psi(x) | p_1 \rangle = \int d^4x e^{-i(p_1 - p_3) \cdot x} \bar{u}(p_3) \Gamma u(p_1) \tag{6.156a}$$

$$\langle 0 | T[\phi(x) \phi(y)] | 0 \rangle = \int \frac{d^4q e^{-iq \cdot (x - y)}}{(2\pi)^4} \frac{i}{q^2 - \mu^2} \tag{6.156b}$$

$$\langle p_4 | \int d^4y \bar{\psi}(y) \Gamma \psi(y) | p_2 \rangle = \int d^4y e^{-i(p_2 - p_4) \cdot y} \bar{u}(p_4) \Gamma u(p_2) \tag{6.156c}$$

into Eq. (6.155), we obtain

$$\begin{aligned}
 T_{fi} &= (2\pi)^4 \delta^4(p_1 - p_3 + q) (2\pi)^4 \delta^4(p_2 - p_4 - q) \int \frac{d^4q}{(2\pi)^4} \\
 &\quad \times \bar{u}(p_3) \Gamma u(p_1) \frac{g^2}{q^2 - \mu^2} \bar{u}(p_4) \Gamma u(p_2) \\
 &= (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \\
 &\quad \times \bar{u}(p_3) \Gamma u(p_1) \frac{g^2}{q^2 - \mu^2} \bar{u}(p_4) \Gamma u(p_2) \Big|_{q=p_3-p_1}
 \end{aligned} \tag{6.157}$$

Scalar Meson Exchange

$$\Gamma = 1$$

To evaluate the amplitude in the nonrelativistic limit, we retain terms only to the lowest order in the 3-momenta:

$$p_1 = (m, \mathbf{p}_1), \quad p_2 = (m, \mathbf{p}_2), \quad p_3 = (m, \mathbf{p}_3), \quad p_4 = (m, \mathbf{p}_4) \quad (6.158)$$

Then

$$\bar{u}_s(p_3)u_r(p_1) = \xi_s^\dagger(\sqrt{p_3 \cdot \sigma} \sqrt{p_1 \cdot \sigma}) \left[\frac{\sqrt{p_1 \cdot \sigma}}{\sqrt{p_1 \cdot \sigma}} \right] \xi_r \simeq 2m\delta_{rs} \quad (6.159a)$$

$$q^2 - \mu^2 = -(\mathbf{q}^2 + \mu^2) = -[(\mathbf{p}_3 - \mathbf{p}_1)^2 + \mu^2] \quad (6.159b)$$

The amplitude becomes

$$T_{fi} = -(2\pi)^4 \delta(p_1 + p_2 - p_3 - p_4) \frac{g^2 2m\delta_{rs} 2m\delta_{tu}}{|\mathbf{q}|^2 + \mu^2} \quad (6.160)$$

The amplitude of a scattering by a potential V in quantum mechanics is given by

$$T_{fi} = 2\pi\delta(E_i - E_f) \tilde{V}(q) \quad (6.161)$$

$$\tilde{V}(q) = \int d^3x \psi^\dagger(x) V(x) \psi(x) = \frac{1}{(2\pi)^3} \int d^3x e^{iq \cdot x} V(x)$$

Therefore $\tilde{V}(q)$ for the scalar meson exchange is given by

$$\tilde{V}(q) = -\frac{g^2}{q^2 + \mu^2} \quad (6.162)$$

We have dropped $(2m)^2$ and the δ -function for the momentum conservation. The factor $(2m)^2$ comes from the relativistic wave function normalization and should be dropped for comparison with the nonrelativistic formula. The δ -function originates from integration over the total coordinate $\mathbf{R}$, where

$$\begin{aligned} & \int d^3x e^{i(\mathbf{p}_1 - \mathbf{p}_3) \cdot \mathbf{x}} \int d^3y e^{i(\mathbf{p}_2 - \mathbf{p}_4) \cdot \mathbf{y}} \\ &= \int d^3R e^{i(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_3 - \mathbf{p}_4) \cdot \mathbf{R}} \int d^3r e^{i\mathbf{q} \cdot \mathbf{r}} \\ &= (2\pi)^3 \delta^3(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_3 - \mathbf{p}_4) \int d^3r e^{i\mathbf{q} \cdot \mathbf{r}} \\ & \mathbf{R} = \frac{1}{2}(\mathbf{x} + \mathbf{y}), \quad \mathbf{r} = \mathbf{x} - \mathbf{y} \end{aligned} \quad (6.163)$$

and means the CM frame is unaffected by the internal motion. In the nonrelativistic treatment, only the relative motion is treated, so it does not enter the equation. The potential in the space coordinate is given by

$$V(x) = \frac{1}{(2\pi)^3} \int d^3q e^{iq \cdot x} \frac{-g^2}{q^2 + \mu^2} = -\frac{g^2}{4\pi} \frac{e^{-\mu r}}{r} \quad (6.164)$$

which gives an attractive force. Yukawa postulated the existence of a massive force carrier of the nuclear force and knowing the force range to be approximately 10^{-15} m, he predicted its mass to be $\sim 200 m_e$.

For the pseudoscalar meson exchange, the nuclear force becomes repulsive. We know that the π meson is actually a pseudoscalar meson. How then can the nuclear force be attractive so that it can form a bound state, the deuteron? This is related to the charge exchange force, which will be discussed in connection with isospin exchange in Sect. 13.6.

7

Qed: Quantum Electrodynamics

Quantum electrodynamics (QED) is a mathematical framework for treating the quantum dynamics of electrons with the quantized electromagnetic field. However, it can be applied to the electromagnetic interactions of any charged Dirac particle, including quarks and leptons. It is a gauge theory based on the symmetry $U(1)$ and shares many basic features with more complicated non-Abelian gauge theories that govern the weak and the strong interactions. Thus learning QED is an indispensable step in understanding the dynamical aspect of particle physics.

7.1

 e - μ Scattering

As the next logical step after derivation of the Rutherford and Mott scattering, we treat the scattering process of a spin 1/2 particle by a target also having spin 1/2. When the recoil of the target is not negligible, as is the case for e - μ scattering, we have to use the quantized electromagnetic field. Experimentally, a viable choice would be e - p (electron-proton) scattering, but the proton is a hadron, which has a finite size as well as a large anomalous magnetic moment, and therefore is not a Dirac particle. So we consider e - μ scattering, which is impractical from the experimental point of view, but has applications to e - q (electron-quark) scattering and is useful for analyzing deep inelastic scattering data to probe the inner structure of the hadron. Besides, using crossing symmetry, the process can also be applied to $e^-e^+ \rightarrow \mu^-\mu^+$ and $e^-e^+ \rightarrow q\bar{q}$, (i.e. $e^-e^+ \rightarrow$ hadrons), which are widely used collider experiments.

7.1.1

Cross Section

We define the kinematic variables of e - μ scattering as follows:

$$e(p_1) + \mu(p_2) \rightarrow e(p_3) + \mu(p_4) \quad (7.1)$$

In p_i ($i = 1$ to 4), we include spin orientation as well, unless otherwise specified.

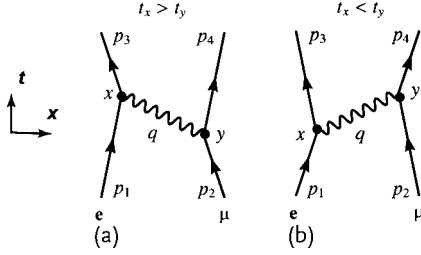


Figure 7.1 e - μ scattering process.

In the lowest order, e - μ scattering occurs as shown in Fig. 7.1, by exchange of a photon. As there are two vertices in the diagram, it is a second order process. The Hamiltonian contains electron as well as muon interactions:

$$\mathcal{H}_I = q(j_e^\nu + j_\mu^\nu)A_\nu = -e(\bar{\psi}_e \gamma^\nu \psi_e + \bar{\psi}_\mu \gamma^\nu \psi_\mu) A_\nu \quad (7.2)$$

The scattering matrix is given by

$$S_{fi} - \delta_{fi} = \frac{(-i)^2}{2!} \int d^4x \int d^4y T[\langle p_3, p_4 | \mathcal{H}_I(x) \mathcal{H}_I(y) | p_1, p_2 \rangle] \quad (7.3)$$

Considering the e -transition $p_1 \rightarrow p_3$ and μ -transition $p_2 \rightarrow p_4$ to occur in different space-time positions, they can be separated. We insert intermediate states n between the two Hamiltonians,

$$1 = \sum_n |n\rangle \langle n| \quad (7.4)$$

where n represents any physical state that can be realized through the interaction. The intermediate state in Fig. 7.1 corresponds to one photon with energy-momentum q^μ and polarization λ ,

$$\sum_n |n\rangle \langle n| = \sum_\lambda \int \frac{d^3q}{(2\pi)^3 2\omega} |q, \lambda\rangle \langle q, \lambda|, \quad \omega = |q| \quad (7.5)$$

Inserting the intermediate states between the two Hamiltonians, $T\langle \dots \rangle$ in Eq. (7.3) is expressed as

$$\sum_\lambda \int \frac{d^3q}{(2\pi)^3 2\omega} T[\langle p_3 | \mathcal{H}_I(x) | p_1, q\lambda \rangle \langle p_4, q\lambda | \mathcal{H}_I(y) | p_2 \rangle] \quad (7.6)$$

There is also a term with $p_1 \leftrightarrow p_2$, $p_3 \leftrightarrow p_4$, but exchanging the integration variable $x \leftrightarrow y$ makes it identical to Eq. (7.6) and merely compensates the factor $2!$. Therefore it is enough to consider Eq. (7.6) only. Referring to Eqs. (6.104a) and (6.104b), Eq. (7.6) can be expressed as

$$\begin{aligned} \langle p_3 | \mathcal{H}_I(x) | p_1, q\lambda \rangle &= -e \bar{u}_s(p_3) \gamma^\mu u_r(p_1) e^{-i(p_1 - p_3) \cdot x} \langle 0 | A_\mu(x) | q\lambda \rangle \\ \langle p_4, q\lambda | \mathcal{H}_I(y) | p_2 \rangle &= -e \bar{u}_m(p_4) \gamma^\nu u_l(p_2) e^{-i(p_2 - p_4) \cdot y} \langle q\lambda | A_\nu(y) | 0 \rangle \end{aligned} \quad (7.7)$$

Then

$$\begin{aligned}
 S_{fi} - \delta_{fi} &= (-i)^2 \int d^4x \int d^4\gamma \left\{ -e \bar{u}_s(p_3) \gamma^\mu u_r(p_1) e^{-i(p_1-p_3) \cdot x} \right\} \\
 &\times \sum_\lambda \int \frac{d^3q}{(2\pi)^3 2\omega} T[\langle 0 | A_\mu(x) | q\lambda \rangle \langle q\lambda | A_\nu(\gamma) | 0 \rangle] \\
 &\times \left\{ -e \bar{u}_m(p_4) \gamma^\nu u_l(p_2) e^{-i(p_2-p_4) \cdot \gamma} \right\}
 \end{aligned} \quad (7.8)$$

As the middle term is the Feynman propagator, we can insert Eq. (6.115) into it. Then

$$\begin{aligned}
 S_{fi} - \delta_{fi} &= (-i)^2 \int d^4x \int d^4\gamma \left\{ -e \bar{u}_s(p_3) \gamma^\mu u_r(p_1) e^{-i(p_1-p_3) \cdot x} \right\} \\
 &\times \frac{1}{(2\pi)^4} \int d^4q e^{-iq \cdot (x-\gamma)} \frac{-ig_{\mu\nu}}{q^2 + i\varepsilon} \\
 &\times \left\{ -e \bar{u}_m(p_4) \gamma^\nu u_l(p_2) e^{-i(p_2-p_4) \cdot \gamma} \right\}
 \end{aligned} \quad (7.9)$$

As integration over x, γ gives

$$\int d^4x \rightarrow (2\pi)^4 \delta^4(p_1 - p_3 + q) \quad (7.10a)$$

$$\begin{aligned}
 \int d^4\gamma &\rightarrow (2\pi)^4 \delta^4(p_2 - p_4 - q) \\
 \frac{1}{(2\pi)^4} \int d^4q \left[\iint d^4x d^4\gamma \right] &\rightarrow (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4)
 \end{aligned} \quad (7.10b)$$

the matrix element of the scattering matrix becomes

$$\begin{aligned}
 S_{fi} &= \delta_{fi} - (2\pi)^4 i \delta^4(p_1 + p_2 - p_3 - p_4) \mathcal{M}_{fi} \\
 &= \delta_{fi} + (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \\
 &\times \bar{u}_s(p_3) (ie\gamma^\mu) u_r(p_1) \left(\frac{-ig_{\mu\nu}}{q^2 + i\varepsilon} \right) \bar{u}_m(p_4) (ie\gamma^\nu) u_l(p_2) \\
 q &= p_3 - p_1 = p_2 - p_4
 \end{aligned} \quad (7.11)$$

which gives the Lorentz invariant matrix element $\mathcal{M}_{fi}$

$$\begin{aligned}
 \mathcal{M}_{fi} &= -j_\mu(e) \frac{g^{\mu\nu}}{q^2 + i\varepsilon} j_\nu(\mu) \\
 &= -e^2 \bar{u}_s(p_3) \gamma^\mu u_r(p_1) \frac{1}{q^2 + i\varepsilon} \bar{u}_m(p_4) \gamma_\mu u_l(p_2)
 \end{aligned} \quad (7.12)$$

Under normal experimental situations, the polarization states of the particle are not measured. Then to obtain the transition amplitude, we have to take the spin sum (average of the initial- and sum of the final-state polarization). We have already done part of the calculation [see Eqs. (6.139)–(6.143)],

$$\begin{aligned}
 L^{\mu\nu} &= \frac{1}{2} \sum_{r,s=1}^2 [\bar{u}_s(p_3) \gamma^\mu u_r(p_1)] [\bar{u}_r(p_1) \gamma^\nu u_s(p_3)] \\
 &= 2 \left[p_3^\mu p_1^\nu + p_3^\nu p_1^\mu + \frac{q^2}{2} g^{\mu\nu} \right]
 \end{aligned} \quad (7.13a)$$

The same calculation for the muon part gives

$$M_{\mu\nu} = 2 \left[p_{4\mu} p_{2\nu} + p_{4\nu} p_{2\mu} + \frac{q^2}{2} g_{\mu\nu} \right] \quad (7.13b)$$

The current conservation law $\partial_\mu j^\mu = 0$ gives

$$q_\mu L^{\mu\nu} = q_\nu L^{\mu\nu} = 0, \quad q_\mu M^{\mu\nu} = q_\nu M^{\mu\nu} = 0 \quad (7.14)$$

which can be explicitly verified from expressions (7.13).

Problem 7.1

Prove Eq. (7.14) by actually multiplying Eq. (7.13) by q_μ .

Using the relation

$$p_4 = p_4 - p_2 + p_2 = -q + p_2 \quad (7.15)$$

and Eq. (7.14), one sees that p_4 in $M_{\mu\nu}$ can be replaced by p_2 :

$$\begin{aligned} N &= L^{\mu\nu} M_{\mu\nu} \\ &= 4 \left[p_3^\mu p_1^\nu + p_3^\nu p_1^\mu + \frac{q^2}{2} g^{\mu\nu} \right] \left[2p_{2\mu} p_{2\nu} + \frac{q^2}{2} g_{\mu\nu} \right] \\ &= 4 \left[4(p_1 \cdot p_2)(p_2 \cdot p_3) + q^2 p_2^2 + q^2(p_1 \cdot p_3) + \left(\frac{q^2}{2}\right)^2 \times 4 \right] \end{aligned} \quad (7.16a)$$

Making use of

$$\begin{aligned} 2(p_1 \cdot p_2) &= (p_1 + p_2)^2 - p_1^2 - p_2^2 = s - m^2 - M^2 \\ 2(p_2 \cdot p_3) &= -(p_3 - p_2)^2 + p_3^2 + p_2^2 \\ &= -(p_1 - p_4)^2 + m^2 + M^2 = m^2 + M^2 - u \\ 2(p_1 \cdot p_3) &= -(p_1 - p_3)^2 + p_1^2 + p_3^2 = 2m^2 - t \end{aligned}$$

N is rewritten as

$$\begin{aligned} N &= 4 \left[(s - m^2 - M^2)(m^2 + M^2 - u) + t(m^2 + M^2) + \frac{t^2}{2} \right] \\ &= 2 \left[(s - m^2 - M^2)^2 + (u - m^2 - M^2)^2 + 2t(m^2 + M^2) \right] \end{aligned} \quad (7.16b)$$

where we used

$$s + t + u = 2(m^2 + M^2), \quad q^2 = t \quad (7.16c)$$

in deriving the last equality. Substituting Eqs. (7.16b) in the square of the matrix element Eq. (7.12), the invariant scattering amplitude (squared) becomes

$$\begin{aligned} |\overline{\mathcal{M}}_{fi}|^2 &= \frac{e^4}{q^4} N \\ &= \frac{4e^4}{q^4} \left[(s - m^2 - M^2)(m^2 + M^2 - u) + t(m^2 + M^2) + \frac{t^2}{2} \right] \end{aligned} \quad (7.17)$$

where $|\overline{\mathcal{M}}_{fi}|^2$ denotes that the average of the initial spin and sum of the final spin state were taken for $|\mathcal{M}_{fi}|^2$. In the LAB frame, neglecting the electron mass ($m = 0$), the invariant Mandelstam variables are expressed as

$$s - M^2 = 2p_1 M, \quad M^2 - u = 2p_3 M, \quad t = -2p_1 p_3 (1 - \cos \theta) \quad (7.18)$$

Inserting the above expression into Eq. (7.16b), the invariant matrix element is

$$|\overline{\mathcal{M}}_{fi}|_{\text{LAB}}^2 = \frac{e^4}{q^4} 16M^2 p_1 p_3 \cos^2 \frac{\theta}{2} \left[1 - \frac{q^2}{2M^2} \tan^2 \frac{\theta}{2} \right] \quad (7.19)$$

As the cross section in LAB is given by Eq. (6.91),

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{LAB}} = \left(\frac{p_3}{p_1} \right)^2 \frac{1}{64\pi^2 M^2} |\overline{\mathcal{M}}_{fi}|_{\text{LAB}}^2 \quad (7.20a)$$

$$= d\sigma_{\text{Mott}} \frac{p_3}{p_1} \left[1 - \frac{q^2}{2M^2} \tan^2 \frac{\theta}{2} \right] \quad (7.20b)$$

$$d\sigma_{\text{Mott}} = \frac{4\alpha^2 E_3^2}{q^4} \cos^2 \frac{\theta}{2} \quad (7.20c)$$

The first term in square brackets $[\dots]$ represents scattering by the electric field created by the target and the second by the magnetic field. As $p_3 = p_1/[1 + p_1(1 - \cos \theta)/M]$, the limit $M \rightarrow \infty$ reproduces the Mott formula.

The cross section in CM is given by Eq. (6.90)

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} = \frac{p_f}{p_i} \frac{1}{64\pi^2 s} |\overline{\mathcal{M}}_{fi}|_{\text{CM}}^2 \quad (7.21)$$

In many cases of high-energy scattering, where $s, u \gg M^2$, we can set $M^2 = 0$. Then

$$\begin{aligned} \left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} &= \frac{p_f}{p_i} \frac{1}{64\pi^2 s} \frac{(4\pi\alpha)^2}{q^4} \times 2(s^2 + u^2) \\ &= \frac{\alpha^2}{q^4} \frac{s^2 + u^2}{2s} = \frac{\alpha^2}{q^4} s \frac{1 + \left(\frac{1 + \cos \theta}{2} \right)^2}{2} \end{aligned} \quad (7.22)$$

This is a formula that appears frequently later in discussions of deep inelastic scattering.

7.1.2

Elastic Scattering of Polarized $e\text{--}\mu$

We consider in this section longitudinally polarized electron–muon scattering. We use here chiral eigenstates to describe the processes. The advantage is that the chirality does not change during the interaction and the calculation is easier than

that using the helicity eigenstate. Besides, it is also the eigenstate of the weak interaction and the derived formula can be applied directly to neutrino scattering. The disadvantage is that the chiral states are not eigenstates of the Hamiltonian, and the scattering amplitude written in terms of the chiral eigenstates does not describe the physical process accurately and hence does not necessarily reflect the properties of helicity eigenstates correctly. However, as was explained in Sect. 4.1.2, the massless particle is a Weyl particle and chirality L(R) state and the helicity states are the same except for a sign change between particles and antiparticles. In most high-energy experiments, the energy is high enough to justify setting the mass of the particle to zero. This is certainly true for electrons ($m \simeq 0.5$ MeV), muons and light quarks (u, d, s) ($m \simeq 5$ to 100 MeV) for typical experiments where the energy is above ~ 1 GeV.

The conservation of chirality in vector-type interactions, which include all the strong, weak and electromagnetic interactions in the Standard Model, was already shown in Sect. 4.3.5. This means

$$\begin{aligned}\bar{\psi}\gamma^\mu\psi &= \bar{\psi}_L\gamma^\mu\psi_L + \bar{\psi}_R\gamma^\mu\psi_R \\ \bar{\psi}_L\gamma^\mu\psi_R &= \bar{\psi}_R\gamma^\mu\psi_L = 0\end{aligned}\quad (7.23)$$

If we make use of the property that the chiral eigenstate does not change after interactions, the unpolarized $e\text{-}\mu$ scattering can be decomposed into the following four processes:

$$e_L\mu_L \rightarrow e_L\mu_L \quad (\text{LL}) \quad (7.24a)$$

$$e_R\mu_L \rightarrow e_R\mu_L \quad (\text{RL}) \quad (7.24b)$$

$$e_L\mu_R \rightarrow e_L\mu_R \quad (\text{LR}) \quad (7.24c)$$

$$e_R\mu_R \rightarrow e_R\mu_R \quad (\text{RR}) \quad (7.24d)$$

Let us consider $e_L\mu_L \rightarrow e_L\mu_L$ first. The scattering amplitude can be obtained from that of unpolarized $e\text{-}\mu$ scattering by the substitution $\gamma^\mu \rightarrow \gamma^\mu(1 - \gamma^5)/2$, hence

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} = \left(\frac{p_f}{p_i} \right) \frac{1}{64\pi^2 s} |\mathcal{M}_{fi}|^2 \quad (7.25a)$$

$$|\mathcal{M}_{fi}|^2 = \frac{e^4}{q^4} K_L^{\mu\nu}(p_1, p_3) K_{L\mu\nu}(p_2, p_4) \quad (7.25b)$$

$$K_L^{\mu\nu}(p_1, p_3) = \frac{1}{4} \sum_{r,s} \bar{u}_s(p_3) \gamma^\mu (1 - \gamma^5) u_r(p_1) \bar{u}_r(p_1) \gamma^\nu (1 - \gamma^5) u_s(p_3) \quad (7.25c)$$

$$= \frac{1}{4} \text{Tr} [\gamma^\mu (1 - \gamma^5) (\not{p}_1 + m_1) \gamma^\nu (1 - \gamma^5) (\not{p}_3 + m_3)] \quad (7.25d)$$

To calculate the trace, we use the properties of the γ^5 matrix:

$$\gamma^\mu \gamma^5 = -\gamma^5 \gamma^\mu, \quad \gamma^5 = -\frac{i}{4!} \varepsilon_{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \quad (7.26a)$$

$$\varepsilon^{0123} = -\varepsilon_{0123} = 1 \quad (7.26b)$$

$$\varepsilon_{\mu\nu\rho\sigma} = \begin{cases} -1 & \mu\nu\rho\sigma = \text{even permutation of } 0123 \\ +1 & \mu\nu\rho\sigma = \text{odd permutation of } 0123 \\ 0 & \text{otherwise} \end{cases} \quad (7.26c)$$

$$\varepsilon^{\mu\nu\rho\sigma} \varepsilon_{\mu\nu}{}^{\gamma\delta} = 2(g^{\rho\delta} g^{\sigma\gamma} - g^{\rho\gamma} g^{\sigma\delta}) \quad (7.26d)$$

Then

$$\text{Tr}[\gamma^5] = \text{Tr}[\gamma^5 \gamma^\mu] = \text{Tr}[\gamma^5 \gamma^\mu \gamma^\nu] = \text{Tr}[\gamma^5 \gamma^\mu \gamma^\nu \gamma^\rho] = 0 \quad (7.27a)$$

$$\text{Tr}[\gamma^5 \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma] = -4i \varepsilon^{\mu\nu\rho\sigma} \quad (7.27b)$$

Using these relations, Eq. (7.25d) can be simplified to give

$$K_L^{\mu\nu}(p_1, p_3) \equiv P^{\mu\nu} + Q^{\mu\nu} \quad (7.28a)$$

$$= 2[\{p_1^\mu p_3^\nu + p_1^\nu p_3^\mu - g^{\mu\nu}(p_1 \cdot p_3)\} + i \varepsilon^{\mu\nu\rho\sigma} p_{1\rho} p_{3\sigma}] \quad (7.28b)$$

$K_L^{\mu\nu}(p_1, p_3) K_{L\mu\nu}(p_2, p_4)$ can be calculated using the following equalities:

$$P^{\mu\nu}(p_1, p_3) P_{\mu\nu}(p_2, p_4) = 8[(p_1 \cdot p_2)(p_3 \cdot p_4) + (p_1 \cdot p_4)(p_2 \cdot p_3)] \quad (7.29a)$$

$$Q^{\mu\nu}(p_1, p_3) Q_{\mu\nu}(p_2, p_4) = 8[(p_1 \cdot p_2)(p_3 \cdot p_4) - (p_1 \cdot p_4)(p_2 \cdot p_3)] \quad (7.29b)$$

$$P^{\mu\nu} Q_{\mu\nu} = P_{\mu\nu} Q^{\mu\nu} = 0 \quad (7.29c)$$

$$\begin{aligned} 4(p_1 \cdot p_2)(p_3 \cdot p_4) &= [(p_1 + p_2)^2 - p_1^2 - p_2^2][(p_3 + p_4)^2 - p_3^2 - p_4^2] \\ &= (s - m_1^2 - m_2^2)(s - m_3^2 - m_4^2) \end{aligned} \quad (7.29d)$$

Similarly

$$4(p_1 \cdot p_4)(p_2 \cdot p_3) = (u - m_1^2 - m_4^2)(u - m_2^2 - m_3^2) \quad (7.29e)$$

Then

$$\begin{aligned} K_L^{\mu\nu}(p_1, p_3) K_{L\mu\nu}(p_2, p_4) &= 16(p_1 \cdot p_2)(p_3 \cdot p_4) \\ &= 4(s - m_1^2 - m_2^2)(s - m_3^2 - m_4^2) \stackrel{m_i \rightarrow 0}{=} 4s^2 \end{aligned} \quad (7.30)$$

Inserting Eq. (7.30) into Eq. (7.25), we obtain

$$d\sigma_{\text{LL}} \equiv \left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} (\text{LL}) = \frac{\alpha^2}{t^2} s \quad (7.31)$$

In calculating the cross section for $e_R\mu_L$ scattering, we only need to replace γ^5 by $-\gamma^5$ in the electron part of $K_L(p_1, p_3)$. Then, $Q^{\mu\nu}$ in $K_L(p_1, p_3)$ becomes $-Q^{\mu\nu}$, but that in the muon part does not change. Then $Q^{\mu\nu}Q_{\mu\nu}$ is changed to $-Q^{\mu\nu}Q_{\mu\nu}$ and the net effect is

$$(p_1 \cdot p_2)(p_3 \cdot p_4) \rightarrow (p_1 \cdot p_4)(p_2 \cdot p_3) \quad \text{or} \quad s \rightarrow u \quad (7.32a)$$

$$\therefore d\sigma_{RL} = \frac{\alpha^2}{t^2} \frac{u^2}{s} = \frac{\alpha^2}{t^2} s \left(\frac{1 + \cos \theta}{2} \right)^2 \quad (7.32b)$$

Others can be calculated similarly:

$$d\sigma_{LR} = d\sigma_{RL}, \quad d\sigma_{RR} = d\sigma_{LL} \quad (7.33)$$

The unpolarized scattering cross section is simply the sum of all the above cross sections divided by 4 to account for average L, R fractions in the beam:

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} = 2 \times \left[\frac{1}{4} \frac{\alpha^2}{t^2} \frac{s^2 + u^2}{s} \right] = \frac{\alpha^2}{t^2} \frac{s}{2} \left[1 + \left(\frac{1 + \cos \theta}{2} \right)^2 \right] \quad (7.34)$$

As $t = q^2$, this agrees with Eq. (7.22).

The above expression can be understood in terms of helicity conservation. In the CM frame, the $e_L\mu_L$ and $e_R\mu_L$ two-body systems are in the helicity eigenstates of the total angular momentum $J = 0, 1$. The system scattered by angle θ can be obtained from the incoming system by rotation if there are no other changes. The rotation matrix elements (see Appendix E) are given by

$$d_{j_z', j_z}^j = \langle j, j_z' | e^{-iJ_y \theta} | j, j_z \rangle \quad (7.35a)$$

$$d_{0,0}^0 = 1, \quad d_{1,1}^1 = \frac{1}{2}(1 + \cos \theta) \quad (7.35b)$$

The helicity states are described in Fig. 7.2. For LL or RR scattering, the sum of the J_z is zero and is conserved at any angle, but for RL or LR scattering $J_z = \pm 1$ and it is inverted at the scattering angle $\theta = \pi$, violating angular momentum conservation. The rotation matrix reflects this conservation law and gives the suppression factor $(1 + \cos \theta)^2/2$ in Eq. (7.35).

Note, that $e_L \leftrightarrow e_R$ exchange gives $s \leftrightarrow u$ in the scattering matrix. As $s \leftrightarrow u$ exchange can be alternatively realized by crossing ($p_1 \rightarrow -p_3$, $p_3 \rightarrow -p_1$), particle $\leftrightarrow$ antiparticle exchange of one or other partner in the scattering and that of $R \leftrightarrow L$ are equivalent as far as the kinematics of the scattering is concerned. In $e\text{-}\mu$ scattering the crossing can be distinguished by seeing the charge change of the particle, but for neutral particles, such as neutrinos, looking at the angular distribution difference is one way to tell particles from antiparticles.

7.1.3

$e^-e^+ \rightarrow \mu^-\mu^+$ Reaction

It is possible and not so difficult to calculate the $e\bar{e} \rightarrow \mu\bar{\mu}$ production cross section using the standard procedure we have learned so far [see Eq. (C.17)]. Here,

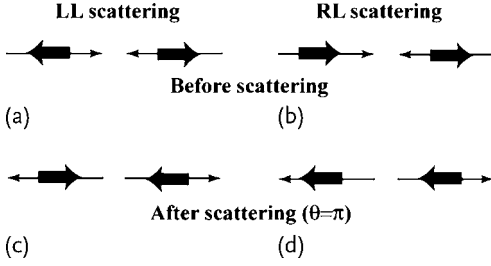


Figure 7.2 The difference of the angular distributions by helicity conservation. For LL scattering, sum of the angular components in the $e\text{-}\mu$ two-body system is zero and it is conserved at any angle. For RL scatter-

ing $J_z = 1/2 + 1/2$ before scattering, but after scattering at the scattering angle $\theta = \pi$, $J_z = -1$ and the angular momentum is not conserved. The suppression factor $(1 + \cos \theta)^2/4$ reflects this fact.

instead, we use crossing to derive the cross section. The $e\mu$ scattering amplitude as shown in Fig. 7.3a can be converted to $e\bar{e} \rightarrow \mu\bar{\mu}$ reaction (Fig. 7.3b) by crossing ($p_3 \rightarrow -p_3$, $p_2 \rightarrow -p_2$).

We attach a subscript t to the crossed channel variables to distinguish those before crossing,

$$s = (p_1 + p_2)^2 \rightarrow (p_1 - p_2)^2 = u_t \quad (7.36a)$$

$$t = (p_1 - p_3)^2 = q^2 \rightarrow (p_1 + p_3)^2 = s_t \quad (7.36b)$$

$$u = (p_1 - p_4)^2 \rightarrow (p_1 - p_4)^2 = t_t \quad (7.36c)$$

The scattering angle θ_t in the crossed channel is defined by the angle between e and μ . The spin-averaged scattering matrix element (squared) is then converted to

$$|\overline{\mathcal{M}}_{fi}|^2(e\mu \rightarrow e\mu) \sim \frac{s^2 + u^2}{t^2} \rightarrow |\overline{\mathcal{M}}_{fi}|^2(e\bar{e} \rightarrow \mu\bar{\mu}) \sim \frac{t_t^2 + u_t^2}{s_t^2} \quad (7.37)$$

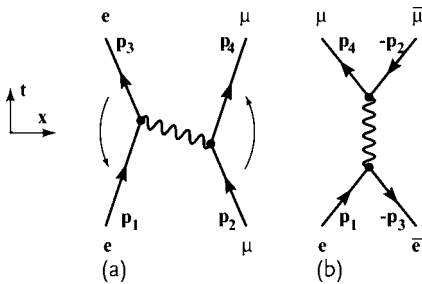


Figure 7.3 Crossing ($p_1 \rightarrow -p_1$, $p_3 \rightarrow -p_3$) converts $e\mu$ scattering (a) to $e\bar{e} \rightarrow \mu\bar{\mu}$ reaction (b).

Dropping the subscript, the cross section for $e\bar{e} \rightarrow \mu\bar{\mu}$ becomes

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} = \frac{\alpha^2}{s} \frac{t^2 + u^2}{2s^2} = \frac{4\pi\alpha^2}{3s} \left[\frac{3}{16\pi} (1 + \cos^2 \theta) \right] \quad (7.38a)$$

$$\sigma_{\text{TOT}} = \int d\sigma = \frac{4\pi\alpha^2}{3s} = \frac{87}{s[\text{GeV}^2]} \text{ nb} \quad (\text{nb} = 10^{-33} \text{ cm}^2) \quad (7.38b)$$

This is a basic cross section used in electron–positron colliding accelerator experiments.

Problem 7.2

Show if the mass of the produced particle is not neglected

$$\begin{aligned} \sigma_{\text{TOT}} &= \frac{4\pi\alpha^2}{3s} F(\beta) \\ F(\beta) &= \beta \left(\frac{3 - \beta^2}{2} \right), \quad \beta = \text{particle velocity in CM} \end{aligned} \quad (7.39)$$

7.2

Compton Scattering

Scattering Amplitude

The photon–electron scattering is known as the Compton scattering. Historically, it is famous as the process to have shown the particle nature of the light wave and have proved Einstein’s light quantum hypothesis proposed in 1905. Here, we calculate the process with later application to QCD in mind. The variables in the reaction are defined as

$$\gamma(k) + e(p) \rightarrow \gamma(k') + e(p') \quad (7.40)$$

In the lowest order, two processes shown in Fig. 7.4 contribute.

The scattering matrix is expressed as

$$S_{fi} - \delta_{fi} = \frac{(-iq)^2}{2!} \int d^4x d^4y \langle f | T[j^\mu(x) A_\mu(x) j^\nu(y) A_\nu(y)] | i \rangle \quad (7.41)$$

From now on we will omit δ_{fi} from the scattering matrix where there is no confusion. Separating the photon and electron parts

$$\begin{aligned} S_{fi} &= (-iq)^2 \int d^4x d^4y T[\langle k' | A_\mu(x) | 0 \rangle \langle p' | j^\mu(x) j^\nu(y) | p \rangle \langle 0 | A_\nu(y) | k \rangle] \\ &= (-iq)^2 \int d^4x d^4y \varepsilon_\mu(k')^* \langle p' | T[j^\mu(x) j^\nu(y)] | p \rangle \varepsilon_\nu(k) e^{i(k' \cdot x - k \cdot y)} \end{aligned} \quad (7.42a)$$

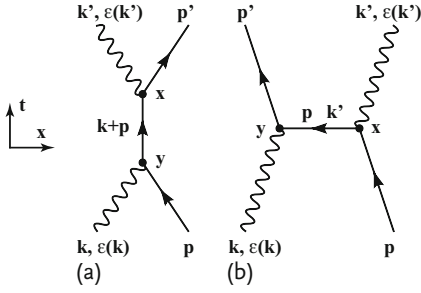


Figure 7.4 Two diagrams that contribute to Compton scattering.

$$\langle p' | T[j^\mu(x) j^\nu(y)] | p \rangle = \langle p' | T[\bar{\psi}(x) \gamma^\mu \psi(x) \bar{\psi}(y) \gamma^\nu \psi(y)] | p \rangle \quad (7.42b)$$

where the polarization index λ has been suppressed for simplicity. The factor $1/2$ is canceled by an identical contribution obtained by $x \leftrightarrow y$ in Eq. (7.41). Corresponding to two annihilation operators ψ , there are two ways of making the matrix element. When $x^0 - y^0 > 0$, writing the matrix elements suffixes explicitly

$$\begin{array}{cc} (1) & (2) \\ \underbrace{\langle p' | \bar{\psi}_\alpha(x) \gamma^\mu_{\alpha\beta} \psi_\beta(x) \bar{\psi}_\gamma(y) \gamma^\nu_{\gamma\delta} \psi_\delta(y) | p \rangle}_{(2)} & \underbrace{\langle p' | \bar{\psi}_\alpha(x) \gamma^\mu_{\alpha\beta} \psi_\beta(x) \bar{\psi}_\gamma(y) \gamma^\nu_{\gamma\delta} \psi_\delta(y) | p \rangle}_{(1)} \end{array} \quad (7.43)$$

In the bracket labeled (1), the incoming particle $|p\rangle$ is annihilated by ψ_δ at point y and the outgoing particle $\langle p'|$ is created by $\bar{\psi}_\alpha$ at point x . In bracket (2), $|p\rangle$ is annihilated by ψ_β at point x and the outgoing particle $\langle p'|$ is created by $\bar{\psi}_\gamma$ at point y . Therefore, bracket (1) corresponds to diagram (a) of Fig. 7.4 and bracket (2) to (b). The matrix element of bracket (1) can be written as

$$\begin{aligned} \langle p' | T[j^\mu(x) j^\nu(y)] | p \rangle &|_{(a) \ x^0 > y^0} \\ &= \bar{u}_\alpha(p') \gamma^\mu_{\alpha\beta} \langle 0 | \psi_\beta(x) \bar{\psi}_\gamma(y) | 0 \rangle \gamma^\nu_{\gamma\delta} u_\delta(p) e^{i(p' \cdot x - p \cdot y)} \end{aligned} \quad (7.44)$$

When $x^0 - y^0 < 0$

$$\langle p' | T[j^\mu(x) j^\nu(y)] | p \rangle = \langle p' | j^\nu(y) j^\mu(x) | p \rangle \quad (7.45)$$

This is the same as Eq. (7.43) if the arguments are interchanged:

$$x \leftrightarrow y, \quad \mu \leftrightarrow \nu, \quad \alpha\beta \leftrightarrow \gamma\delta \quad (7.46)$$

$$\begin{array}{cc} (1) & (2) \\ \underbrace{\langle p' | \bar{\psi}_\gamma(y) \gamma^\nu_{\gamma\delta} \psi_\delta(y) \bar{\psi}_\alpha(x) \gamma^\mu_{\alpha\beta} \psi_\beta(x) | p \rangle}_{(2)} & \underbrace{\langle p' | \bar{\psi}_\gamma(y) \gamma^\nu_{\gamma\delta} \psi_\delta(y) \bar{\psi}_\alpha(x) \gamma^\mu_{\alpha\beta} \psi_\beta(x) | p \rangle}_{(1)} \end{array} \quad (7.47)$$

This time diagram (a) corresponds to bracket (2), and its matrix element is given by

$$\begin{aligned} \langle p' | T[j^\mu(x) j^\nu(y)] | p \rangle_{(a)} &= \bar{u}_\alpha(p') \gamma_{\alpha\beta}^\mu \langle 0 | -\bar{\psi}_\gamma(y) \psi_\beta(x) | 0 \rangle \gamma_{\gamma\delta}^\nu u_\delta(p) e^{i(p' \cdot x - p \cdot y)} \end{aligned} \quad (7.48)$$

Combining Eqs. (7.44) and (7.48)

$$\langle p' | T[j^\mu(x) j^\nu(y)] | p \rangle_{(a)} = \bar{u}(p') \gamma^\mu \langle 0 | T[\psi(x) \bar{\psi}(y)] | 0 \rangle \gamma^\nu u(p) e^{i(p' \cdot x - p \cdot y)} \quad (7.49)$$

$\langle | \dots | \rangle$ is just the Feynman propagator. Using Eq. (6.126), the scattering matrix element corresponding to diagram (a) of Fig. 7.4 becomes

$$\begin{aligned} S_{fi}|_{(a)} &= \int d^4x d^4y e^{i(k' + p') \cdot x - i(k + p) \cdot y} \varepsilon_\mu(k')^* \bar{u}(p') (-i q \gamma^\mu) \\ &\times \left[\frac{i}{(2\pi)^4} \int d^4q e^{-iq(x-y)} \frac{\not{q} + m}{q^2 - m^2 + i\varepsilon} \right] (-i q \gamma^\nu) u(p) \varepsilon_\nu(k) \end{aligned} \quad (7.50a)$$

$$= (2\pi)^4 \delta^4(k + p - k' - p') \quad (7.50b)$$

$$\times \varepsilon_\mu(k')^* \bar{u}(p') (-i q \gamma^\mu) \frac{i(\not{p} + \not{k} + m)}{(p + k)^2 - m^2 + i\varepsilon} (-i q \gamma^\nu) u(p) \varepsilon_\nu(k) \quad (7.50c)$$

The scattering matrix for diagram (b) can be obtained similarly, but is more easily obtained by crossing diagram (a), i.e. by interchanging the variables $k \leftrightarrow -k'$, $\varepsilon(k')^* \leftrightarrow \varepsilon(k)$. Then the total scattering amplitude is given by

$$\mathcal{M} = \mathcal{M}_a + \mathcal{M}_b \quad (7.51a)$$

$$= e^2 \bar{u}(p') \left[\not{\varepsilon}'^* \frac{\not{p} + \not{k} + m}{(p + k)^2 - m^2} \not{\varepsilon} + \not{\varepsilon} \frac{\not{p} - \not{k}' + m}{(p - k')^2 - m^2} \not{\varepsilon}'^* \right] u(p) \quad (7.51b)$$

where $\varepsilon = \gamma^\mu \varepsilon_\mu(k)$, $\varepsilon' = \gamma^\mu \varepsilon_\mu(k')$ are used for brevity. The calculations that follow are familiar trace algebra. But what will become important for later applications is the high-energy behavior at $s \gg m^2$. So we will set $m = 0$, which considerably simplifies the calculation.

Summation of the Photon Polarization Assuming that neither photon nor electron are polarized, we take the average of the initial and sum of the final states. As discussed in Sect. 5.3.4, the polarization sum can be extended to all four states, giving

$$\sum_{\lambda=\pm} \varepsilon_\mu(\lambda)^* \varepsilon_\nu(\lambda) = \delta_{ij} - \frac{k^i k^j}{|k|^2} \rightarrow \sum_{\lambda=\text{all}} \varepsilon_\mu(\lambda)^* \varepsilon_\nu(\lambda) = -g_{\mu\nu} \quad (7.52)$$

This is because the polarization vector always appears coupled with the conserved current. Since the conserved current satisfies

$$\partial_\mu j^\mu = 0 \rightarrow k_\mu j^\mu = k^0 j^0 - |\mathbf{k}| j^3 = 0 \quad (7.53)$$

in the coordinate system where the z-axis is taken along the momentum direction. Then the polarization sum of the matrix elements (squared) is rewritten as

$$\begin{aligned} \sum_\lambda |\mathcal{M}|^2 &\sim \sum_\lambda |\varepsilon_\mu(\mathbf{k}, \lambda) j^\mu \dots|^2 \sim -g_{\mu\nu} j^\mu j^{\nu\dagger} \times (\dots) \\ -g_{\mu\nu} j^\mu j^{\nu\dagger} &= \underbrace{-j^0 j^{0\dagger} + j^3 j^{3\dagger}}_{j^0 j^{0\dagger} \left(-1 + \frac{(k^0)^2}{|\mathbf{k}|^2}\right)} + \underbrace{j^1 j^{1\dagger} + j^2 j^{2\dagger}}_{\mathbf{j}_\perp \cdot \mathbf{j}_\perp} \\ &= \frac{k^2}{k^2} |j^0|^2 + \mathbf{j}_\perp \cdot \mathbf{j}_\perp^\dagger = \mathbf{j}_\perp \cdot \mathbf{j}_\perp^\dagger = \left(\delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2}\right) j^i j^{j\dagger} \end{aligned} \quad (7.54)$$

The first term vanishes because for the real photon $k^2 = 0$.

As the degrees of freedom of both the photon polarization and electron spin are two, we divide by 4 in taking the spin-polarization average:

$$\begin{aligned} |\overline{\mathcal{M}}|^2 &= \frac{1}{4} \sum |\mathcal{M}_a + \mathcal{M}_b|^2 = \frac{1}{4} \sum (|\mathcal{M}_a|^2 + |\mathcal{M}_b|^2 + 2 \operatorname{Re} \mathcal{M}_a \mathcal{M}_b^*) \\ &= 2e^2 \left(-\frac{u}{s} - \frac{s}{u}\right) \end{aligned} \quad (7.55a)$$

We prove the first term. Setting $m = 0$ and picking up only the numerators, we have

$$\sum_{r,s,\lambda} |\mathcal{M}_a|^2 \sim e^4 \sum \bar{u}_s(p') \not{\epsilon}'^* (\not{p} + \not{k}) \not{\epsilon} u_r(p) \bar{u}_r(p) \not{\epsilon}^* (\not{p} + \not{k}) \not{\epsilon}' u_s(p') \quad (7.56a)$$

Using (see Appendix C for trace calculation formula)

$$\sum_\lambda \not{\epsilon}^* \dots \not{\epsilon} = \sum_\lambda \varepsilon_\mu^* \gamma^\mu \dots \varepsilon_\nu \gamma^\nu = (-g_{\mu\nu}) \gamma^\mu \dots \gamma^\nu = -\gamma_\mu \dots \gamma^\mu \quad (7.56b)$$

$$\gamma_\mu \not{Q} \gamma^\mu = -2 \not{Q} \quad (7.56c)$$

$$\sum_\lambda \not{\epsilon}^* \left(\sum_r u_r(p) \bar{u}_r(p) \right) \not{\epsilon} = -\gamma_\mu (\not{p} + m) \gamma^\mu \simeq 2 \not{p} \quad (7.56d)$$

$$\sum_\lambda \not{\epsilon}^* \left(\sum_s u_r(p') \bar{u}_s(p') \right) \not{\epsilon}' \simeq 2 \not{p}' \quad (7.56e)$$

$$\begin{aligned} \therefore \frac{1}{4} \sum_{r,s,\lambda} |\mathcal{M}_a|^2 &= e^4 \operatorname{Tr}[(\not{p} + \not{k}) \not{p} (\not{p} + \not{k}) \not{p}'] \\ &= e^4 \operatorname{Tr}[\not{p} \not{p} \not{p}' + \not{p} \not{p} \not{k} \not{p}' + \not{k} \not{p} \not{p}' + \not{k} \not{p} \not{k} \not{p}'] \\ &= e^4 \operatorname{Tr}[\not{k} \not{p} \not{k} \not{p}'] = 4e^4 [2(k \cdot p)(k \cdot p') - k^2(p \cdot p')] \\ &= -2e^4 su \end{aligned} \quad (7.56f)$$

Here we have used the relations

$$\not{p}' \not{p}' = p'^2 = m^2 = 0, \quad \not{k} \not{k} = k^2 = 0, \quad s = 2(k \cdot p), \quad u = -2(k \cdot p') \quad (7.56g)$$

Recovering other factors, and using that the denominators give $(p + k)^2 - m^2 \simeq s^2$

$$|\overline{\mathcal{M}}_a|^2 = -2e^4 \frac{u}{s} \quad (7.56h)$$

The interference term vanishes. Then the cross section is obtained from Eq. (6.92)

$$\frac{d\sigma}{dt} = \frac{|\overline{\mathcal{M}}_{fi}|^2}{16\pi(s - m^2)^2} \simeq \frac{2\pi\alpha^2}{s^2} \left(-\frac{u}{s} - \frac{s}{u} \right) \quad (7.57)$$

Problem 7.3

When Eq. (7.51) is written as $\varepsilon'^*_{\mu} \mathcal{M}^{\mu\nu} \varepsilon_{\nu}$, gauge symmetry requires (Sect. 5.3.4) invariance under the transformations

$$\varepsilon'^*_{\mu} \rightarrow \varepsilon'^*_{\mu} + c k'_{\mu}, \quad \varepsilon_{\nu} \rightarrow \varepsilon_{\nu} + c k_{\nu} \quad (7.58)$$

Therefore the following equalities should hold:

$$k'_{\mu} \mathcal{M}^{\mu\nu} = k_{\nu} \mathcal{M}^{\mu\nu} = 0 \quad (7.59)$$

Prove the equality explicitly using Eq. (7.51)

Problem 7.4

Show that gauge invariance does not hold with either $\mathcal{M}_a$ or $\mathcal{M}_b$ alone, but holds with both terms added together.

Problem 7.5

Show that if we assume that the incoming photon is virtual and has mass $k^2 = -Q^2$, then the interference term does not vanish, and Eq. (7.55a) becomes

$$|\overline{\mathcal{M}}|^2 = 2e^4 \left(-\frac{u}{s} - \frac{s}{u} + \frac{2Q^2 t}{su} \right) \quad (7.60)$$

This equation appears frequently in QCD.

Problem 7.6*

If we assume the electron mass is nonzero, the calculation can be carried out in the LAB frame. Writing the variables as

$$k = (\omega, \mathbf{k}), \quad k' = (\omega', \mathbf{k}'), \quad p = (m, \mathbf{0}), \quad \varepsilon = (0, \boldsymbol{\epsilon})$$

then the following equalities are satisfied:

$$\varepsilon \cdot p = \varepsilon' \cdot p = \varepsilon \cdot k = \varepsilon' \cdot k' = 0 \quad (7.61a)$$

$$\varepsilon \cdot p' = -\varepsilon \cdot k', \quad \varepsilon' \cdot p' = \varepsilon' \cdot k \quad (7.61b)$$

Taking the average of the electron spin but leaving the photon polarization as it is, derive after a straightforward but long calculation

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{LAB}} = \frac{\alpha^2}{4m^2} \left(\frac{\omega'}{\omega} \right)^2 \left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\varepsilon \cdot \varepsilon')^2 \right] \quad (7.62)$$

This is known as the Klein–Nishina formula.

7.3

Bremsstrahlung

Scattering Amplitude*

Bremsstrahlung (braking radiation in German) is a process in which a charged particle emits a photon when its path is bent by the electric field of the nucleus. It can be considered as Compton scattering by a virtual photon (γ^*) produced by the nucleus:

$$\gamma^*(q) + e(p_i) \rightarrow \gamma(k) + e(p_f) \quad (7.63)$$

The scattering amplitude is obtained from that of the Compton scattering Eq. (7.51) by replacing the incoming photon (q) with an external field $A^\mu = (Ze/r, 0, 0, 0)$.

$$\varepsilon_\mu(q) e^{-iq \cdot \gamma} \rightarrow \frac{Ze}{4\pi r} \delta_{\mu 0} = \frac{Ze}{(2\pi)^3} \int d^4 q' e^{iq' \cdot \gamma} \frac{\delta_{\mu 0} \delta(q^0)}{|q|^2} \quad (7.64)$$

$$q = p_f - p_i + k$$

$$\begin{aligned} S_{fi} = & +2\pi i \delta(E_i - E_f - \omega) \frac{Ze}{|q|^2} \bar{u}(p_f) \\ & \times \left[(ie\not{\epsilon}^*) \frac{i}{\not{p}_f + \not{k} - m} (ie\gamma^0) + (ie\gamma^0) \frac{i}{\not{p}_i - \not{k} - m} (ie\not{\epsilon}^*) \right] u(p_i) \\ = & 2\pi i \delta(E_i - E_f - \omega) \bar{u}(p_f) e\Gamma u(p_i) A^0 \end{aligned} \quad (7.65)$$

$$\Gamma = \not{\epsilon}^* \frac{(\not{p}_f + \not{k} + m)}{(p_f + k)^2 - m^2} \gamma^0 + \gamma^0 \frac{(\not{p}_i - \not{k} + m)}{(p_i - k)^2 - m^2} \not{\epsilon}^*,$$

$$A^0 = -\frac{Ze^2}{|q|^2}$$

The unpolarized cross section becomes

$$d\sigma(p_i \rightarrow p_f + k) = 2\pi\delta(E_i - E_f - \omega) \frac{Z^2 e^6}{2E_i |\mathbf{v}_i|} \overline{\sum} |\bar{u}_f \Gamma u_i|^2 \times \frac{1}{|\mathbf{q}|^4} \frac{d^3 p_f}{2E_f (2\pi)^3} \frac{d^3 k}{2\omega (2\pi)^3} \quad (7.66)$$

where $\overline{\sum}$ takes the average of the initial and the sum of the final electron spin states and the sum of the photon polarization states. The trace we want to calculate is

$$\begin{aligned} \overline{\sum} |\bar{u}_f \Gamma u_i|^2 &= \frac{1}{2} \sum_{\varepsilon} \text{Tr} \left[\left(\not{\varepsilon}^* \frac{\not{p}_f + \not{k} + m}{2p_f \cdot k} \gamma^0 - \gamma^0 \frac{\not{p}_i - \not{k} + m}{2p_i \cdot k} \not{\varepsilon}^* \right) \right. \\ &\quad \times (\not{p}_i + m) \\ &\quad \times \left. \left(\gamma^0 \frac{\not{p}_f + \not{k} + m}{2p_f \cdot k} \not{\varepsilon} - \not{\varepsilon} \frac{\not{p}_i - \not{k} + m}{2p_i \cdot k} \gamma^0 \right) (\not{p}_f + m) \right] \\ &\equiv \frac{1}{8} (T_1 + T_2 + T_3) \end{aligned} \quad (7.67)$$

We have rewritten the denominator using the relation

$$(p_f + k)^2 - m^2 = 2(k \cdot p_f), \quad (p_i + k)^2 - m^2 = 2(k \cdot p_i) \quad (7.68)$$

Choosing $\epsilon^0 = 0$, the calculation of the traces yields

$$\begin{aligned} T_1 &= (p_f \cdot k)^{-2} \sum_{\varepsilon} \text{Tr} \\ &\quad \times \{ \not{\varepsilon} (\not{p}_f + \not{k} + m) \gamma^0 (\not{p}_i + m) \gamma^0 (\not{p}_f + \not{k} + m) \not{\varepsilon} (\not{p}_f + m) \} \\ &= 8(k \cdot p_f)^{-2} \\ &\quad \times \sum_{\varepsilon} [2(\varepsilon \cdot p_f)^2 \{m^2 + 2E_i E_f + 2E_i \omega - (p_i \cdot p_f) - (k \cdot p_i)\} \\ &\quad + 2(\varepsilon \cdot p_f)(\varepsilon \cdot p_i)(k \cdot p_f) + 2E_i \omega(k \cdot p_f) - (k \cdot p_i)(k \cdot p_f)] \end{aligned} \quad (7.69a)$$

$$T_2 = T_1(p_i \leftrightarrow -p_f) \quad (7.69b)$$

$$\begin{aligned} T_3 &= -(p_f \cdot k)^{-1} (p_i \cdot k)^{-1} \sum_{\varepsilon} \text{Tr} \left[\gamma^0 (\not{p}_i - \not{k} + m) \not{\varepsilon} (\not{p}_i + m) \gamma^0 \right. \\ &\quad \times (\not{p}_f + \not{k} + m) \not{\varepsilon} (\not{p}_f + m) + (p_i \leftrightarrow -p_f) \left. \right] \\ &= 16(k \cdot p_f)^{-1} (k \cdot p_i)^{-1} \sum_{\varepsilon} \left[(\varepsilon \cdot p_i)(\varepsilon \cdot p_f) \{ (k \cdot p_i) - (k \cdot p_f) \} \right. \\ &\quad + 2(p_i \cdot p_f) - 4E_i E_f - 2m^2 \} \\ &\quad + (\varepsilon \cdot p_f)^2 (k \cdot p_i) - (\varepsilon \cdot p_i)^2 (k \cdot p_f) + (k \cdot p_i)(k \cdot p_f) - m^2 \omega^2 \\ &\quad \left. + \omega \{ \omega(p_i \cdot p_f) - E_i(p_f \cdot k) - E_f(k \cdot p_i) \} \right] \end{aligned} \quad (7.69c)$$

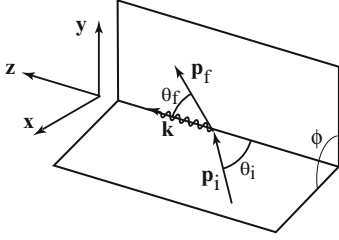


Figure 7.5 Bremsstrahlung. The emitted photon is in the z direction and the emitted electron is in the y - z plane. The incoming electron is in a plane rotated by angle ϕ from the y - z plane.

The reaction does not take place in a plane, so we choose a coordinate frame as described in Fig. 7.5.

We take the z -axis along the photon momentum $\mathbf{k}$, the emitted electron momentum $\mathbf{p}_f$ in the y - z plane and the incoming electron momentum $\mathbf{p}_i$ in a plane rotated by an angle ϕ from the y - z plane. In this coordinate system, the parameters are

$$\begin{aligned}
 k^\mu &= (\omega, 0, 0, k) \quad \omega = |\mathbf{k}| = k \\
 p_f^\mu &= (E_f, 0, p_f \sin \theta_f, p_f \cos \theta_f) \\
 p_i^\mu &= (E_i, p_i \sin \theta_i \sin \phi, p_i \sin \theta_i \cos \phi, p_i \cos \theta_i) \\
 \varepsilon^1 &= (0, 1, 0, 0), \quad \varepsilon^2 = (0, 0, 1, 0)
 \end{aligned} \tag{7.70}$$

where $p_i = |\mathbf{p}_i|$, $p_f = |\mathbf{p}_f|$. Then the cross section is expressed as

$$\begin{aligned}
 d\sigma &= \frac{Z^2 \alpha^3}{4\pi^2} \frac{p_f}{p_i |\mathbf{q}|^4} \frac{d\omega}{\omega} d\Omega_k d\Omega_f \\
 &\times \left[\frac{p_f^2 \sin^2 \theta_f}{(E_f - p_f \cos \theta_f)^2} (4E_i^2 - |\mathbf{q}|^2) + \frac{p_i^2 \sin^2 \theta_i}{(E_i - p_i \cos \theta_i)^2} \right. \\
 &\times (4E_f^2 - |\mathbf{q}|^2) + 2\omega^2 \frac{p_i^2 \sin^2 \theta_i + p_f^2 \sin^2 \theta_f}{(E_f - p_f \cos \theta_f)(E_i - p_i \cos \theta_i)} \\
 &\left. - 2 \frac{p_i p_f \sin \theta_i \sin \theta_f \cos \phi}{(E_f - p_f \cos \theta_f)(E_i - p_i \cos \theta_i)} (4E_i E_f - |\mathbf{q}|^2 + 2\omega^2) \right]
 \end{aligned} \tag{7.71}$$

This is known as the Bethe–Heitler formula.

7.3.1

Soft Bremsstrahlung

Now we consider the process of soft Bremsstrahlung. Here we are only interested in the soft photon limit $\omega = k^0 \rightarrow 0$. Then the photon momentum $\mathbf{k}$ in the numerator of the scattering amplitude Eq. (7.65) can be ignored. Then using the

relations

$$\begin{aligned}
 (\not{p}_i + m) \not{\epsilon}^*(k) u(p_i) &= [2(\epsilon^* \cdot p_i) - \not{\epsilon}^*(\not{p}_i - m)] u(p_i) \\
 &= 2(\epsilon^* \cdot p_i) u(p_i) \\
 \bar{u}(p_f) \not{\epsilon}^*(k)(\not{p}_f + m) &= [2(\epsilon^* \cdot p_f) - (\not{p}_f - m) \not{\epsilon}^*] \\
 &= 2\bar{u}(p_f)(\epsilon^* \cdot p_f)
 \end{aligned} \tag{7.72}$$

the Bremsstrahlung amplitude can be reexpressed as

$$\begin{aligned}
 S_{fi} &= 2\pi i \delta(E_i - E_f - \omega) \bar{u}(p_f) \gamma^0 u(p_i) A^0 \\
 &\times \left[e \left(\frac{(\epsilon^* \cdot p_f)}{(k \cdot p_f)} - \frac{(\epsilon^* \cdot p_i)}{(k \cdot p_i)} \right) \right] A^0 = -\frac{Ze^2}{|q|^2}
 \end{aligned} \tag{7.73}$$

If we approximate $q = p_f - p_i + k \simeq p_f - p_i$ in A^0 , this is just the elastic (i.e. Mott) scattering amplitude [see Eq. (6.136)] times a factor for emission of a photon. Thus the Bremsstrahlung cross section for soft photon emission is the Mott scattering cross section times the probability P_γ of emitting a photon:

$$\begin{aligned}
 d\sigma[e(p_i) \rightarrow e(p_f) + \gamma(k)] &= d\sigma_{\text{Mott}}(p_i \rightarrow p_f) dP_\gamma \\
 dP_\gamma &= \frac{d^3k}{2\omega(2\pi)^3} \sum_{\lambda=1,2} e^2 \left| \frac{(\epsilon^\lambda \cdot p_f)}{(k \cdot p_f)} - \frac{(\epsilon^\lambda \cdot p_i)}{(k \cdot p_i)} \right|^2
 \end{aligned} \tag{7.74}$$

Since the polarization sum can be extended to all $\lambda = 0, \dots, 3$, we have $\sum_\lambda \epsilon_\mu^\lambda \epsilon_\nu^\lambda = -g_{\mu\nu}$, and we can write

$$P_\gamma = \frac{\alpha}{\pi} \int \frac{2d\omega}{\omega} \int \frac{d\Omega}{4\pi} \left[\frac{(p_i \cdot p_f)}{(\hat{k} \cdot p_i)(\hat{k} \cdot p_f)} - \frac{1}{2} \left\{ \frac{m^2}{(\hat{k} \cdot p_i)^2} + \frac{m^2}{(\hat{k} \cdot p_f)^2} \right\} \right] \tag{7.75}$$

where $\hat{k} = \mathbf{k}/|\mathbf{k}|$. The integrand goes to ∞ as $\omega \rightarrow 0$. This is known as the infrared divergence. The resolution of this question comes when the one-loop corrections are added properly. We shall discuss the infrared divergence in detail in Sect. 8.1.3. Note that the number of photons increases as $\omega \rightarrow 0$, but the emitted energy remains finite. Here, we simply cut off at $\omega = \mu$, assuming the photon has a small finite mass. The total photon emission probability also diverges formally if integrated over all the momentum range. But remember we have limited $|\mathbf{k}| \ll |\mathbf{q}| \simeq |\mathbf{p}_i - \mathbf{p}_f|$ in making the approximation, so the integral should have an upper limit around $|\mathbf{q}|$. Terms in the braces $\{\dots\}$ are easily calculated to give

$$\int \frac{d\Omega}{4\pi} \frac{m^2}{(\hat{k} \cdot p)^2} = \frac{m^2}{2E^2} \int_{-1}^{+1} d\cos\theta \frac{1}{(1 - \beta \cos\theta)^2} = 1 \tag{7.76}$$

The first term in square brackets $[\dots]$ can be integrated using Feynman's parameterization

$$\frac{1}{ab} = \int_0^1 dx \frac{1}{[ax + b(1-x)]^2} \tag{7.77}$$

Writing $(\hat{k} \cdot p_i) = E_i(1 - \hat{k} \cdot \beta_i) \dots$,

$$\begin{aligned}
 & \int \frac{d\Omega}{4\pi} \frac{(p_i \cdot p_f)}{(\hat{k} \cdot p_i)(\hat{k} \cdot p_f)} \\
 &= \int_0^1 dx \int_{-1}^{+1} \frac{d \cos \theta_k}{2} \frac{1 - \beta_i \cdot \beta_f}{[1 - \hat{k} \cdot \{\beta_f x + \beta_i(1-x)\}]^2} \\
 &= \int_0^1 dx \frac{1 - \beta_i \beta_f \cos \theta}{1 - |\beta_f x + \beta_i(1-x)|^2} \\
 &\simeq \int_0^1 dx \frac{1 - \beta^2 \cos \theta}{1 - \beta^2 + 4\beta^2 \sin^2(\theta/2)x(1-x)} \quad (7.78) \\
 &= \int_0^1 dx \frac{m^2 - q^2/2}{m^2 - q^2 x(1-x)} \\
 &\simeq \begin{cases} 1 + [-q^2/(3m^2)] + O((m^2/(-q^2))^2) & m^2/(-q^2) \gg 1 \\ \ln(-q^2/(m^2)) + O(m^2/(-q^2)) & m^2/(-q^2) \ll 1 \end{cases} \\
 &-q^2 = -(p_f - p_i)^2 \simeq q^2 = 4p^2 \sin^2(\theta/2)
 \end{aligned}$$

where the integration over θ_k was carried out choosing the z-axis along $\beta_f x + \beta_i(1-x)$. The soft Bremsstrahlung cross section is then

$$\begin{aligned}
 \frac{d\sigma}{d\Omega_f} &= \left(\frac{d\sigma}{d\Omega} \right)_{\text{elastic}} \frac{\alpha}{\pi} A_{\text{IR}}(q^2) \ln \frac{\omega_{\text{max}}^2}{\mu^2} \\
 A_{\text{IR}}(q^2) &= \int_0^1 dx \frac{m^2 - q^2/2}{m^2 - q^2 x(1-x)} - 1 \quad (7.79) \\
 &= \begin{cases} \frac{-q^2}{3m^2} + O\left(\left(\frac{-q^2}{m^2}\right)^2\right) & \text{Nonrelativistic} \\ \ln\left(\frac{-q^2}{m^2}\right) - 1 + O\left(\frac{m^2}{-q^2}\right) & \text{Extremely relativistic} \end{cases}
 \end{aligned}$$

The angular distribution of the emitted photon is interesting. When the electron energy is much larger than its mass, i.e. $\beta \rightarrow 1$, the denominators of the expressions in the bracket of Eq. (7.75) go to zero as $\theta \rightarrow 0$ and the radiation is strongly peaked forward:

$$\begin{aligned}
 1 - \beta \cos \theta &= 1 - \left(1 - \frac{1}{\gamma^2(1+\beta)}\right) \cos \theta \simeq \frac{\theta^2}{2} + \frac{1}{2\gamma^2} \\
 \therefore \frac{1/2}{1 - \beta \cos \theta} &\simeq \frac{1}{\theta^2 + \left(\frac{1}{\gamma^2}\right)} = \frac{1}{\theta^2 + \left(\frac{m}{E}\right)^2} \quad (7.80)
 \end{aligned}$$

which means about 1/2 of the total radiation is within a narrow angle of $\theta < 1/\gamma$. The emission, therefore, is strongly concentrated in the initial and the final electron directions.

Note, the soft photon emission probability Eq. (7.79) factorizes as the product of the elastic scattering and the photon emission probability. In the soft photon

limit, the momentum loss of the electron is negligible. Then the electron does not change its state after the soft photon emission and the emission probability of the second soft photon is the same as the first. This means the soft emission is statistically independent of the elastic scattering and we expect a Poisson distribution for the n -photon emission probability. The phenomenon is treated in more detail in Sect. 8.1.4. The electron current is then a classical current source of the photon emission and we expect a semi-classical treatment would produce the same result.

Problem 7.7

Consider a classical source where the momentum of a moving charged particle changes instantaneously at $t = 0$ ¹⁾. The current can be expressed as

$$j^\mu(x) = e \int_{-\infty}^{\infty} d\tau \frac{dx^\mu}{d\tau} \delta^4(x^\mu - x^\mu(\tau))$$

$$x^\mu(\tau) = \begin{cases} (p_i^\mu/m)\tau & \tau < 0 \\ (p_f^\mu/m)\tau & \tau > 0 \end{cases} \quad (7.81)$$

(1) Show that

$$\tilde{j}(k)^\mu = \int d^4x e^{ik \cdot x} j^\mu(x) = ie \left[\frac{p_f^\mu}{(k \cdot p_f)} - \frac{p_i^\mu}{(k \cdot p_i)} \right] \quad (7.82)$$

We may assume the integrand of Eq. (7.81) is multiplied by $e^{-\epsilon|\tau|}$ to assure convergence.

(2) Show that the one-photon emission probability agrees with dP_γ given in Eq. (7.74)

Hint: Calculate $S_{fi} - \delta_{fi} = -i \int d^4x \langle \gamma(k) | A_\mu(x) | 0 \rangle j^\mu(x)$, treating the photon quantum mechanically.

7.4

Feynman Rules

Until now, we have used Feynman diagrams to aid calculation, but there is a deeper, in fact, one-on-one correspondence between the diagrams and the scattering amplitudes. With some experience, the calculations are made much easier. We summarize here the correspondence so that one can reconstruct the scattering amplitude easily by simply looking at the Feynman diagrams. As the scattering amplitude is expressed as $S = T \exp(-i \int d^4x \mathcal{H}_I)$, once we know the interaction Hamiltonian, the corresponding terms are uniquely defined. The electromagnetic interaction Hamiltonian is obtained from the Lagrangian ($\mathcal{H}_I = -\mathcal{L}_I$) by the replacement

1) This is an idealization. In reality, the interaction takes a finite amount of time Δt , in which case the emitted photon energy would be limited $\omega < 1/\Delta t$, providing a natural cutoff for emitted photon energy.

$$\partial_\mu \rightarrow \partial_\mu + iqA_\mu$$

$$\text{spin } 1/2 \text{ particle} \quad -i\mathcal{H}_1 = -iq\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu(x) \quad (7.83a)$$

$$\text{scalar particle} \quad -i\mathcal{H}_1 = qA_\mu[\varphi^\dagger\partial^\mu\varphi - \partial^\mu\varphi^\dagger\varphi] + ie^2A_\mu A^\mu\varphi^\dagger\varphi \quad (7.83b)$$

(1) **Feynman diagrams** First draw all the topologically different diagrams for the desired process.

(2) **External lines** (Fig. 7.6)

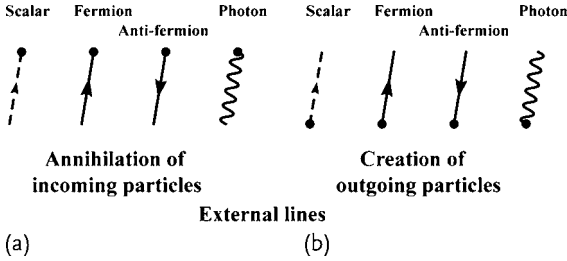


Figure 7.6 Basic elements of the Feynman diagram: external lines.

Assign wave functions to the external lines (incoming or outgoing particles)

		creation		annihilation
scalar boson	$:\varphi^\dagger(x)$	$\rightarrow e^{ik'\cdot x}$	$\varphi(x)$	$\rightarrow e^{-ik\cdot x}$
electron	$:\bar{\psi}(x)$	$\rightarrow \bar{u}(p')e^{ip'\cdot x}$	$\psi(x)$	$\rightarrow u(p)e^{-ip\cdot x}$
positron	$:\psi(x)$	$\rightarrow v(p')e^{ip'\cdot x}$	$\bar{\psi}(x)$	$\rightarrow \bar{v}(p)e^{-ip\cdot x}$
photon	$:A_\mu(x)$	$\rightarrow \varepsilon_\mu(k')^* e^{ik'\cdot x}$	$A_\mu(x)$	$\rightarrow \varepsilon_\mu(k)e^{-ik\cdot x}$

(7.84)

(3) **Vertices** (Fig. 7.7)

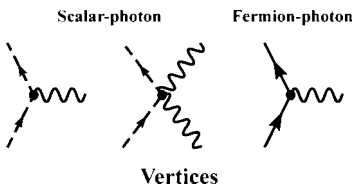


Figure 7.7 Basic elements of the Feynman diagram: vertices.

To the vertices assign

scalar-photon 3-point vertex	$\rightarrow -iq(k + k')^\mu$	
scalar-photon 4-point vertex	$\rightarrow 2iq^2 g^{\mu\nu}$	(7.85)
electron-photon 3-point vertex	$\rightarrow -iq\gamma^\mu$	

(4) **Internal lines** (Fig. 7.8) To the internal lines that connect two vertices, assign the Feynman propagators

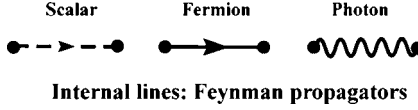


Figure 7.8 Basic elements of the Feynman diagram: internal lines.

scalar	$\langle 0 T[\varphi(x)\varphi(y)] 0 \rangle$	$\rightarrow i\Delta_F(k^2) = \frac{i}{k^2 - m^2 + i\epsilon}$
photon	$\langle 0 T[A_\mu(x)A_\nu(y)] 0 \rangle$	$\rightarrow iD_F(k^2) = \frac{-ig_{\mu\nu}}{k^2 + i\epsilon}$
electron	$\langle 0 T[\psi(x)\bar{\psi}(y)] 0 \rangle$	$\rightarrow iS_F(k^2) = \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}$
massive vector ²⁾	$\langle 0 T[W^\mu(x)W^\nu(y)] 0 \rangle$	$\rightarrow i\Delta_F^{\mu\nu}(k^2)$ $= \frac{-i(g^{\mu\nu} - k^\mu k^\nu / m^2)}{k^2 - m^2 + i\epsilon}$

(7.86)

(5) Momentum assignment (Fig. 7.9)

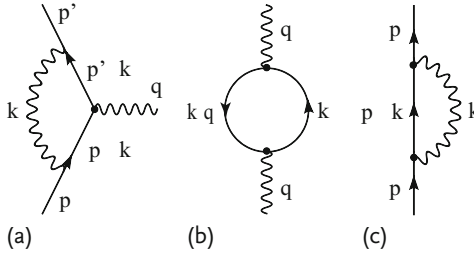


Figure 7.9 Momentum assignment to loop Feynman diagrams. From left to right: vertex correction, photon self-energy and fermion self-energy. For the fermion loops -1 should

be multiplied for each loop. The variable k in the figure is not constrained by δ functions at vertices hence has to be integrated.

Assign k_i to every boson and p_i to every fermion.

external lines	$\rightarrow$	determined by experimental conditions.
internal lines	$\rightarrow$	Starting from external lines, assign k_i, p_i to all the internal lines so as to make the total sum $= 0$ at the vertices.
loop lines	$\rightarrow$	$\int \frac{d^4 k}{(2\pi)^4}$
	$\rightarrow$	multiply by -1 for each fermion loop.

(7.87)

The momentum can be assigned to all the internal lines starting from the external lines, whose momenta are decided by experimental conditions. Space-time

- 2) The difference from the photon is its mass and longitudinal polarization, which gives an extra term in the numerator of the propagator.
- 3) Tree diagrams are those which do not contain loop diagrams.

integrals produce δ -functions at each vertex, which ensure that the sum of the energy momentum is conserved. The δ -functions constrain all the internal momenta in tree diagrams³⁾ and disappear by integration, except the one that ensures the conservation of overall energy-momentum. When loops appear as higher order effects, their internal momentum is not restricted by the δ -function and has to be integrated, hence the factor $\int d^4k/(2\pi)^4$. Examples are shown in Fig. 7.9a–c, where k has to be integrated. For the fermion loops, -1 should be multiplied for each loop.

8

Radiative Corrections and Tests of QED*

8.1

Radiative Corrections and Renormalization*

So far we have calculated the matrix elements only in the lowest order (tree approximation). Most of the essential features of the process are in it, and it is not the author's intention to get involved with a theoretical labyrinth by going into higher orders. However, the appearance of the infinity was inevitable when we tried to calculate higher order corrections in the perturbation series. It is an inherent difficulty deeply rooted in the structure of quantized field theory. We should at least understand the concept of how those infinities are circumvented to make QED a self-consistent mathematical framework. We may as well say that the history of theoretical particle physics has been the struggle to cope with and overcome the infinity problem. The quest is still going on. The standard remedy is “renormalization”, which extracts finite results by subtracting infinity from infinity. The process is rather artificial, but has stood the long and severe test of observations. Today, renormalizability is considered an essential ingredient for the right theory. The higher order corrections are indispensable in comparing theoretical predictions with precision experiments and serve as crucial tests for the internal consistency of the theory. The author considers it important that the reader understand at least the philosophy behind it. For this purpose, we describe the essence of the radiative corrections in order to understand the renormalization prescription.¹⁾

8.1.1

Vertex Correction

We consider first the effect of radiative corrections on the strength of the coupling constant. The strength of the electromagnetic constant is typically measured by a low-energy elastic scattering by nuclei referred to as Mott scattering. The scattering

- 1) This chapter deals with the self-consistency of QED. The discussion is mathematical and rather formal, except for Sect. 8.2, which summarizes the validity of QED. Readers may skip this chapter; it is not compulsory for understanding later chapters. They may come back when they have gained enough understanding and when they are inclined to.

matrix element can be expressed as [see Eq. (6.136)]

$$\mathcal{M}_{\text{Mott}} = -iq\bar{u}(p')\gamma^\mu u(p) \frac{-ig_{\mu\nu}}{q^2} [qj^\nu(q)] \quad (8.1)$$

j^μ is the source of an external field or that created by a static scattering center, which is symbolically denoted by a cross in Fig. 8.1.

So far we have treated the interaction of an electron with the field only in the lowest order, which is depicted as diagram (b). Higher order effects include the vertex correction (c), the photon propagator correction (d), and electron propagator corrections (e,f). Emission of very low energy photons (g,h) is physically indistinguishable from the process that radiates none and should be included in a consistent treatment. We first consider the vertex correction (c) to the process. Applying the Feynman rules described in the previous section, $-iq\gamma^\mu$ in Eq. (8.1) has an extra contribution

$$\begin{aligned} -iq\gamma^\mu &\rightarrow -iq(\gamma^\mu + \Gamma^\mu) \\ \Gamma^\mu &= \int \frac{d^4k}{(2\pi)^4} \frac{-ig_{\rho\sigma}}{k^2} \\ &\times \left[(-iq\gamma^\rho) \frac{i(\not{p}' - \not{k} + m)}{(p' - k)^2 - m^2 + i\epsilon} \gamma^\mu \frac{i(\not{p} - \not{k} + m)}{(p - k)^2 - m^2 + i\epsilon} (-iq\gamma^\sigma) \right] \end{aligned} \quad (8.2)$$

As one can see easily by counting the powers of k , the integral diverges like $\sim \log k^2$ as $k \rightarrow \infty$, which is called the ultraviolet divergence. It also diverges as $k \rightarrow 0$, which is called the infrared divergence. The latter has a logical and physically justifiable interpretation and hence is not a fundamental difficulty.

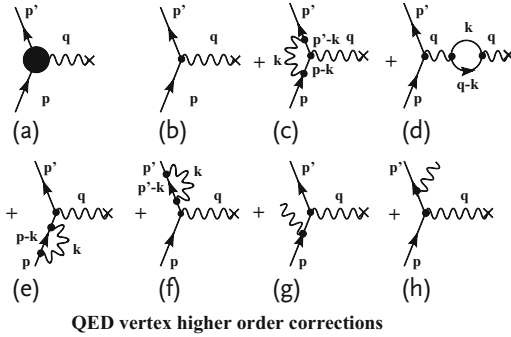


Figure 8.1 Corrections to the vertex function. The effective vertex including overall effects is symbolically represented as the black circle in diagram (a), which in the lowest order is given by the vertex in (b). Higher order effects include the vertex correction (c), the photon propagator correction (d), and the fermion propagator corrections (e, f). Emission of very

low energy photons (g, h) are physically indistinguishable from the process that radiates none. Two kinds of divergences, ultraviolet and infrared, occur once diagrams involving loops are included. The former is treated by renormalization and the latter is canceled by (g, h).

8.1.2

Ultraviolet Divergence

The ultraviolet divergence is inherently connected with the infinite degrees of freedom of the field and rather complicated to handle. A prescription is the renormalization developed by Tomonaga, Schwinger and Feynman.²⁾ To understand it, let us calculate Eq. (8.2) and see how the difficulties arise. Throughout this calculation we tacitly assume that the vertex function is sandwiched by two on-shell spinors. A current operator sandwiched by on-shell spinors is generally expressed as

$$\begin{aligned} \langle p' | j^\mu(x) | p \rangle &\sim \langle p' | \tilde{j}^\mu(q) | p \rangle e^{iq \cdot x} \quad q^\mu = p'^\mu - p^\mu \\ &\sim \bar{u}(p') \left[\gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu} q_\nu}{2m} F_2(q^2) \right] u(p) e^{iq \cdot x} \end{aligned} \quad (8.3)$$

From Lorentz invariance alone, terms proportional to q^μ and $p'^\mu + p^\mu$ should exist. But the current conservation constrains

$$\partial_\mu j^\mu(x) \sim q_\mu \tilde{j}^\mu(q) = 0 \quad (8.4)$$

So there should be no term proportional to q^μ . γ^μ sandwiched with $\bar{u}(p')$ and $u(p)$ gives

$$\begin{aligned} 2m\gamma^\mu &= \not{p}'\gamma^\mu + \gamma^\mu\not{p} = 2p'^\mu - \gamma^\mu(\not{p}' - \not{p}) = 2p'^\mu - \gamma^\mu\not{q} \\ &= 2p^\mu + (\not{p}' - \not{p})\gamma^\mu = 2p^\mu + \not{q}\gamma^\mu = (p'^\mu + p^\mu) - \frac{1}{2}[\gamma^\mu, \not{q}] \\ &= (p'^\mu + p^\mu) + i\frac{1}{2}[\gamma^\mu, \gamma^\nu]q_\nu = (p'^\mu + p^\mu) + i\sigma^{\mu\nu}q_\nu \end{aligned} \quad (8.5)$$

where we have used $\not{p}u(p) = \bar{u}(p')\not{p}' = m$. The relation is called the Gordon equality and shows that $p'^\mu + p^\mu$ can be expressed in terms of γ^μ and $i\sigma^{\mu\nu}q_\nu$. In the static limit, $p'^\mu = p^\mu = (m, 0)$, $q^2 \rightarrow 0$, $iq_\mu = (0, -i\mathbf{q}) \sim (0, \nabla)$, and γ^μ coupled with the electromagnetic field reduces to

$$\langle p' | e \int d^3x e^{iq \cdot x} \bar{\psi}(x) \gamma^\mu \psi(x) A_\mu(x) | p \rangle \xrightarrow{q^2 \rightarrow 0} \left[e\phi + \frac{e}{2m} \boldsymbol{\sigma} \cdot \mathbf{B} \right] \quad (8.6)$$

This expression means γ^μ includes both the electric charge and the magnetic moment of a Dirac particle with $g = 2$. Then $F_1(q^2)$ and $F_2(q^2)$ describe the dynamic part of the charge and the anomalous magnetic moment, which arise from the inner structure of the particle. They are constrained by Eq. (8.6)

$$F_1(0) = 1, \quad F_2(0) = \chi = g - 2 \quad (8.7)$$

From the above argument we expect Γ^μ defined by Eq. (8.2) to have the same structure as Eq. (8.3) with $F_1(0) = 0$ because Γ^μ is defined as a correction to γ^μ .

2) Any standard textbook on quantized field theory has a detailed description of it. We quote some representative books: [63, 64, 215, 322, 338].

In order to calculate Eq. (8.2), we begin with the Feynman parameter trick:

$$\frac{1}{a_1 a_2 \cdots a_n} = (n-1)! \int_0^\infty \cdots \int_0^\infty \prod_{i=1}^n dz_i \frac{\delta(1 - \sum_i z_i)}{(\sum_i a_i z_i)^n} \quad (8.8)$$

We put

$$a_1 = k^2 - \mu^2 + i\varepsilon, \quad a_2 = (p' - k)^2 - m^2 + i\varepsilon, \quad a_3 = (p - k)^2 - m^2 + i\varepsilon \quad (8.9)$$

where we have introduced a small photon mass μ temporarily to avoid the infrared divergence. Then Eq. (8.2) is expressed as

$$\begin{aligned} \Gamma^\mu &= -2ie^2 \int \frac{d^4 k}{(2\pi)^4} \int_0^\infty dz_1 dz_2 dz_3 \delta(1 - z_1 - z_2 - z_3) \\ &\quad \times \frac{N^\mu(k)}{(k'^2 - M^2 + i\varepsilon)^3} \end{aligned} \quad (8.10a)$$

$$\begin{aligned} N^\mu(k) &= \gamma_\nu (\not{p}' - \not{k} + m) \gamma^\mu (\not{p}' - \not{k} + m) \gamma^\nu \\ k' &= k - p' z_2 - p z_3 \\ M^2 &= m^2 (1 - z_1)^2 + \mu^2 z_1 - q^2 z_2 z_3 \end{aligned} \quad (8.10b)$$

Now we change the integration variable from k to k' , and rewrite the numerator in terms of k' . Then terms in odd powers of k' vanish because of symmetry. Rewriting k' as k , the numerator becomes

$$\begin{aligned} N^\mu(k) &\rightarrow N^\mu(k + p' z_2 + p z_3) = N_1 + N_2 + (\text{terms linear in } k^\mu) \\ N_1 &= -2[-k^2 \gamma^\mu + 2k^\mu k^\nu \gamma_\nu] \\ N_2 &= \gamma_\nu [\not{p}'(1 - z_2) - \not{p} z_3 + m] \gamma^\mu [-\not{p}' z_2 + \not{p}(1 - z_3) + m] \gamma^\nu \end{aligned} \quad (8.11)$$

After straightforward but long calculations of Dirac traces, using relations

$$\begin{aligned} \gamma_\nu \gamma^\mu \gamma^\nu &= -2\gamma^\mu, \\ \gamma_\nu \not{A} \not{B} \gamma^\nu &= 4(A \cdot B), \\ \gamma_\nu \not{A} \not{B} \not{C} \gamma^\nu &= -2 \not{C} \not{B} \not{A} \\ \gamma^\mu \not{q} &= q^\mu + i\sigma^{\mu\nu} q_\nu \\ \not{q} \gamma^\mu &= q^\mu - i\sigma^{\mu\nu} q_\nu \end{aligned} \quad (8.12)$$

and Gordon decomposition, N^μ becomes

$$\begin{aligned} N^\mu &= N_1 + N_2 = -2[-k^2 \gamma^\mu + 2k^\mu k^\nu \gamma_\nu + q^2(1 - z_2)(1 - z_3) \gamma^\mu \\ &\quad + m^2(1 - 4z_1 + z_1^2) \gamma^\mu + z_1(1 - z_1) i\sigma^{\mu\nu} m q_\nu \\ &\quad + (2 - z_1)(z_2 - z_3) m q^\mu] \end{aligned} \quad (8.13)$$

The term proportional to q^μ is forbidden by current conservation. Here, the coefficient of q^μ is odd for the exchange $z_2 \leftrightarrow z_3$ and therefore vanishes after integration over z_2 and z_3 , as it must. The first two terms are divergent, as can be seen by

counting the powers of k . Obviously we need to separate the divergent and finite parts and extract a meaningful result out of them. The simplest method is simply to stop integration at some finite cutoff value Λ . But usually that destroys the built-in symmetry. For QED, the Pauli–Villars regularization method which respects the Lorentz invariance has frequently been used.³⁾ But for non-Abelian gauge theory it is unsatisfactory because it does not respect the gauge symmetry and the preferred method is dimensional regularization, where the space-time dimensionality is analytically continued from the real physical value, four, to a complex value D . For real $D < 4$, the integral of the form $\int d^D k / (k^2 + \dots)^2$ converges. When we recover the original dimension $D = 4 - \varepsilon$, $\varepsilon \rightarrow 0$, the divergence reappears in the form of a pole at $D = 4$, which can be subtracted. Other poles appear at $D = 6, 8, \dots$, but it turns out that the integral is analytic in the complex D -plane and we can make analytic continuation to any complex values except those poles. The relevant formulas in this case are [see Eq. (D.20) and discussions in Appendix D]

$$\int \frac{d^D k}{[k^2 - M^2 + i\varepsilon]^n} = i(\pi)^{D/2} \frac{(-1)^n}{(M^2)^{n-D/2}} \frac{\Gamma(n - \frac{D}{2})}{\Gamma(n)} \quad (8.15a)$$

$$\int \frac{k^2 d^D k}{[k^2 - M^2 + i\varepsilon]^n} = i(\pi)^{D/2} (-1)^{n-1} \frac{D/2}{(M^2)^{n-1-D/2}} \frac{\Gamma(n - 1 - \frac{D}{2})}{\Gamma(n)} \quad (8.15b)$$

$$\int \frac{k^\mu k^\nu d^D k}{[k^2 - M^2 + i\varepsilon]^n} = \frac{g^{\mu\nu}}{D} \int \frac{k^2 d^D k}{[k^2 - M^2 + i\varepsilon]^n} \quad (8.15c)$$

where Γ is the well-known gamma function defined by

$$\Gamma(x) = \int_0^\infty dt e^{-t} t^{x-1} \quad (8.16)$$

Setting $n = 3$, defining an infinitesimal number $\varepsilon = 4 - D$ and using

$$\Gamma\left(\frac{\varepsilon}{2}\right) = \frac{2}{\varepsilon} - \gamma_E + O(\varepsilon) \quad (8.17a)$$

$$x^{\varepsilon/2} = e^{(\varepsilon/2)\ln x} = 1 + \frac{\varepsilon}{2} \ln x + O(\varepsilon^2) \quad (8.17b)$$

we obtain

$$\int \frac{d^4 k k^2}{[k^2 - M^2 + i\varepsilon]^3} = i\pi^2 [\Delta - \ln M^2] \quad (8.18a)$$

$$\Delta = \frac{2}{\varepsilon} - \gamma_E \quad (8.18b)$$

3) In this method the propagator is replaced

$$\frac{1}{k^2 - m^2} \rightarrow \frac{1}{k^2 - m^2} - \frac{1}{k^2 - \Lambda^2} = \frac{m^2 - \Lambda^2}{(k^2 - m^2)(k^2 - \Lambda^2)} \quad (8.14)$$

and the integral made finite. After the calculation, $\Lambda \rightarrow \infty$ is taken to recover the original form. The divergent terms appear as Λ^2 or $\ln \Lambda$.

$$\int \frac{d^4 k k^\mu k^\nu}{[k^2 - M^2 + i\varepsilon]^3} = \frac{g^{\mu\nu}}{4} \int \frac{d^4 k k^2}{[k^2 - M^2 + i\varepsilon]^3} \quad (8.18c)$$

$$\int \frac{d^4 k}{[k^2 - M^2 + i\varepsilon]^3} = -\frac{i\pi^2}{2} \frac{1}{M^2} \quad (8.18d)$$

where γ_E is the Euler constant.⁴⁾ The divergent part is included in Δ . Inserting the above equalities, we obtain

$$\begin{aligned} \Gamma^\mu = & \frac{\alpha}{2\pi} \int_0^\infty dz_1 dz_2 dz_3 \delta(1 - z_1 - z_2 - z_3) \left[(\Delta - \ln M^2) \gamma^\mu \right. \\ & + \gamma^\mu \left\{ \frac{m^2(1 - 4z_1 + z_1^2) + q^2(1 - z_2)(1 - z_3)}{M^2} \right\} \\ & \left. + \frac{i\sigma^{\mu\nu} q_\nu}{2m} \frac{2m^2 z_1(1 - z_1)}{M^2} \right] \end{aligned} \quad (8.20)$$

Though we could separate the divergent part, it is by no means unique. If we add some finite part to the infinity, it is still infinity. To determine the infinite part, we note that, in the static limit, the current has to have the form of Eq. (8.6) and Γ^μ has to vanish in the limit $q^2 \rightarrow 0$. This suggests a prescription to subtract terms evaluated at $q^2 = 0$. Then we can write

$$\gamma^\mu + \Gamma^\mu = \gamma^\mu + [\Gamma^\mu(q^2) - \Gamma^\mu(0)] = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}}{2m} F_2(q^2) \quad (8.21a)$$

$$\begin{aligned} F_1(q^2) = & 1 + \frac{\alpha}{2\pi} \int_0^\infty dz_1 dz_2 dz_3 \delta(1 - z_1 - z_2 - z_3) \\ & \times \left[\ln \left(\frac{m^2(1 - z_1)^2 + \mu^2 z_1}{m^2(1 - z_1)^2 - q^2 z_2 z_3 + \mu^2 z_1} \right) \right. \\ & + \frac{m^2(1 - 4z_1 + z_1^2) + q^2(1 - z_2)(1 - z_3)}{m^2(1 - z_1)^2 - q^2 z_2 z_3 + \mu^2 z_1} \\ & \left. - \frac{m^2(1 - 4z_1 + z_1^2)}{m^2(1 - z_1)^2 + \mu^2 z_1} \right] \end{aligned} \quad (8.21b)$$

$$\begin{aligned} F_2(q^2) = & \frac{\alpha}{2\pi} \int_0^\infty dz_1 dz_2 dz_3 \delta(1 - z_1 - z_2 - z_3) \\ & \times \left[\frac{2m^2 z_1(1 - z_1)}{m^2(1 - z_1)^2 + \mu^2 z_1 - q^2 z_2 z_3 - i\varepsilon} \right] \end{aligned} \quad (8.21c)$$

The expressions no longer contains the ultraviolet divergence. The treatment of the finite part of $F_1(q^2)$ will be postponed until we discuss corrections to the electron propagator. The calculation of $F_2(q^2)$ is easier. It does not include the ultraviolet

4) Defined as

$$\gamma_E = \lim_{N \rightarrow \infty} \left[\sum_{n=1}^N \frac{1}{n} - \ln N \right] = 0.57721566 \dots \quad (8.19)$$

or infrared divergence. We can safely set $\mu = 0$. To evaluate the static anomalous magnetic moment, we can set $q^2 = 0$ and

$$\begin{aligned} F_2(0) &= \frac{\alpha}{2\pi} \int_0^\infty dz_1 dz_2 dz_3 \delta(1 - z_1 - z_2 - z_3) \left[\frac{2m^2 z_1(1 - z_1)}{m^2(1 - z_1)^2} \right] \\ &= \frac{\alpha}{2\pi} \int_0^1 dz_1 \int_0^{1-z_1} dz_2 \frac{2z_1}{(1 - z_1)} = \frac{\alpha}{2\pi} \end{aligned} \quad (8.22)$$

Equation (8.7) tells us that the real value of the magnetic moment of the electron $\mu = g(e/2m)s$ is not given by $g = 2$ but $(g - 2)/2 = F_2(0) = \alpha/2\pi \simeq 0.0011614$.

8.1.3

Infrared Divergence

$F_1(q^2)$ needs more elaborate calculation. Here we concentrate on the dominant part as $\mu^2 \rightarrow 0$. It is given as

$$\begin{aligned} F_1(q^2) &\simeq 1 + \frac{\alpha}{2\pi} \int dz_1 dz_2 dz_3 \delta(1 - z_1 - z_2 - z_3) \\ &\quad \times \left[\frac{m^2(1 - 4z_1 + z_1^2) + q^2(1 - z_2)(1 - z_3)}{m^2(1 - z_1)^2 - q^2 z_2 z_3 + \mu^2 z_1} \right. \\ &\quad \left. - \frac{m^2(1 - 4z_1 + z_1^2)}{m^2(1 - z_1)^2 + \mu^2 z_1} \right] \end{aligned} \quad (8.23)$$

The divergence occurs in the integration region where $z_1 \rightarrow 1$, $z_2, z_3 \rightarrow 0$. In this region, we can set $z_1 = 1$, $z_2 = z_3 = 0$ in the numerator of Eq. (8.23). We can also set $\mu^2 z_1 \rightarrow \mu^2$. The integral then becomes

$$\begin{aligned} F_1(q^2) &\simeq 1 + \frac{\alpha}{2\pi} \int_0^1 dz_1 \int_0^{1-z_1} dz_2 \\ &\quad \times \left[\frac{-2m^2 + q^2}{m^2(1 - z_1)^2 - q^2 z_2(1 - z_1 - z_2) + \mu^2} - \frac{-2m^2}{m^2(1 - z_1)^2 + \mu^2} \right] \end{aligned} \quad (8.24)$$

By changing the variables

$$\xi = \frac{z_2}{1 - z_1}, \quad \eta = 1 - z_1, \quad dz_1 dz_2 = -\eta d\xi d\eta \quad (8.25)$$

we obtain

$$\begin{aligned} F_1(q^2) &\simeq 1 + \frac{\alpha}{2\pi} \int_0^1 d\xi \int_0^1 \frac{d(\eta^2)}{2} \\ &\quad \times \left[\frac{-2m^2 + q^2}{[m^2 - q^2 \xi(1 - \xi)]\eta^2 + \mu^2} - \frac{-2m^2}{m^2 \eta^2 + \mu^2} \right] \\ &= 1 - \frac{\alpha}{2\pi} \int_0^1 d\xi \\ &\quad \times \left[\frac{m^2 - q^2/2}{m^2 - q^2 \xi(1 - \xi)} \ln \left\{ \frac{m^2 - q^2 \xi(1 - \xi)}{\mu^2} \right\} - \ln \left(\frac{m^2}{\mu^2} \right) \right] \end{aligned} \quad (8.26)$$

In the limit $\mu^2 \rightarrow 0$, the parameter ξ in the logarithm is not an important part of the integrand. We have already encountered the coefficient of the log function as part of the soft photon emission probability [see Eq. (7.79)]:

$$\begin{aligned} A_{\text{IR}}(q^2) &\equiv \int_0^1 d\xi \frac{m^2 - q^2/2}{m^2 - q^2\xi(1-\xi)} - 1 \\ &\simeq \begin{cases} [-q^2/(3m^2)] + O\{(-q^2/m^2)^2\} & -q^2 \ll m^2 \\ \ln(-q^2/m^2)\{1 + O(m^2/(-q^2))\} - 1 & -q^2 \gg m^2 \end{cases} \end{aligned} \quad (8.27)$$

The final expression for $F_1(q^2)$ becomes

$$F_1(q^2) = 1 - \frac{\alpha}{2\pi} \begin{cases} [-q^2/(3m^2)] \ln(m^2/\mu^2) & -q^2 \ll m^2 \\ \ln(-q^2/m^2) \ln(-q^2/\mu^2) & -q^2 \gg m^2 \end{cases} \quad (8.28)$$

The q^2 dependence of the second line is known as the Sudakov double logarithm. Since this is a second-order correction to the vertex, the contribution to the cross section for the elastic scattering can be written (for the case $q^2 \gg m^2$) as

$$d\sigma(p \rightarrow p') = d\sigma_0 |F_1(q^2)|^2 = d\sigma_0 \left[1 - \frac{\alpha}{\pi} \ln \frac{-q^2}{m^2} \ln \frac{-q^2}{\mu^2} + O(\alpha^2) \right] \quad (8.29)$$

The cross section for the soft bremsstrahlung (contribution from Fig. 8.1g,h) was given in Eq. (7.79):

$$d\sigma(p \rightarrow p') = d\sigma_0 \frac{\alpha}{\pi} \ln \frac{-q^2}{m^2} \ln \frac{\omega_{\text{max}}^2}{\mu^2} \quad (8.30)$$

When the two cross sections are added together, we find

$$d\sigma(p \rightarrow p') = d\sigma_0 \left[1 - \frac{\alpha}{\pi} \ln \frac{-q^2}{m^2} \ln \frac{-q^2}{\omega_{\text{max}}^2} + O(\alpha^2) \right] \quad (8.31)$$

which is a finite, converging result independent of μ^2 , as desired.

Summary of the Infrared Divergence The infrared divergence is induced by low-energy ($k \rightarrow 0$) photons (soft photons). Diagram (c) in Fig. 8.1 is a third-order [$O(e^3)$] effect but it is indistinguishable from the lowest order effect. Therefore, their contributions have to be added to the lowest order vertex diagram (b) and squared to derive a measurable quantity. Diagram (c) produces an $O(e^6)$ effect by itself, but also produces $O(e^4)$ by interfering with (b). This is what we have in Eq. (8.29). Bremsstrahlung processes that emit very low energy photons (g,h) cannot be distinguished from those which emit none, and their contributions have to be added (after squaring the amplitude). This effect is of $O(e^4)$. We have shown that the contribution from emission of a real soft photon and that from the soft-photon loops

compensate each other and no infinities appear. This is true for all orders of perturbations in QED and is demonstrated in the next section. In fact the infrared compensation works for any gauge theory.

The reader may consider that the photon-emitting process is physically distinct from that with no emission. It is certainly true that high-energy photons can be measured and the process is distinct, but consider experimentally. The passage of the electron is measured typically using counters of some kind and its momentum by bending in a magnetic field. The counters have a finite size and finite energy resolution, and low-energy photons that fly close to the electron are measured together with the electron, and cannot be separated. If you say this is the problem of the detector and not the principle, I would argue that even in principle one could never make a detector with infinite accuracy. The detector cannot avoid counting $1, 2, 3, \dots, n$ extra photons, and the sum of all of them becomes finite theoretically. Therefore, for the calculation of the infrared contributions, knowledge of the experimental setup becomes necessary in principle. In practice, the calculation is carried out with some cutoff. If the soft photons are not observed, their contribution can be integrated over the variables. When one calculates the total cross section or inclusive processes (in which one measures some finite number of particles and does not observe the rest), one can derive a finite result independent of the experimental conditions. Thus, we can consider that the infrared divergence is not an inherent problem.

8.1.4

Infrared Compensation to All Orders*

Here we show that the infrared divergence due to photon emission is compensated by that of virtual photon exchange to all orders. The most comprehensive treatment of the infrared divergence is given in [397]. Multiple emission of soft photons was first considered by Bloch and Nordsieck [67] classically. The discussion here follows closely that of Weinberg [382].

We have learned that the soft bremsstrahlung process was well approximated by elastic scattering of an electron at a nucleus multiplied by the probability of emitting a soft photon [$e(p) \rightarrow e(p') + \gamma(k)$]. The cross section in the soft-photon limit was given by Eq. (7.74)

$$\begin{aligned}
 d\sigma(p \rightarrow p' + k) &= d\sigma(p \rightarrow p') P_\gamma \\
 P_\gamma &= \int \frac{d^3k}{2\omega(2\pi)^3} \sum_{\lambda=1,2} e^2 \left| \frac{(\epsilon^\lambda \cdot p_f)}{(k \cdot p_f)} - \frac{(\epsilon^\lambda \cdot p_i)}{(k \cdot p_i)} \right|^2 \\
 &\simeq \frac{\alpha}{\pi} \ln \frac{-q^2}{m^2} \ln \frac{\omega_{\max}^2}{\mu^2} \quad -q^2 \gg m^2
 \end{aligned} \tag{8.32}$$

where μ is the small photon mass introduced to avoid the infrared divergence and $\omega_{\max}$ is the maximum photon energy the detector is capable of detecting. We have seen that the infrared divergence vanishes to order α if we include the virtual photon contribution to the vertex.

Here we intend to show that the cancellation is effective at every order of photon emission. This may seem an impossible task, since there are a variety of ways in which the infrared divergence can enter the various Feynman diagrams. However, the problem is made tractable by extracting only those in which soft photons are emitted from the incoming or outgoing electrons on the mass shell.

The Feynman diagram where a total of n soft photons are emitted from on-shell electrons has the structure depicted in Fig. 8.2a. Only soft photons emitted from (nearly) on-shell electrons contribute, because the denominator of the electron propagator before (or after) the photon emission is given as

$$\frac{1}{(p \pm k_1)^2 - m^2 + i\epsilon} = \frac{1}{(p^2 - m^2) \pm 2(p \cdot k_1)} \quad (8.33)$$

and becomes large only when $p^2 - m^2 = 0$ and $k_1 \rightarrow 0$. Contributions of hard photons or soft photons emitted from deeply off-shell electrons, where $(p \pm k)^2 - m^2 \gg m^2$, are negligible.

First we consider a process where all the soft photons are emitted from the outgoing electron. The amplitude has the following form, where we have numbered

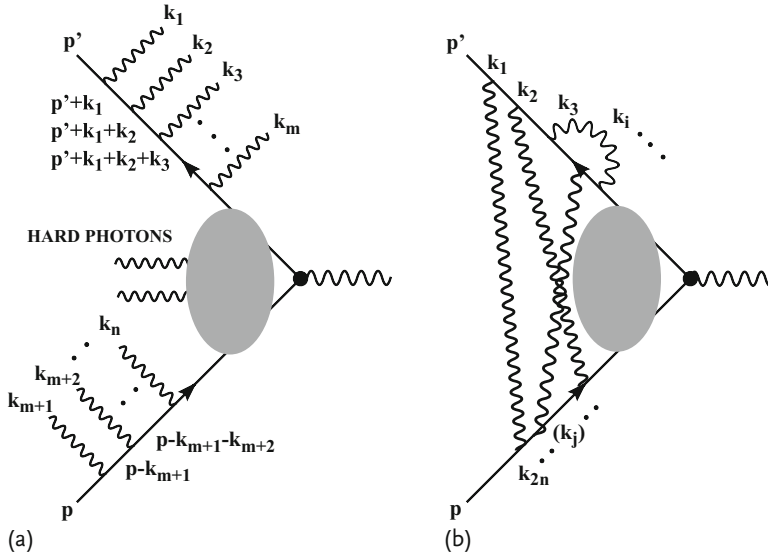


Figure 8.2 Diagrams that contribute to the infrared divergence. (a) Only soft photons emitted from on-shell electrons, namely those incoming and outgoing electrons at the outer legs give large contribution. Soft photons or hard photons from deep off-shell electrons

(denoted by the gray blob) do not contribute much. (b) Only soft-photon internal lines exchanged by on-shell electrons contribute to cancel the divergence created by real photon emission.

the photons as in Fig. 8.2:

$$\begin{aligned} & \bar{u}(p') \cdots (-i q \not{\epsilon}_j^*) \frac{i(\not{p}' + \not{k}_1 + \cdots + \not{k}_j) + m}{2(p' \cdot k_1 + \cdots + k_j)} (-i q \not{\epsilon}_{j+1}^*) \\ & \times \frac{i(\not{p}' + \not{k}_1 + \cdots + \not{k}_{j+1}) + m}{2(p' \cdot k_1 + \cdots + k_{j+1})} \cdots u(p) \end{aligned} \quad (8.34)$$

We are only interested in the soft-photon limit ($\omega = |k^0| \rightarrow 0$). Therefore, we can neglect $\not{k}$ terms in the numerator. Next we move $\not{p}'$ to the left and use $\bar{u}(p')(\not{p}' - m) = 0$. Namely

$$\begin{aligned} & \bar{u}(p') \not{\epsilon}^*(\not{p}' + m) \not{\epsilon}^*(\not{p}' + m) \cdots u(p) \\ & = \bar{u}(p') [2(\epsilon^* \cdot p') - (\not{p}' - m) \not{\epsilon}^*] (\not{p}' + m) \cdots u(p) \\ & = \bar{u}(p') 2(\epsilon^* \cdot p') \not{\epsilon}^*(\not{p}' + m) \cdots u(p) \\ & = \bar{u}(p') 2(\epsilon^* \cdot p') 2(\epsilon^* \cdot p') \cdots u(p) \end{aligned} \quad (8.35)$$

Thus all the numerators can be replaced with $(2\epsilon^* \cdot p')$. Furthermore, we can use the formula

$$\begin{aligned} & \sum_{\text{perm}} \frac{1}{(p' \cdot k_1)} \frac{1}{(p' \cdot (k_1 + k_2))} \cdots \frac{1}{(p' \cdot (k_1 + \cdots + k_n))} \\ & = \frac{1}{(p' \cdot k_1)} \frac{1}{(p' \cdot k_2)} \cdots \frac{1}{(p' \cdot k_n)} \end{aligned} \quad (8.36)$$

which can easily be proven for $n = 2$:

$$\frac{1}{(p' \cdot k_1)} \frac{1}{(p' \cdot (k_1 + k_2))} + (1 \leftrightarrow 2) = \frac{1}{(p' \cdot k_1)} \frac{1}{(p' \cdot k_2)} \quad (8.37)$$

The result can be proved for general n by induction. Then the scattering amplitude with subsequent n -photon emission that is effective at $k_j \rightarrow 0$ can be written as

$$\bar{u}(p') i \mathcal{M}_{el} u(p) \prod_{j=1}^n \frac{e(\epsilon^* \cdot p')}{(p' \cdot k_j)} \quad (8.38a)$$

Next we assume all the soft photons are emitted from the incoming electron and repeat the same procedure. We obtain

$$\bar{u}(p') i \mathcal{M}_{el} u(p) \prod_{j=1}^n \left(-\frac{e(\epsilon^* \cdot p)}{(p \cdot k_j)} \right) \quad (8.38b)$$

which can be obtained from Eq. (8.38a) by the substitutions

$$p' \rightarrow p, \quad k_j \rightarrow -k_j$$

as can be seen in the lower part of Fig. 8.2a. In reality, the n photons can be emitted from either the outgoing or the incoming electron. For this case, we have

$$\begin{aligned}
 A_{\text{real}} &= \bar{u}(p') i \mathcal{M}_{el} u(p) \sum_{\text{perm.}} \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} e^{\frac{(\epsilon^* \cdot p')}{(p' \cdot k_i)}} \left(-e^{\frac{(\epsilon^* \cdot p)}{(p \cdot k_j)}} \right) \\
 &= \bar{u}(p') i \mathcal{M}_{el} u(p) \prod_{j=1}^n e^{\left[\frac{(\epsilon^* \cdot p')}{(p' \cdot k_j)} - \frac{(\epsilon^* \cdot p)}{(p \cdot k_j)} \right]}
 \end{aligned} \tag{8.39}$$

Then the probability of emitting n photons is obtained by squaring the amplitude, taking the polarization sum, which gives $\sum_{\lambda} (\epsilon \cdot A)(\epsilon^* \cdot B) = -(A \cdot B)$, and multiplying by the final-state phase space volume. Note $1/n!$ comes in because of the Bose statistics of the photon. Summing over all n 's, we obtain

$$\begin{aligned}
 d\sigma &= d\sigma(p \rightarrow p') \sum_{n=0}^{\infty} \frac{1}{n!} R^n = d\sigma(p \rightarrow p') \exp[R] \\
 R &= \int \frac{d^3 k}{(2\pi)^3 2\omega} e^2 \sum_{\lambda} \left(\frac{(\epsilon_{\lambda}^* \cdot p')}{(p' \cdot k)} - \frac{(\epsilon_{\lambda}^* \cdot p)}{(p \cdot k)} \right)^2 \\
 &\stackrel{(7.79)}{\simeq} \frac{\alpha}{\pi} \ln \left(\frac{-q^2}{m^2} \right) \ln \left(\frac{\omega_{\text{max}}^2}{\mu^2} \right) \quad \text{for } -q^2 = -(p' - p)^2 \gg m^2
 \end{aligned} \tag{8.40}$$

where ω_{max} is the maximum energy the detector is capable of seeing and μ is the photon mass temporarily introduced to avoid the infrared divergence. Note the one-photon emission probability R was already given as P_{γ} of the bremsstrahlung cross section $d\sigma(p \rightarrow p' + \gamma) = d\sigma(p \rightarrow p') P_{\gamma}$ in Eqs. (7.74) and (7.79).

Next we consider the corresponding internal soft-photon contributions. Corresponding to the n real-photon emission cross section, we need to consider an interference term of order $2n$ (n soft internal lines) with the 0th-order diagram. First we arrange $2n$ external photons arbitrarily on either outgoing or incoming electron lines. Then we pick up a pair (k_i, k_j) of photon lines and connect them to make an internal line. Recollecting that to each external photon line a factor $(\epsilon^* \cdot p')/(p' \cdot k_i)$ or $-(\epsilon^* \cdot p)/(p \cdot k_i)$ is attached, depending on whether the i, j th line is emitted from the outgoing or the incoming electron, we have an amplitude for each internal line

$$A_{\text{internal}} = e^2 \int \frac{d^4 k_i}{(2\pi)^4} \frac{-i}{k_i^2 + i\epsilon} \left(\frac{p_i}{(p_i \cdot k_i)} \right) \cdot \left(\frac{p_j}{-(p_j \cdot k_i)} \right) \tag{8.41}$$

where the dot product is produced by $-g_{\mu\nu}$ ($= \sum \epsilon_{\mu}^* \epsilon_{\nu}$) in the numerator of the photon propagator. p_i and p_j stand for p' or p in the numerator and p' or $-p$ in the denominator, depending on whether the photon i, j is emitted from the outgoing or incoming line. The extra minus sign in the second denominator appears because if we define k as the photon momentum emitted by the line i then k must be absorbed by line j .

In making n internal lines out of $2n$ external lines for all possible combinations, we use Eq. (8.39) to express $2n$ external photon lines, choose lines i and j ($i \neq j$)

and connect pairwise to give

$$\begin{aligned}
 & \sum_{\substack{\text{perm} \\ n_1 + n_2 = 2n}} \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} \frac{(\varepsilon_i^* \cdot p')}{(p' \cdot k_i)} \left(-\frac{(\varepsilon_j^* \cdot p)}{(p \cdot k_j)} \right) \\
 &= \prod_{i=1}^n \prod_{j=1}^n \left[\frac{(\varepsilon_i^* \cdot p')}{p' \cdot k_i} - \frac{(\varepsilon_i^* \cdot p)}{p \cdot k_i} \right] \left[\frac{(\varepsilon_j^* \cdot p')}{p' \cdot k_j} - \frac{(\varepsilon_j^* \cdot p)}{p \cdot k_j} \right] \\
 &\rightarrow -\prod_i^n \left[\left(\frac{p'}{(p' \cdot k_i)} - \frac{p}{(p \cdot k_i)} \right) \cdot \left(\frac{p'}{(p' \cdot k_j)} - \frac{p}{(p \cdot k_j)} \right) \right]_{k_j = -k_i}
 \end{aligned} \tag{8.42}$$

where $\sum \varepsilon_i^{\mu*} \varepsilon_j^\nu|_{i=j} = -g^{\mu\nu}$ was used in going to the third line. We have to divide the sum by $2^n n!$ because, on integrating over k_i , interchanging n virtual photons and two directions of a pair with each other produces the same diagrams. Therefore the sum of all different combinations gives

$$A_{n,\text{virtual}} = \frac{V^n}{n!} \tag{8.43}$$

$$\begin{aligned}
 V &= \frac{e^2}{2} \int \frac{d^4 k}{(2\pi)^4} \frac{-i}{k^2 + i\varepsilon} \left(\frac{p'}{(p' \cdot k)} - \frac{p}{(p \cdot k)} \right) \cdot \left(\frac{p'}{-(p' \cdot k)} - \frac{p}{-(p \cdot k)} \right) \\
 &\simeq -\frac{\alpha}{2\pi} \ln \left(\frac{-q^2}{m^2} \right) \ln \left(\frac{-q^2}{\mu^2} \right) \quad \text{for } -q^2 \gg m^2
 \end{aligned} \tag{8.44}$$

Derivation of the last line can be done following the same logic as was done to calculate the one loop vertex correction function of Eq. (8.2). In fact, one can recognize that V is just its infrared limit which resulted in Eq. (8.28). Virtual photon exchange diagrams may contain m extra external photon emissions, but they factorize and constitute a part of Eq. (8.40). Therefore adding $0, 1, 2, \dots, n, \dots$ virtual photon exchange contributions, we can obtain the correction factor A_{virtual} that multiplies the elastic scattering amplitude. The scattering cross section for emission of any number of photons becomes

$$\begin{aligned}
 \sum_{n=0}^{\infty} d\sigma(p \rightarrow p' + n\gamma) &= d\sigma_0 \left(\sum_{n=0}^{\infty} \frac{R^n}{n!} \right) \left| \sum_{n=0}^{\infty} \frac{V^n}{n!} \right|^2 \\
 &= d\sigma_0 \exp(R) \exp(2V) \\
 &= d\sigma_0 \times \exp \left[\frac{\alpha}{\pi} \ln \left(\frac{-q^2}{m^2} \right) \ln \left(\frac{\omega_{\text{max}}^2}{\mu^2} \right) \right] \\
 &\quad \times \exp \left[-\frac{\alpha}{\pi} \ln \left(\frac{-q^2}{m^2} \right) \ln \left(\frac{-q^2}{\mu^2} \right) \right] \\
 &= d\sigma_0 \exp \left[-\frac{\alpha}{\pi} \ln \left(\frac{-q^2}{m^2} \right) \ln \left(\frac{-q^2}{\omega_{\text{max}}^2} \right) \right] \quad \text{for } q^2 \gg m^2
 \end{aligned} \tag{8.45}$$

This expression depends on the detector sensitivity, but no longer on the infrared cutoff μ^2 . Unlike the $O(\alpha)$ correction, which could go to $-\infty$,⁵⁾ the correction factor is finite and lies between 0 and 1. If we set $\omega_{\min}$ instead of $\omega_{\max}$ as the minimum energy the detector can detect, the probability of a scattering cross section with no photon emission is

$$d\sigma(p \rightarrow p' + 0\gamma) = d\sigma_0 \exp \left[-\frac{\alpha}{\pi} \ln \left(\frac{-q^2}{m^2} \right) \ln \left(\frac{-q^2}{\omega_{\min}^2} \right) \right] \quad (8.46)$$

which decreases faster than any power of q^2 . The exponential correction factor contains the Sudakov double logarithm and is known as the Sudakov form factor.

***n*-Photon Emission**

If we set up a detector to observe any number of photons with energy $\omega_{\min} < \omega < \omega_{\max}$, the probability of finding n photons would be proportional to $R^n/n!$, where R is given by integrating the photon energy from $\omega_{\min}$ to $\omega_{\max}$ in Eq. (8.40)

$$\begin{aligned} P_n &\propto \frac{1}{n!} R^n \\ R &= \int \frac{d^3k}{(2\pi)^3 2\omega} e^2 \sum_{\lambda} \left(\frac{(\varepsilon_{\lambda}^* \cdot p')}{(p' \cdot k_i)} - \frac{(\varepsilon_{\lambda}^* \cdot p)}{(p \cdot k_i)} \right)^2 \\ &\simeq \frac{\alpha}{\pi} \ln \left(\frac{-q^2}{m^2} \right) \ln \left(\frac{\omega_{\max}^2}{\omega_{\min}^2} \right) \quad \text{for } -q^2 = -(p' - p)^2 \gg m^2 \end{aligned} \quad (8.47)$$

As the total probability of observing any number of photons should equal 1, we can calculate the normalization constant N from

$$1 = N \sum_{n=0}^{\infty} \frac{R^n}{n!} = N e^R \quad (8.48a)$$

$$P_n = \frac{R^n}{n!} e^{-R}, \quad R = -\frac{\alpha}{\pi} \ln \left(\frac{-q^2}{m^2} \right) \ln \left(\frac{\omega_{\max}^2}{\omega_{\min}^2} \right) \quad (8.48b)$$

This gives a Poisson distribution, as we expected from discussions in Sect. 7.3.1.

8.1.5

Running Coupling Constant

We now turn to other higher order corrections that involve the self-energy of the photon and the electron. Let us consider the photon propagator (Fig. 8.3).

When the contribution of the loop correction (second diagram on the right hand side, rhs) is added, the photon propagator is modified from the lowest order ex-

5) The whole perturbation argument breaks down when $\alpha \ln(m/\mu)$ becomes comparable to unity, which occurs at $\mu \sim e^{-137} m \sim 10^{-60} m$. Physically this is an absurd situation, but mathematically it is a problem.

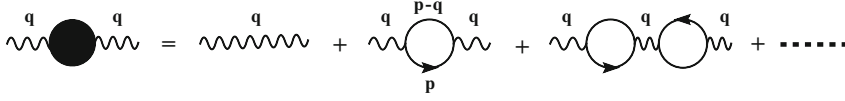


Figure 8.3 When the higher order radiative corrections are added to the lowest propagator D_F , it is dressed to become $iD_F' = iD_F + iD_F(-i\Sigma_{\gamma\gamma})iD_F + \dots$. Here only ladder-type corrections are summed and included.

pression

$$\frac{-ig^{\mu\nu}}{q^2} \rightarrow \frac{g^{\mu\nu}}{iq^2} + \frac{g_\lambda^\mu}{iq^2}(-i\Pi^{\lambda\sigma})\frac{g_\sigma^\nu}{iq^2} = \left[g^{\mu\nu} + \frac{1}{iq^2}(-i\Pi^{\mu\nu}) \right] \frac{1}{iq^2} \quad (8.49a)$$

$$-i\Pi^{\mu\nu} = -\int \frac{d^4 p}{(2\pi)^4} \text{Tr} \left[ie\gamma^\mu \frac{i(\not{p} + m)}{p^2 - m^2 + i\varepsilon} ie\gamma^\nu \frac{i(\not{p} - \not{q} + m)}{(p-q)^2 - m^2 + i\varepsilon} \right] \quad (8.49b)$$

There is an extra minus sign multiplied for each fermion loop. It originates from the anticommutativity of the fermion field. The first term in the first line is the lowest order photon propagator, which is the Feynman propagator of the free field. Counting the powers of k , the integral seems to diverge as $\sim k^2$. However, because of gauge invariance, $\Pi^{\mu\nu}$ satisfies the constraint

$$q_\mu \Pi^{\mu\nu} = q_\nu \Pi^{\mu\nu} = 0 \quad (8.50)$$

which means its functional form has to be of the type

$$\Pi^{\mu\nu} = (q^2 g^{\mu\nu} - q^\mu q^\nu) \Pi_\gamma \quad (8.51)$$

Therefore, the powers of k are effectively reduced by 2 in Π_γ , and it only diverges as $\log k^2$, just like the loop integral of the diagram in Fig. 8.1c. For the sake of later use, we define

$$\Sigma_{\gamma\gamma} \equiv q^2 \Pi_\gamma \quad (8.52a)$$

$$\rightarrow \Pi^{\mu\nu} = \left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2} \right) \Sigma_{\gamma\gamma} \quad (8.52b)$$

If we expand $\Pi_\gamma(q^2)$ in Taylor series⁶⁾

$$\Pi_\gamma(q^2) = \Pi_\gamma(0) + q^2 \frac{\partial}{\partial q^2} \Pi_\gamma(q^2) \Big|_{q^2=0} + \dots \equiv \Pi_\gamma(0) + \hat{\Pi}_\gamma(q^2) \quad (8.53)$$

By definition $\hat{\Pi}_\gamma(0) = 0$. The power of k^2 in the integral Eq. (8.49b) which contributes to the coefficients of the Taylor series reduces as the power of q^2 increases. Since the constant term $\Pi_\gamma(0)$ contains the highest power of k^2 , if it diverges at

6) It can be shown that generally $\Pi_\gamma(q^2)$ does not have a pole at $q^2 = 0$.

most logarithmically, we expect that coefficients of the higher q^2 powers do not diverge and are finite.

Indeed, actual calculation (see Appendix D for details) shows that the divergent term is included only in $\Pi_\gamma(0)$:

$$\Pi_\gamma(q^2) = \Pi_\gamma(0) - \frac{2\alpha}{\pi} \int_0^1 dz z(1-z) \ln \left[1 - \frac{q^2}{m^2} z(1-z) \right] \quad (8.54)$$

$$\Pi_\gamma(0) = \frac{\alpha}{3\pi} (\Delta - \ln m^2) \quad (8.55)$$

Δ is the divergent term and is given by Eq. (8.18). If we use dimensional regularization, and carry out the integration in D -dimensional space, the integral converges for $D - \varepsilon$, $\varepsilon > 0$. The divergence comes back when we make $\varepsilon \rightarrow 0$ as Δ contains a pole term $1/\varepsilon$. The second term is well defined and finite and behaves like

$$\hat{\Pi}_\gamma(q^2) = \begin{cases} \frac{\alpha q^2}{15\pi m^2} & -q^2 \rightarrow 0 \\ -\frac{\alpha}{3\pi} \ln \frac{-q^2}{m^2} & -q^2 \rightarrow \infty \end{cases} \quad (8.56)$$

Including the second-order effect, the lowest order Mott scattering amplitude Eq. (8.1) is modified to

$$ie\bar{u}(p')\gamma^\mu u(p) \left[\frac{g_{\mu\nu}}{iq^2} + \frac{g_{\mu\lambda}}{iq^2} (-i\Pi^{\lambda\sigma}) \frac{g_{\sigma\nu}}{iq^2} \right] [ej^\nu(q)] \quad (8.57a)$$

The term proportional to $q^\mu q^\nu$ vanishes when coupled with the conserved current ($q_\mu j^\mu = 0$ or $q^\mu \bar{u}\gamma_\mu u = 0$). The above expression gives

$$\begin{aligned} ie\bar{u}(p')\gamma^\mu u(p) \frac{-ig_{\mu\nu}}{q^2} [1 - \Pi_\gamma(0) - \hat{\Pi}_\gamma(q^2)] [ej^\nu(q)] \\ \simeq ie\bar{u}(p')\gamma^\mu u(p) \frac{-ig_{\mu\nu}}{q^2} [1 + \Pi_\gamma(0)]^{-1} [1 - \hat{\Pi}_\gamma(q^2)] [ej^\nu(q)] \end{aligned} \quad (8.57b)$$

Then defining

$$\Pi_\gamma(0) = \frac{\alpha}{3\pi} (\Delta - \ln m^2) = \frac{1}{Z_3} - 1 \equiv \delta Z_3 \quad (8.58a)$$

$$e_R = e [1 + \Pi_\gamma(0)]^{-1/2} \sim Z_3^{1/2} e \quad (8.58b)$$

we rewrite Eq. (8.57) to give

$$\rightarrow ie_R \bar{u}(p')\gamma^\mu u(p) \frac{-ig_{\mu\nu}}{q^2} [1 - \hat{\Pi}_\gamma(q^2)] [e_R j^\nu(q)] \quad (8.59a)$$

$$\alpha \rightarrow \alpha_R = \frac{e_R^2}{4\pi} \quad (8.59b)$$

We shall call e the bare coupling constant and e_R the renormalized coupling constant. The bare coupling constant, which appears in the original Lagrangian, is

then not the true coupling constant. The true coupling constant should include all the higher order corrections and is an observable because it appears in measurable quantities such as cross sections. Here we have treated $\delta Z_3 = \Pi_\gamma(0)$ as a small number in the perturbation series in terms of α , though actually it is divergent.

In effect, we have subtracted infinity from infinity to obtain a finite expression, which is not allowed because it is an ambiguous procedure. That is what we learned in high-school mathematics. The justification is as follows. Higher order corrections for the photon come in because of the interactions with charged particles. But if $q^2 = 0$, the photon is on shell, and it is a freely propagating wave. The field is completely decided by the equation of motion without a source. We know the propagator for the free photon. It is the Green's function to the free Klein–Gordon equation, which is $-ig^{\mu\nu}/q^2$. Whatever form the modified propagator takes, it has to respect this condition. Therefore, in principle, the correction to the lowest order term must vanish at $q^2 = 0$ or $\Pi_\gamma(0) = 0$ must hold. In practice it does not, so we subtract it as redundant.

Now, the photon can make an electron–positron pair at any time and acquire an extra self-energy. So applying the replacement Eq. (8.49) to the photon propagator successively, the net effect is to change $D_F \rightarrow D'_F$ (to dress the photon), the diagrams of which are depicted in Fig. 8.3. Only ladder-type diagrams are collected here. More complex diagrams have to be included if we go to higher orders.

Then the expression for D'_F is given by

$$\begin{aligned}
 D_F'^{\mu\nu}(q^2) &= \frac{g^{\mu\nu}}{iq^2} \left[1 + (-i\Sigma_{\gamma\gamma}) \frac{1}{iq^2} + (-i\Sigma_{\gamma\gamma}) \frac{1}{iq^2} (-i\Sigma_{\gamma\gamma}) \frac{1}{iq^2} + \cdots \right] \\
 &= \frac{g^{\mu\nu}}{i[q^2 + \Sigma_{\gamma\gamma}(q^2)]} = \frac{g^{\mu\nu}}{iq^2} \frac{1}{1 + \Pi_\gamma(q^2)} \\
 &= \frac{g^{\mu\nu}}{iq^2} \frac{1}{1 + \Pi_\gamma(0) + \hat{\Pi}_\gamma(q^2)} \equiv \frac{g^{\mu\nu}}{iq^2} \frac{1}{Z_3^{-1} + \hat{\Pi}_\gamma(q^2)} \\
 &= \frac{g^{\mu\nu}}{iq^2} \frac{Z_3}{1 + Z_3 \hat{\Pi}_\gamma(q^2)} \tag{8.60}
 \end{aligned}$$

As the propagator is always sandwiched between two vertices, it accompanies e^2 [see Eq. (8.49)]. Therefore, we move the Z_3 factor to them and rename

$$e^2 \rightarrow e_R^2 = Z_3 e^2 \tag{8.61}$$

$\hat{\Pi}_\gamma$ itself is proportional to e^2 and appears as multiplied by Z_3 . Therefore we can replace e^2 in $\hat{\Pi}_\gamma$ by e_R^2 , too.

$$Z_3 \hat{\Pi}_\gamma(q^2, e^2) = \hat{\Pi}_\gamma(q^2, e_R^2) \tag{8.62}$$

Although Z_3 is a diverging constant, if we consider Z_3 and consequently e_R to be finite, the photon propagator can still be expressed as $g^{\mu\nu}/q^2$, but the coupling constant modified to

$$e_R^2(q^2) \equiv \frac{Z_3 e^2}{1 + Z_3 \hat{\Pi}_\gamma(q^2, e^2)} = \frac{e_R^2}{1 + \hat{\Pi}_\gamma(q^2, e_R^2)} \tag{8.63}$$

e_R^2 is now a function of q^2 and is referred to as the effective or running coupling constant. Apart from the treatment of the infinite constants, all the higher order corrections to the photon propagator result in simply making the coupling constant run. From the above argument, the ultraviolet problem, as far as the photon correction is concerned, can be contained in the coupling constant, a procedure called renormalization.

Finally, we give a form frequently quoted as the running coupling constant of QED at high energy. Substituting Eq. (8.56) in Eq. (8.63) we have at large $Q^2 = -q^2 > 0$

$$\alpha(Q^2) = \frac{\alpha}{1 - \frac{\alpha}{3\pi} \ln \frac{Q^2}{m_e^2}} \quad (8.64)$$

where $\alpha = \alpha(0)$. The screening due to vacuum polarization gives the smallest value of α at infinite distance. Conversely α becomes larger at larger value of Q^2 or at shorter distance.

8.1.6

Mass Renormalization

Next let us consider the radiative corrections to the electron (Fermion) propagator. As an example consider the Compton scattering amplitude. Its inner line corresponds to the propagator

$$i S_F(p) = \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} = \frac{i}{\not{p} - m} \quad (8.65)$$

The next higher order correction is a process to emit and reabsorb a photon, as shown in Fig. 8.4. Using the same argument for the photon propagator, we may sum all the ladder-type repetition of the self-energy:

$$i S_F(p) \rightarrow i S'_F(p) = \frac{i}{\not{p} - m} + \frac{i}{\not{p} - m} (i \Sigma_e(p)) \frac{i}{\not{p} - m} + \dots \quad (8.66a)$$

$$i \Sigma_e(p) = \int \frac{d^4 k}{(2\pi)^4} \frac{-ig_{\mu\nu}}{k^2 - \mu^2 + i\epsilon} (ie\gamma^\mu) \frac{i}{\not{p} - \not{k} - m + i\epsilon} (ie\gamma^\nu) \quad (8.66b)$$

where we introduced a small photon mass μ to avoid the infrared divergence. As we can see by counting the powers of k , the integral diverges as $k \rightarrow \infty$. $\Sigma_e(p)$ can be calculated again using the same regularization method as the photon (see

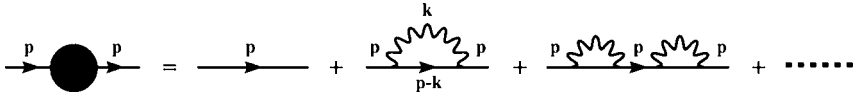


Figure 8.4 $i S'_F = i S_F + i S_F (-i \Sigma_e) i S_F + \dots$

Appendix D). Then $S'_F(p)$ can be rewritten as

$$S'_F(p) = \frac{i}{\not{p} - m + i\varepsilon} \left(1 + \Sigma_e(p) \frac{1}{\not{p} - m} \right)^{-1} = \frac{i}{\not{p} - m + \Sigma_e(p)} \quad (8.67)$$

Note, the pole of the propagator has shifted, meaning the mass value has changed as the result of higher order corrections. Then the original mass term m , called the bare mass, that appeared in the Lagrangian is not the true mass, but a parameter having no physical meaning. The true mass should include all the corrections and the propagator should have the pole at the correct (or observed) mass, which we denote by m_R . Then we perform the Taylor expansion around $\not{p} = m_R$ in the same spirit as we did for the photon propagator around $q^2 = 0$:

$$\Sigma_e(p) = \delta m + [\delta Z_2 + \hat{\Pi}_e(\not{p})](\not{p} - m_R) \quad (8.68a)$$

$$\delta m = \Sigma_e(p)|_{\not{p}=m_R} = \Sigma_e(m) + O(\alpha^2) \simeq -\frac{3\alpha}{4\pi} m \Delta \quad (8.68b)$$

$$\delta Z_2 \equiv \frac{1}{Z_2} - 1 = \left. \frac{\partial \Sigma_e(p)}{\partial \not{p}} \right|_{\not{p}=m_R} \simeq +\frac{\alpha}{4\pi} \Delta \quad (8.68c)$$

$$\hat{\Pi}_e(m_R) = 0 \quad (8.68d)$$

where $\simeq$ means that some finite contributions are omitted. The diverging quantities are contained in δm and δZ_2 . After they are subtracted, the remaining $\hat{\Pi}_e$ is finite. $\hat{\Pi}_e(m_R) = 0$ since it is the third term of the Taylor expansion around $\not{p} = m_R$. Using the above quantities, S'_F (8.66) can be rewritten as

$$\begin{aligned} S'_F(p) &= \frac{i}{\not{p} - m + \Sigma_e(p)} = \frac{i}{[Z_2^{-1} + \hat{\Pi}_e(p^2)](\not{p} - m_R)} \\ &= \frac{i Z_2}{[1 + Z_2 \hat{\Pi}_e(p^2)](\not{p} - m_R)} \\ m_R &= m - \delta m \end{aligned} \quad (8.69)$$

As the propagator accompanies e^2 also, the replacement $e^2 \rightarrow e'^2_R = Z_2 e^2$ removes the diverging Z_2 , in the same way as we removed Z_3 . Therefore

$$i e^2 S_F(p) \rightarrow i e'^2_R S'_F = \frac{e'^2_R}{1 + \hat{\Pi}_e(p^2)} \frac{1}{\not{p} - m_R} \quad (8.70a)$$

$$e'^2_R = Z_2 e^2, \quad m_R = m - \delta m \quad (8.70b)$$

Thus we have shown that the renormalization effect is to make the replacement $m \rightarrow m_R$ and multiply the electron propagator by Z_2 and the photon propagator Z_3 . Z_2 and Z_3 are absorbed in the definition of the running coupling constant:

$$e^2 \rightarrow \frac{e_R^2}{(1 + \hat{\Pi}_e(p^2)(1 + \hat{\Pi}_\gamma(q^2))} \times F_1(q^2) \quad (8.71)$$

where F_1 is a vertex correction Eq. (8.24). Note, when the higher order corrections are applied to the external lines, $p^2 = p'^2 = m_R^2$ or $q^2 = 0$ for the photon, $\hat{\Pi}_e(m_R) = \hat{\Pi}_\gamma(0) = 0$ holds. Therefore, the effect of the renormalization on the external line is only to change $m \rightarrow m_R$ and multiply the wave functions by Z_2 or Z_3 (see arguments in Sect. 8.1.8).

8.1.7

Ward–Takahashi Identity

The correction to the vertex function was calculated from Eq. (8.2)

$$\bar{u}(p')\gamma^\mu u(p) \rightarrow \bar{u}(p')[\gamma^\mu + \Gamma^\mu]u(p) \quad (8.72)$$

where Γ^μ is given in Eq. (8.20) and includes a divergent term. It can be separated and removed in a similar way as was done for the propagator. Namely, when we take the limit

$$q = p' - p \rightarrow 0, \quad p' \rightarrow m, \quad p \rightarrow m \quad (8.73)$$

then

$$\Gamma^\mu(p'^2, p^2) = \Gamma^\mu(m^2, m^2) + (\text{finite term}) \equiv (Z_1^{-1} - 1)\gamma^\mu + (\text{finite term}) \quad (8.74)$$

As a consequence, the vertex is modified by higher order corrections to

$$\gamma^\mu \rightarrow Z_1^{-1} \times \gamma^\mu (1 + \text{finite term}) \quad (8.75)$$

Actually, Z_1 can be obtained without explicitly calculating Eq. (8.2). The trick is to use the following relation:

$$\frac{\partial}{\partial p_\mu} S_F(p) = \frac{\partial}{\partial p_\mu} \frac{1}{\not{p} - m} = -\frac{1}{\not{p} - m} \gamma^\mu \frac{1}{\not{p} - m} \quad (8.76)$$

and

$$S_F(p) S_F^{-1}(p) = 1 \quad (8.77)$$

Differentiating Eq. (8.77) and using Eq. (8.76), we obtain

$$\frac{\partial}{\partial p_\mu} S_F^{-1}(p) = -S_F^{-1}(p) \left(\frac{\partial}{\partial p_\mu} S_F(p) \right) S_F^{-1} = \gamma^\mu \quad (8.78)$$

The meaning of Eqs. (8.76) and (8.78) is that the derivative of a propagator with respect to the momentum is equivalent to attaching a zero-energy photon or converting the propagator to a vertex with zero momentum transfer. The argument

can be generalized and any vertex function can be obtained by differentiating the corresponding propagator:

$$\Gamma^\mu(p, p) = \frac{\partial}{\partial p_\mu} S_F'^{-1}(p) - \gamma^\mu = \frac{\partial \Sigma_e(p)}{\partial p_\mu} \quad (8.79)$$

which is referred to as the Ward–Takahashi identity. It is easily seen that the relation holds between the lowest corrections Eq. (8.2) and Eq. (8.66b). Differentiating Eq. (8.68a) and comparing with Eq. (8.74), we immediately see

$$Z_1 = Z_2 \quad (8.80)$$

8.1.8

Renormalization of the Scattering Amplitude

To see where all the renormalized constants Z_1 , Z_2 , Z_3 and δm end up in the final scattering amplitude, we consider an example, Compton scattering (Fig. 8.5).

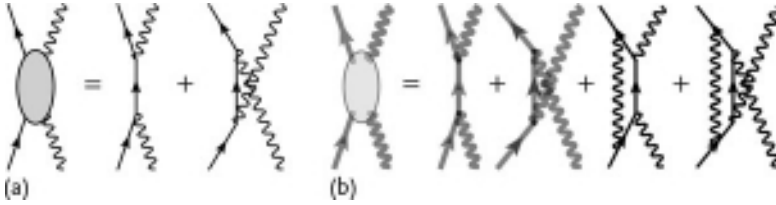


Figure 8.5 (a) Compton scattering in the lowest order. (b) Up to fourth order with dressing. See Fig. 8.6 for dressing.

The scattering amplitude corresponding to the first Feynman diagram on the rhs in Fig. 8.5a is expressed as

$$\mathcal{M} \sim \varepsilon'^*_{\mu} \bar{u}'(ie\gamma^\mu) \frac{i}{\not{p} + \not{k} - m} (ie\gamma^\nu) u \varepsilon_\nu \quad (8.81)$$

where we have suppressed the spin and polarization suffixes and have written $\bar{u}_s(p') \rightarrow \bar{u}'$, $u_r(p) \rightarrow u$, $\varepsilon_\mu(k', \lambda') \rightarrow \varepsilon'$. Taking into account higher order corrections, the internal as well as external lines of the electron propagator acquire Z_2 in addition to $m \rightarrow m - \delta m$, the photon propagator Z_3 , and the vertex function Z_1^{-1} . Namely, after including the corrections, the scattering amplitude is modified to

$$\begin{aligned} \mathcal{M} \sim & (Z_3 \varepsilon'^*) (Z_2 \bar{u}') (i Z_1^{-1} e \gamma^\mu) \left(\frac{i Z_2}{\not{p} + \not{k} - m + \delta m} \right) \\ & \times (i Z_1^{-1} e \gamma^\nu) (Z_2 u) (Z_3 \varepsilon_\nu) (1 + \dots) \end{aligned} \quad (8.82)$$

where $\dots$ are contributions from $\hat{\Pi}_e(p^2)$, $\hat{\Pi}_\gamma(q^2)$, finite terms in the vertex correction and others. We divide Z_2 , Z_3 into two as $Z_2^{1/2} \times Z_2^{1/2}$, $Z_3^{1/2} \times Z_3^{1/2}$ and move

them to combine them with the wave function and the coupling constant e . Then

$$\begin{aligned}\mathcal{M} &\sim (Z_3^{1/2} \varepsilon'_\mu)^* (Z_2^{1/2} \bar{u}') (i Z_1^{-1} Z_2 Z_3^{1/2} e \gamma^\mu) \frac{i}{\not{p} - m_R} \\ &\quad \times (i Z_1^{-1} Z_2 Z_3^{1/2} e \gamma^\nu) (Z_2^{1/2} u) (Z_3^{1/2} \varepsilon) (1 + \cdots) \\ &= \varepsilon_R'^* \bar{u}_R' (i e_R \gamma^\mu) \frac{i}{\not{p} - m_R} (i e_R \gamma^\nu) u_R \varepsilon_R (1 + \cdots)\end{aligned}\quad (8.83)$$

The modified expression does not contain diverging constants (Z_i 's, δm). In summary, by the replacements

$$e \rightarrow e_R = Z_1^{-1} Z_2 Z_3^{1/2} e = Z_3^{1/2} e \quad (8.84a)$$

$$m \rightarrow m_R = m - \delta m \quad (8.84b)$$

$$\psi \rightarrow \psi_R = Z_2^{1/2} \psi \quad (8.84c)$$

$$A_\mu \rightarrow A_{R\mu} = Z_3^{1/2} A_\mu \quad (8.84d)$$

all the diverging constants are absorbed in the redefinition of the coupling constant, mass, and wave function normalizations. Corrections to the second Feynman diagram in Fig. 8.5a can be treated similarly. We call the propagators and the vertex functions with corrections “dressed”. Then the two lowest order amplitudes written in terms of dressed particles are modified to the two diagrams written in thin bold lines in Fig. 8.5b. The dressed diagram contains diagrams as depicted in Fig. 8.6.

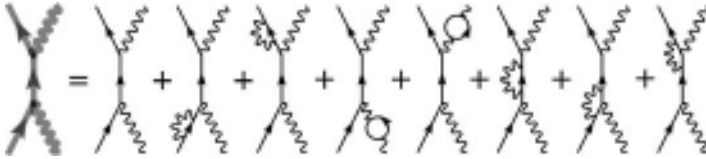


Figure 8.6 Irreducible diagrams such as the first term on the rhs can be dressed by taking the self-energy and vertex corrections. Reducible diagrams to the right of irreducible

diagrams can be included in the irreducible diagram simply by dressing the propagators and the vertices of the irreducible diagram.

There are other higher order effects that cannot be expressed by simply dressing the particle, for example the third and fourth diagrams in Fig. 8.5b. Their contribution is finite and is included in $(1 + \cdots)$. They are referred to as irreducible diagrams, which are defined as ones that cannot be separated by cutting a single line. They have to be calculated separately. Examples of reducible diagrams are shown on the rhs of Fig. 8.6. The reducible diagrams in Fig. 8.6 (diagrams on the rhs) can be included in the irreducible diagram (the first diagram) by simply dressing its propagators and vertices.

The prescription developed above seems rather artificial. All the infinities are contained in the mass m_R , the coupling constant e_R^2 and the wave function normalization. We simply swept the dust under the carpet. However, what we measure in experiments are masses and coupling constants with all the theoretical corrections included. Then m , α that appear in the Lagrangian should be regarded as mere parameters, and measured numbers, such as $m = 0.511 \text{ MeV}$, $\alpha = 1/137, \dots$ should be assigned to $m_R = m - \delta m$ and $e_R = e - \delta e$. As renormalized theory is written purely in terms of e_R and m_R , which are observed quantities, it does not include any infinities and is mathematically consistent in itself. QED, or quantum electrodynamics, thus established by Tomonaga, Schwinger and Feynman was the first successful gauge field theory that could give consistent and unambiguous predictions starting from first principles. Nevertheless, ad hoc treatments of infinities, such as renormalization, can only be justified if they agree with experiment.

8.2

Tests of QED

8.2.1

Lamb Shift

The first term in Eq. (8.59) of the Mott scattering amplitude with radiative correction represents the usual Coulomb force in the static limit. The second term is the QED correction and is referred to as the vacuum polarization effect. In quantized field theories, the vacuum is not an empty space but constantly creating and annihilating charged particle pairs. Under the influence of the applied external electromagnetic effect, those pairs naturally polarize. The vacuum behaves like a dielectric material and its effect appears as a part of the radiative correction. One phenomenon in which this correction appears is the energy shift of the hydrogen levels. They can be calculated by including the Coulomb potential in the Dirac equation, but without using quantized field theory the levels $2s_{1/2}$ and $2p_{1/2}$ are degenerate. Experimentally, there is a gap between the two energy levels (see Fig. 4.1), which is called the Lamb shift [248]. The energy levels are calculated by taking matrix elements of the effective Coulomb potential between the two hydrogen wave functions. The effective Coulomb potential can be obtained by considering higher order corrections to the Mott scattering amplitude. In addition to the vacuum polarization effect, they also include vertex correction, together with the infrared and the anomalous magnetic moment effects, as we saw in Sect. 8.1.2. Altogether, the net effect on the Mott amplitude is given by

$$\bar{u}(p') \left[\gamma^\mu \left\{ 1 - \frac{\alpha q^2}{3\pi m^2} \left(\ln \frac{m}{\mu} - \frac{3}{8} - \frac{1}{5} \right) \right\} + \frac{\alpha}{2\pi} \frac{i\sigma^{\mu\nu} q_\nu}{2m} \right] \frac{1}{|q|^2} u(p) \quad (8.85)$$

The first term in the braces $\{\dots\}$ represents the first-order Mott scattering amplitude. The first and second terms in the parentheses $(\dots)$ are the vertex corrections, $-1/5$ is the vacuum polarization and the last term in the square brackets $[\dots]$ is

the contribution due to the anomalous magnetic moment. The correction includes terms proportional to q^2 , which for the static field is ∇^2 acting on the space coordinate of the field. The operation of ∇^2 on the Coulomb potential produces $\delta^3(0)$. Namely, only the wave function at the origin contributes, hence it affects the s state but not the p state. Therefore, the correction removes the degeneracy between $s_{1/2}$ and $p_{1/2}$. In configuration space, the effective potential shift becomes [215]

$$\Delta V_{\text{eff}} = \frac{4\alpha^2}{3m^2} \left(\ln \frac{m}{\mu} - \frac{3}{8} - \frac{1}{5} + \frac{3}{8} \right) \delta^3(\mathbf{x}) + \frac{\alpha^2}{4\pi m^2} \frac{\boldsymbol{\sigma} \cdot \mathbf{L}}{r^3} \quad (8.86)$$

(1) (2) (3) (4) (4')

The vertex correction terms [(1) and (2)] are larger than the vacuum polarization effect (3) in the Lamb shift. The infrared cut is determined by extrapolating its contribution and connecting it smoothly with that obtained by nonrelativistic treatment, which is finite. The contribution from the anomalous magnetic moment [terms (4) and (4')] can be calculated by extracting the lowest order matrix element of the effective Hamiltonian

$$\mathcal{H} = -\kappa \frac{e}{2m} \bar{\psi}(x) \left[\frac{1}{2} \sigma^{0i} F_{0i} \right] \psi(x), \quad \kappa = \frac{\alpha}{2\pi} \quad (8.87)$$

If we express the energy shift by its corresponding frequency ($\Delta E = h\Delta\nu$), the vertex correction amounts to 1010 MHz, the vacuum correction to -27.1 MHz, and the anomalous magnetic moment to 68 MHz giving a total of 1051 MHz which is accurate to about 6 MHz. Since then, higher order corrections have been calculated, so the accuracy is now ~ 0.01 MHz.

Theory total	1057.8445 ± 0.0026	MHz [279]
Experiment	1057.8450 ± 0.0090	MHz

As one can see, the measurement is precise enough to test the correction. It was QED's first major success to convince people that it is the right theory for the electromagnetic interactions of the electron.

8.2.2

$g - 2$

In the radiative correction to the vertex, a term proportional to $q_\nu \sigma^{\mu\nu}$ appears. This is just $\boldsymbol{\sigma} \cdot \nabla \times \mathbf{A} \sim \boldsymbol{\mu} \cdot \mathbf{B}$ and is a higher order correction to the magnetic moment that alters the value of $g - 2$ from zero. The statement that the Dirac particle has $g = 2$ is no longer correct when higher order QED effects are included. We have calculated the first loop correction to the vertex and obtained a correction that amounts to $\alpha/2\pi$ [see Eq. (8.22)]. Calculations of even higher orders have since

been performed [219, 232] and have been compared with the experimental values:

$$\begin{aligned} \left(\frac{g-2}{2}\right)_e &= \frac{\alpha}{2\pi} - 0.328\,478\,444\,002\,90\,(60) \left(\frac{\alpha}{\pi}\right)^2 \\ &+ 1.181\,234\,016\,828\,(19) \left(\frac{\alpha}{\pi}\right)^3 \\ &- 1.914\,4(35) \left(\frac{\alpha}{\pi}\right)^4 + 0.0(3.8) \left(\frac{\alpha}{\pi}\right)^5 + \dots \end{aligned} \quad (8.88a)$$

Inserting the value of the fine structure constant, the theoretical value can be compared with the experimental data:

$$(1\,159\,652\,180.85 \pm 0.76) \times 10^{-12} \quad \text{experiment [219, 311]} \quad (8.88b)$$

The experimental value nowadays is so accurate that it perfectly agrees with the theoretical value. In fact, the theoretical error is larger due to uncertainty in the fine structure constant. Therefore, the experimental value is used to determine the most accurate value of the fine structure constant [219]:

$$\alpha^{-1} = 137.035\,999\,069\,(96) \quad [\text{relative error} = 0.70 \text{ ppb}^7] \quad (8.89)$$

The experimental value of $g-2$ is no longer a value to be tested but a standard.

Similar calculations have also been made for the muon [311]

$$\begin{aligned} \left(\frac{g-2}{2}\right)_\mu &= \frac{\alpha}{2\pi} + 0.765\,857\,410 \left(\frac{\alpha}{\pi}\right)^2 + 24.050\,509 \left(\frac{\alpha}{\pi}\right)^3 \\ &+ 130.805 \left(\frac{\alpha}{\pi}\right)^4 \dots \end{aligned} \quad (8.90a)$$

The formula gives

$$\begin{aligned} &= (116\,584\,718.10 \pm 0.16) \times 10^{-11} \quad \text{QED only} \\ &\quad (+154 \pm 2) \times 10^{-11} \quad \text{electroweak} \\ &\quad (+6\,916 \pm 46) \times 10^{-11} \quad \text{hadronic} \end{aligned} \quad (8.90b)$$

$$\begin{aligned} &= (116\,591\,788 \pm 46) \times 10^{-11} \quad \text{theory total} \\ &= (116\,592\,080 \pm 54) \times 10^{-11} \quad \text{experiment} \end{aligned} \quad (8.90c)$$

The different value of the muon $g-2$ reflects the difference of the mass. The error in the muon is much larger than that of the electron. This is because for the muon the hadronic as well as weak interaction correction is sizeable (Fig. 8.7b–e). The hadronic correction enters in the process in which a virtual photon becomes hadrons (i.e. quark pairs). The hadronic correction uses the high-energy $e^-e^+ \rightarrow$ hadrons data.

Although the procedure of subtracting infinity from infinity is aesthetically unsatisfactory, a theory that can reproduce the experimental result to the accuracy of

7) ppb: parts per billion = 10^{-9} .

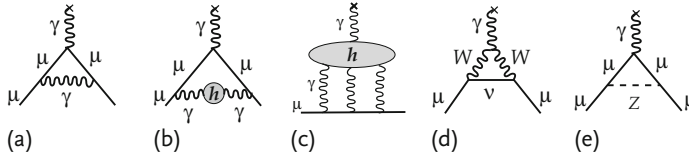


Figure 8.7 Feynman diagrams for $g - 2$ of the muon. (a) Schwinger term [$O(\alpha)$ correction], (b) hadronic contribution, (c) hadronic light-by-light contribution, (d,e) lowest order $W^\pm$, Z contribution.

10^{-10} must have some truth in it. That we trust in the renormalization procedure has firm foundations experimentally. In fact, the confidence in the gauge theory starting from QED is so solid that any deviation of the experimental measurement of $g - 2$ is interpreted as an effect of unknown new physics. In fact, a small discrepancy from the theoretical prediction was detected in measurements of the $g - 2$ of the muon [281]. It could be an experimental error, but if real it could signal a possible new physics, for instance, supersymmetry (see Chap. 19) and is one of the hot topics of the day (as of 2010). Nowadays, the precision measurement of a physical object to look for a possible deviation from the theoretical prediction is considered a valid and powerful method to search for new physics.

8.2.3

Limit of QED Applicability

QED is a quantized version of the well-established classical Maxwell theory. At the macroscopic level, it can be applied to the magnetic field in the periphery of Jupiter and to the galactic magnetic field. Namely, it is valid at the scale of at least a few hundred light years. What about at the microscopic level? The fact that there was a limit to the applicability of classical electromagnetism at the turn of the 20th century led to the introduction of quantum theory. We can test the validity of QED at small distances or at high energy by performing $e^- + e^+ \rightarrow \mu^- + \mu^+$ and Bhabha scattering ($e^- + e^+$ scattering), etc. If we parametrize the difference of the electron propagator between the theory and the experiment by $F = 1 + s/\Lambda^2$, where s is the total energy squared, we can estimate quantitatively the applicability limit of QED by the value of Λ . At the present experimental level, we see no deviation from QED. The present experimental lower limit of Λ is larger than 1 TeV. In other words, we know that QED is valid down to the length scale of 10^{-20} m. Therefore, QED is valid in scales over the range of 40 orders of magnitude and is still extending its validity. QED's triumph is remarkable.

8.2.4

E821 BNL Experiment

Although the measured value $g - 2$ of the electron is much more precise and is used to define the prime fine structure constant α , the muon is much more sensitive to

effects other than QED, including the electroweak effect (i.e. the contribution of Feynman diagrams that contain $W^\pm$, Z^0) and hadronic interactions and possibly new physics. The contribution of the physics of the energy scale M to $g - 2$ is proportional to $(m/M)^2$ and hence the effect is $\sim (m_\mu/m_e)^2 \simeq 4.3 \times 10^4$ larger for the muon. Even though $a_\mu \equiv (g - 2)/2$ is ~ 800 times less precise than a_e , it is still ~ 50 times more sensitive to new physics.

Since hadronic as well as weak effects on $g - 2$ can be calculated accurately using the now established Standard Model (SM), the deviation, if any, of a measurement from the prediction of SM could be a possible clue to new physics. For this reason, the precise measurement of $g - 2$ of the muon not only serves as a crucial test of QED but is also at the forefront of experimental research in high-energy physics. We describe briefly how to measure the value of $g - 2$ experimentally.

Principle of the Measurement A charged elementary particle with half-integer spin has a magnetic dipole moment $\boldsymbol{\mu}$ with spin s :

$$\boldsymbol{\mu} = g_s \left(\frac{q}{2m} \right) \mathbf{s} \quad (8.91)$$

where $q = \pm e$ is the charge of the particle and the proportionality constant g_s is referred to as the Landé g -factor. In a homogeneous magnetic field where the magnetic energy is given by $\mathcal{H} = -\boldsymbol{\mu} \cdot \mathbf{B}$, the spin receives a torque from the field, which is given by $\mathbf{N} = \boldsymbol{\mu} \times \mathbf{B}$, and the spin makes a precessive motion on the axis parallel to the magnetic field. In the particle's rest frame the precession (Larmor) frequency is given by

$$\omega_L = g_s \frac{qB}{2m} \quad (8.92)$$

but when the particle is moving, a relativistic effect called Thomas precession is added and the net frequency is given by

$$\omega_s = \omega_L + \omega_T = g_s \frac{qB}{2m} + \frac{qB}{\gamma m} (1 - \gamma) \quad (8.93)$$

The particle makes a circular orbit perpendicular to the magnetic field. Assuming $\mathbf{v} \cdot \mathbf{B} = 0$, its angular velocity is given by the cyclotron frequency ω_c and the spin precession frequency relative to the orbit becomes

$$\begin{aligned} \omega_a = \omega_s - \omega_c &= \underbrace{g_s \frac{qB}{2m}}_{\text{Larmor}} + \underbrace{\frac{qB}{m\gamma} (1 - \gamma)}_{\text{Thomas}} - \underbrace{\frac{qB}{m\gamma}}_{\text{cyclotron}} \\ &= \left(\frac{g_s - 2}{2} \right) \frac{qB}{m} \equiv a_\mu \frac{qB}{m} \end{aligned} \quad (8.94)$$

Since, m (the muon mass) is known and the strength of the magnetic field can be fixed by experimental conditions, what we need to determine the anomaly a_μ is to measure the spin precession frequency. But at the precision level of parts per billion (ppb), B and m have to be measured with full care, too.

Thomas precession may be understood as follows. Imagine an airplane flying in a circular orbit. Its motion may be considered as a series of longitudinal advances by $d\ell_{\parallel} \simeq r\Delta\theta$ accompanied by a transverse shift $d\ell_{\perp} = \Delta\theta d\ell_{\parallel}$. Inside the airplane (variables denoted by a prime), the longitudinal distance is Lorentz contracted ($d\ell'_{\parallel} = d\ell_{\parallel}/\gamma$), but the transverse distance is not ($d\ell'_{\perp} = d\ell_{\perp}$), giving the angular shift ($\Delta\theta' = \gamma\Delta\theta$). When the airplane comes back to the original point, the rotation angle in the airplane frame (comoving frame) has an excess $2\pi(\gamma - 1)$ relative to the ground frame. Therefore, the rate of extra advance rotation per period (angular frequency) is given by

$$\frac{\omega_{\text{com}}}{\omega_{\text{circ}}} = \frac{2\pi(\gamma - 1)/T}{2\pi/T} = \gamma - 1 \quad (8.95)$$

This is a shift of a reference axis with which the rotation angle of a spinning object is measured. The advance shift of the reference means the spin precession frequency of a muon moving in a circular orbit with the velocity $v = 2\pi r\omega$ in the laboratory frame is smaller by $\gamma - 1$. Since the circular motion of the muon is due to the Lorentz force of the magnetic field and its angular frequency (cyclotron frequency ω_c) itself becomes smaller by γ because of the relativistic mass increase, the net change in the angular frequency due to the Thomas precession becomes $\Delta\omega_{\text{Thomas}} = -(qB/m)(\gamma - 1)/\gamma$.

Note: The Thomas precession is a purely kinematical effect independent of the dynamics of the spinning object and is common to objects in quantum mechanics and classical tops or gyroscopes.

Experimental Method

(i) Measurement of ω_a : Firstly, we have to prepare a large enough number of polarized muons and a method to measure their direction. Polarized muons can be obtained by collecting decay muons emitted in the forward direction of flying pions. The weak interaction mediated by the gauge boson $W^{\pm}$ acts only on the left-handed particles. Because of this, the decay neutrino from π^+ in the reaction

$$\pi^+ \rightarrow \mu^+ + \nu_{\mu} \quad (8.96a)$$

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_{\mu} \quad (8.96b)$$

has helicity -1 and by angular momentum conservation μ^+ has also helicity -1 , namely is 100% polarized in the pion rest frame. Consequently, muons emitted in the forward direction in the laboratory frame along the high-energy pion beam are also highly polarized. The left-handed nature of the weak interaction also works in muon decay and $e^{\pm}$ are preferentially emitted antiparallel (parallel) to the polarization of $\mu^{\pm}$. Therefore the electron intensity at a fixed direction relative to the muon momentum is modulated by the muon spin precession (see Fig. 8.8). Name-

ly, the muon has a built-in polarization analyzer. The modulation frequency directly reflects ω_a in Eq. (8.94).

(ii) Measurement of B : The magnetic field B is measured by observing the Larmor frequency of stationary protons, $\omega_p = g_p(eB/2m_p)$, in a nuclear magnetic resonance (NMR) probe. This requires the value of μ_μ/μ_p to extract a_μ from the data. Fortunately, it can be determined precisely from the measurement of the hyperfine structure in muonium (μ - e bound state). From Eq. (8.94) one obtains

$$a_\mu = \frac{\mathcal{R}}{\mu_\mu/\mu_p - \mathcal{R}}, \quad \mathcal{R} = \frac{\omega_a}{\omega_p} \quad (8.97)$$

(iii) Focusing of the muon beam: To store muons in a circular orbit and retain them for the duration of the measurement, focusing of the beam is necessary. This can be achieved by making gradients on the magnetic field, but then the homogeneity of the field is sacrificed, which is the most critical factor in the experiment to obtain high precision at ppb level. Therefore, E821 (and also previous CERN experiments [36]) used an electrostatic quadrupole field focusing. With the presence of the electric field, the precession frequency is modified to [216, 276]

$$\omega_a = \frac{q}{m} \left[a_\mu \mathbf{B} - \left(a_\mu - \frac{1}{\gamma^2 - 1} \right) \mathbf{v} \times \mathbf{E} \right] \quad (8.98)$$

For $\gamma_{\text{magic}} = 29.3$ ($p_\mu = 3.09 \text{ GeV}/c$), the second term vanishes and a_μ reduces to Eq. (8.94) again.

Figure 8.9a shows schematically the muon injection and storage in the $g - 2$ ring. The trajectory of the injected beam is modified by three kicker modules to become a circle that stays in the ring. The spin precesses $\sim 12^\circ/\text{circle}$ (Fig. 8.9b). The decay electrons are measured by 24 lead-scintillator calorimeters distributed along the inner circle of the ring. Figure 8.8 shows the measured distribution of electron counts. The result gives

$$a_\mu^{\text{exp}} = 1.165\,920\,80\,(63) \times 10^{-3} \quad (8.99)$$

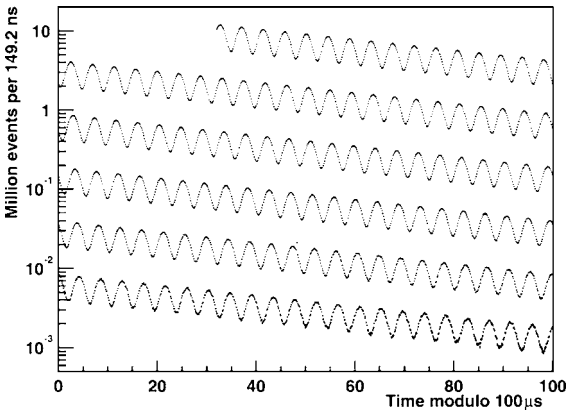


Figure 8.8 Distribution of electron counts versus time for 3.6×10^9 muon decays in the R01 μ^- data-taking period. The data are wrapped around modulo $100 \mu\text{s}$. Data from [281].

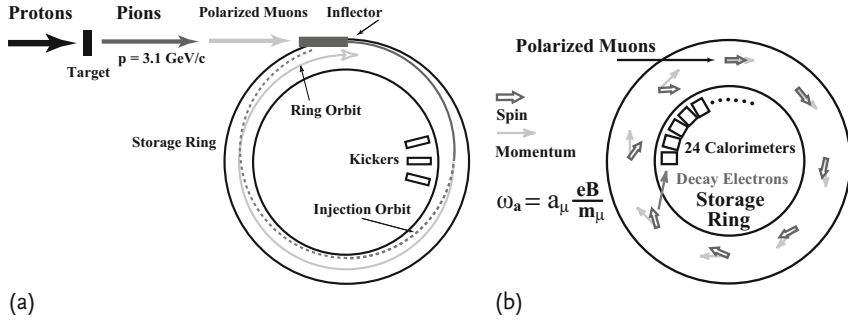


Figure 8.9 (a) Schematic diagram of the E821 muon storage ring [281]. Polarized muons with $\sim 3.1 \text{ GeV}/c$ enter the storage ring through a field-free channel. The three kicker modulators provide a short current pulse, which gives the muon bunch a transverse 10 mrad kick and places it in the ring orbit.

(b) Schematic diagram of the muon spin precession and the electron detector layout. Note the spin precession rate is exaggerated. In reality it is $\sim 12^\circ/\text{turn}$. Electrons are preferentially emitted parallel or antiparallel to the muon spin direction. Figure from [219].

where the number in parenthesis denotes error in the last two digits. The experimental value is compared with the theoretical prediction

$$a_\mu^{\text{theor}} = 1.165\,917\,93\,(68) \times 10^{-3} \quad (8.100)$$

which is 3.2σ ($\sigma = \text{standard deviation}$) off. Whether the deviation is a possible sign of new physics is under much debate.

9

Symmetries

When one does not know the principles or the fundamental equations that govern the phenomena one observes, one often obtains clues to them by looking at the symmetries or selection rules they exhibit. Historically, Maxwell constructed his unified theory of electromagnetism by assuming the existence of the displacement current, which was required by current conservation. Discovery of special relativity was inspired by Lorentz invariance of Maxwell's equations. History also shows that disclosed symmetries, more often than not, turn out to be approximate. Isospin and flavor $SU(3)$, which led to the discovery of the quark model, are such examples. Today what people believe to be strictly conserved are based on gauge symmetries. Energy-momentum, angular momentum and charge conservations are in this category. The conservation of baryon and lepton numbers, although phenomenologically valid, are not considered as strict laws. Their mixing is an essential feature of the grand unified theories (GUTs). Symmetries and conservation laws are intimately connected. Whether based on gauge symmetry or not, apparent symmetries and conservation laws can be powerful tools for discovering new physics.

Symmetry means one cannot tell the difference before and after a certain transformation. For the case of continuous space-time translation, this means there is no special position or absolute coordinate system in setting space-time coordinates for describing a physical phenomenon. The same experiments done at two different places (say in Tokyo and in New York) or done at different times (today or tomorrow) should produce the same result. In other words, translational invariance means the existence of certain nonobservables, i.e. absolute position and time. So far, we have elaborated on the relation between space-time translational symmetry and the resultant energy-momentum conservation. However, it is a common feature of symmetry. Similar relations can be derived for a variety of symmetry transformations. We list some of them here in Table 9.1 before proceeding further.

Symmetries come in two categories, first, those related to space-time structure, which are often referred to as external, and second, those related to internal symmetries, such as electric charge, isospin and color charge. Mathematically speaking, the latter leave the Lagrangian density invariant after the symmetry operation apart from a total derivative. It gives a surface integral that vanishes under normal circumstances. Space-time symmetries change the coordinate values, so only the

Table 9.1 Example symmetries and conservation laws.

Nonobservables	Symmetry transformation	Conserved observables or selection rules
Space-time symmetry		
Absolute position	$\mathbf{x} \rightarrow \mathbf{x} + \mathbf{a}$	Momentum
Absolute time	$t \rightarrow t + b$	Energy
Absolute direction	$\theta \rightarrow \theta + \alpha$	Angular momentum
Absolute velocity	Lorentz transformation	Lorentz invariance
Right or left	$\mathbf{x} \rightarrow -\mathbf{x}$	Parity
Internal symmetry		
Particle identity	Permutation	Bose–Einstein or Fermi–Dirac statistics
Charge + or –	$Q \rightarrow -Q$	C invariance
Absolute phase	$\varphi \rightarrow e^{-iQ\alpha} \varphi$	Charge
– among quarks	$SU(3)$ gauge transformation	Color charge
Approximate symmetry		
– among leptons ^a	$\psi \rightarrow e^{-iL\alpha} \psi$	Lepton number
– among quarks	$\psi \rightarrow e^{-iB\alpha} \psi$	Baryon number
Near equal mass of u, d	Flavor $SU(2)$	Isospin
– of u, d, s	Flavor $SU(3)$	Unitary spin

^a Phenomenologically, lepton and baryon numbers are conserved.

action, which is an integral of a Lagrangian over all space-time, remains invariant. In other classifications, they are either continuous or discrete. We discuss the continuous space-time symmetries first.

9.1

Continuous Symmetries

Where there is a symmetry, there exists a corresponding conservation law. What this means is that when an equation of motion or the Lagrangian is invariant under some symmetry operation, there exists a physical observable that does not change as a function of time. We have already shown some examples in Sect. 5.1.4 in the form of Noether’s theorem. Here, we start by comparing how the symmetry and the invariance of the equation of motion are related in classical, quantum mechanical and quantum field theoretical treatments.

9.1.1

Space and Time Translation

Classical Mechanics In the Lagrangian formalism, the equation of motion is given as the Euler–Lagrange formula

$$\frac{d}{dt} \left(\frac{\partial L(q_i, \dot{q}_i)}{\partial \dot{q}_i} \right) - \frac{\partial L(q_i, \dot{q}_i)}{\partial q_i} = 0 \quad (9.1)$$

If the Lagrangian is translationally invariant, namely,

$$L(q_i) = L(q_i + a_i) \quad (9.2)$$

or equivalently $\partial L / \partial q_i = 0$, the Euler–Lagrange formula Eq. (9.1) tells us that the conjugate momentum defined by

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad (9.3)$$

is a constant of time. Therefore, translational invariance in space coordinates leads to the existence of the conjugate momentum, which is time invariant. Since the Lagrangian formalism is valid in a general coordinate system, this statement is general. If q_i is an angular variable, the conservation of angular momentum follows. The time derivative of the Hamiltonian is expressed as

$$\begin{aligned} \frac{dH}{dt} &= \frac{d}{dt} \left(\dot{q} \frac{\partial L}{\partial \dot{q}} - L \right) = \ddot{q} \frac{\partial L}{\partial \dot{q}} + \dot{q} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \left(\frac{\partial L}{\partial t} + \frac{\partial L}{\partial q} \dot{q} + \frac{\partial L}{\partial \dot{q}} \ddot{q} \right) \\ &= -\frac{\partial L}{\partial t} \end{aligned} \quad (9.4)$$

where the Euler–Lagrange equation was used to arrive at the last equation. The above expression tells us that the Hamiltonian is a constant of motion if the Lagrangian does not contain time variables explicitly. Namely, translational invariance in time leads to energy conservation.

Quantum Mechanics Observables in quantum mechanics are expressed as matrix elements of hermitian operators. If there is a unitary operator $\hat{U}^{(1)}$ that does not have explicit time dependence, and if the transformed wave function $\psi' = \hat{U}\psi$ obeys the same Schrödinger equation, then

$$\begin{aligned} i \frac{\partial}{\partial t} \psi &= \hat{H} \psi \rightarrow i \frac{\partial}{\partial t} \hat{U} \psi = \hat{U} \hat{H} \hat{U}^{-1} \hat{U} \psi \\ \therefore i \frac{\partial}{\partial t} \psi' &= \hat{H} \psi' \rightarrow \hat{U} \hat{H} \hat{U}^{-1} = \hat{H} \quad \text{or} \quad [\hat{H}, \hat{U}] = 0 \end{aligned} \quad (9.5)$$

- 1) The symmetry operation has to be unitary because the transition probability of the state before and after symmetry transformation has to be the same. $(\langle \psi' | \psi' \rangle = \langle \psi | \hat{U}^\dagger \hat{U} | \psi \rangle = \langle \psi | \psi \rangle)$. We attach a hat $\hat{}$ to differentiate operators from numbers when we discuss objects in quantum mechanics.

The operator $\hat{U}$ commutes with the Hamiltonian. If $\hat{U}$ is a continuous function of some parameter ε and $\hat{U} \rightarrow 1$ when $\varepsilon \rightarrow 0$, then for infinitesimal ε

$$\hat{U} \simeq 1 - i\varepsilon \hat{Q} \quad (9.6)$$

As $\hat{U}$ is unitary, $\hat{Q}$ is hermitian. The operator $\hat{Q}$ in $\hat{U} = e^{-i\hat{Q}\alpha}$, where α is a continuous parameter, is called the generator of the transformation. Since, the hermitian operator in quantum mechanics is an observable, $\hat{Q}$ is an observable. As $\hat{Q}$ commutes with the Hamiltonian, the time derivative of the expectation value is given by

$$\frac{d\langle \hat{Q} \rangle}{dt} = \frac{d}{dt} \langle \psi | \hat{Q} | \psi \rangle = i \langle \psi | [\hat{H}, \hat{Q}] | \psi \rangle = 0 \quad (9.7)$$

Therefore, if there is a continuous symmetry represented by a unitary operator $\hat{U}$ that leaves the equation of motion invariant, its generator $\hat{Q}$ is a constant of time.

The hamiltonian is a time translation operator. This is obvious from

$$\psi(t) = e^{-i\hat{H}t} \psi(0) \rightarrow \psi(t+a) = e^{-ia\hat{H}} \psi(t) \quad (9.8)$$

To find the corresponding space translation operator $\hat{U}(a) = e^{-ia\hat{Q}}$ that is defined by

$$\psi(x) \rightarrow \psi(x+a) = \hat{U}(a)\psi(x) \quad (9.9a)$$

we make the displacement infinitesimal:

$$\begin{aligned} \hat{U}(\epsilon)\psi(x) &= [1 - i\epsilon\hat{Q} + O(\epsilon^2)]\psi(x) = \psi(x+\epsilon) \\ &= \psi(x) + \epsilon \frac{\partial}{\partial x} \psi(x) + O(\epsilon^2) \end{aligned} \quad (9.9b)$$

$$\therefore \hat{Q} = i \frac{\partial}{\partial x} = -\hat{p}$$

A Lorentz invariant expression can be obtained by combining Eqs. (9.8) and (9.9):

$$\psi(x^\mu + a^\mu) = e^{-ia^\mu P_\mu} \psi(x^\mu) \quad (9.10a)$$

$$P_\mu = i(\partial_0, \nabla) \quad (9.10b)$$

Problem 9.1

For a translational operator $\hat{p}$ where $O(x)$ is a function of $\hat{p}$ and x :

$$O(x+a) = e^{ia\hat{p}} O(x) e^{-ia\hat{p}} \quad (9.11a)$$

where $O(x)$ is a function of $\hat{p}$ and x , prove

$$\nabla O(x) = i[\hat{p}, O(x)] \quad (9.11b)$$

Translational Invariance in Quantum Field Theory

Covariant version of the Heisenberg equation of motion is expressed as

$$\partial_\mu O = i[P_\mu, O] \quad (9.12a)$$

$$O(x + a) = e^{i a^\mu P_\mu} O(x) e^{-i a^\mu P_\mu} \quad (9.12b)$$

The field theoretical version of the energy momentum operator is given by

$$P^\mu = \int d^3x T^{\mu 0} \quad (9.13)$$

where $T^{\mu\nu}$ is the energy-momentum tensor that can be derived from the Lagrangian using Noether's theorem. Here, P^μ and $O(x)$ are functions of the field operators. Note, corresponding to the negative sign of space variables in the Lorentz-invariant expression ($a^\mu P_\mu = a^0 P^0 - \mathbf{a} \cdot \mathbf{P}$), $P_\mu = (P^0, -\mathbf{P})$ is used.

We first prove that Eq. (9.12b) is equivalent to the Heisenberg equation (9.12a). To prove the equivalence, it is enough to show that one can be derived from the other and vice versa. Consider a^μ in Eq. (9.12b) as a small number ε^μ , expand both sides of the equation and compare terms of $O(\varepsilon)$:

$$(1 + i \varepsilon^\mu P_\mu) O(x) (1 - i \varepsilon^\mu P_\mu) \simeq O(x) + \varepsilon^\mu \partial_\mu O \quad (9.14a)$$

$$\therefore i[P_\mu, O(x)] = \partial_\mu O(x) \quad (9.14b)$$

Conversely, to derive Eq. (9.12b) from Eq. (9.12a), we use the Baker–Campbell–Hausdorff (BCH) formula.

Baker–Campbell–Hausdorff formulae Let A, B be operators.

Theorem 9.1

$$\begin{aligned} e^A B e^{-A} &= B + [A, B] + \frac{1}{2!} [A, [A, B]] + \cdots \\ &\quad + \frac{1}{n!} \overbrace{[A, [A, [A, \cdots [A, B]]]]}^{nA's} + \cdots \end{aligned} \quad (9.15)$$

Corollary 9.1

$$[B, e^{-A}] = e^{-A} \left([A, B] + \frac{1}{2} [A, [A, B]] + \cdots \right) \quad (9.16a)$$

$$[e^A, B] = \left([A, B] + \frac{1}{2} [A, [A, B]] + \cdots \right) e^A \quad (9.16b)$$

Theorem 9.2

Assuming $[A, [A, B]] = [B, [B, A]] = 0$

$$e^{A+B} = e^{-\frac{1}{2}[A, B]} e^A e^B = e^A e^B e^{-\frac{1}{2}[A, B]} \quad (9.17a)$$

$$e^A e^B = e^{[A, B]} e^B e^A \quad (9.17b)$$

Corollary 9.2

$$e^{A/2} e^B e^{A/2} = e^{A+B} \quad (9.18)$$

Problem 9.2

Prove Theorem 9.1 and Theorem 9.2

Replace A by $i P_\mu a^\mu$, B by O , and make use of Eq. (9.12a), then

$$\begin{aligned} e^{i P_\mu a^\mu} O(x) e^{-i P_\mu a^\mu} &= O(x) + a^\mu \partial_\mu O(x) + \cdots + \frac{a^n}{n!} \partial_\mu^{(n)} O(x) + \cdots \\ &= O(x + a) \end{aligned} \quad (9.19)$$

It remains to prove Eqs. (9.12) with P_μ expressed as Eq. (9.13). We consider the case of a complex scalar field. Generalization to other fields is straightforward. The energy-momentum operator is given in Eq. (5.93).

$$H = \int d^3x \left[\left(\frac{\partial \varphi^\dagger}{\partial t} \right) \left(\frac{\partial \varphi}{\partial t} \right) + (\nabla \varphi^\dagger \cdot \nabla \varphi) + m^2 \varphi^\dagger \varphi \right] \quad (9.20a)$$

$$\mathbf{P} = - \int d^3x \left[\frac{\partial \varphi^\dagger}{\partial t} \nabla \varphi + \nabla \varphi^\dagger \frac{\partial \varphi}{\partial t} \right] \quad (9.20b)$$

Using the equal time commutation relation

$$[\varphi(x), \pi(y)]_{t_x=t_y} = [\varphi(x), \dot{\varphi}^\dagger(y)]_{t_x=t_y} = i \delta^3(x - y) \quad (9.21)$$

it is easy to verify Eq. (9.12a) if $O = \varphi(x)$ or $O = \pi(x)$.

Then for any operators A, B that satisfy

$$\partial_\mu A = i[P_\mu, A], \quad \partial_\mu B = i[P_\mu, B] \quad (9.22a)$$

$$i[P_\mu, AB] = i[P_\mu, A]B + Ai[P_\mu, B] = (\partial_\mu A)B + A\partial_\mu B = \partial_\mu(AB) \quad (9.22b)$$

it follows that the equation is valid if (AB) is replaced with any polynomials of A and B . Since P_μ is a conserved operator that does not include any space-time

coordinates, applying any function of space-time differentiation $g(\partial)$ on both sides of the equation gives

$$i[P_\mu, \{g(\partial) O(A, B)\}] = \partial_\mu \{g(\partial) O(A, B)\} \quad (9.23)$$

Therefore, if O is a local operator, i.e. a polynomial of $\varphi(x)$, $\pi(x)$ and their derivatives, it satisfies

$$\partial_\mu O = i[P_\mu, O] \quad (9.24)$$

q.e.d.

In summary, the notion that for the space-time translational symmetry there exists a corresponding conserved quantity, the energy-momentum, is valid in quantum field theory as it was in classical field theory. In quantum field theory, the energy-momentum operators are the generators of the translational transformation. Note also that the Heisenberg equation holds even if the fields are interacting, because the equal time commutation relation, which was the basis of the proof, is also valid for interacting fields. The validity is ensured by the fact that they are connected by a unitary transformation [see Eq. (6.7)].

9.1.2

Rotational Invariance in the Two-Body System

Discussions in the previous section have shown that energy-momentum conservation is a result of translational invariance in Cartesian coordinates. As the Lagrangian formalism is valid for a general coordinate system, translational invariance in polar coordinates $\phi \rightarrow \phi + \alpha$ results in another conserved quantity, the angular momentum. Expressed in the original Cartesian coordinates

$$\mathbf{p} \times \mathbf{r} = -i\partial_\phi = -i(x\partial_y - y\partial_x) = \mathbf{r} \times \frac{\nabla}{i} = \mathbf{L}_z \quad (9.25)$$

In Chap. 3 we have shown that Lorentz invariance leads to the existence of spin angular momentum, which acts on spin components of the field and is an intrinsic property of particles, the quanta of the field. Determining the spin is an essential step in the study of elementary particle physics and the method is based on the consideration of rotational invariance. Experimentally, the most frequently used reactions are two-body scatterings. If rotational invariance holds, angular momentum is conserved. Then the scattering amplitude can be decomposed into partial waves of definite angular momentum, and by analyzing the angular distributions of the two-body scattering state the spin of the system can be determined. We have already learned in quantum mechanics that the scattering amplitude in the center of mass (CM) frame can be expanded in partial waves:

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 \quad (9.26a)$$

$$f(\theta) = \frac{1}{2ip} \sum_J (2J+1) [S^J(E) - 1] P_J(\cos\theta) \quad (9.26b)$$

where $\mathbf{p}$ is the momentum in CM, $S^J(E)$ is the partial element of the scattering matrix with angular momentum J and $P_J(\theta)$ is the Legendre function of $\cos \theta$. The above formula is valid for a spinless particle. It can be extended to particles having spin [217]. For the treatment, it is convenient to work in the eigenstates where the incoming and outgoing states are specified by their helicity

$$h = \frac{\mathbf{J} \cdot \mathbf{p}}{|\mathbf{p}|} \quad (9.27)$$

instead of J_z , which is a component of the angular momentum along the fixed z -axis. The helicity is by its definition rotationally invariant.

What we are dealing with is a two-particle system having momentum and helicity $\mathbf{p}_1, \lambda_1$ and $\mathbf{p}_2, \lambda_2$, respectively:

$$|\Psi\rangle = |\mathbf{p}_1 \lambda_1; \mathbf{p}_2 \lambda_2\rangle = \psi_1(\mathbf{p}_1, \lambda_1) \psi_2(\mathbf{p}_2, \lambda_2) \quad (9.28)$$

Let us first separate the motion of the two-body reaction

$$A(p_1) + B(p_2) \rightarrow C(p_3) + D(p_4) \quad (9.29)$$

into that of the CM frame itself and the relative motion in it. Defining the momentum of the CM system $\mathbf{P}$ and the relative momentum $\mathbf{p}$ in the initial and final states by

$$\mathbf{P}_i = \mathbf{p}_1 + \mathbf{p}_2, \quad \mathbf{P}_f = \mathbf{p}_3 + \mathbf{p}_4 \quad (9.30a)$$

$$\mathbf{p}_i = \frac{1}{2}(\mathbf{p}_1 - \mathbf{p}_2), \quad \mathbf{p}_f = \frac{1}{2}(\mathbf{p}_3 - \mathbf{p}_4) \quad (9.30b)$$

and working in the CM frame, in which $\mathbf{P}_i = \mathbf{P}_f = (E, \mathbf{0})$, state vectors in CM can be expressed by their relative momentum in polar coordinates. We rewrite the two-particle state as

$$|p \theta \phi; \lambda_1 \lambda_2\rangle = \psi_1(p \lambda_1) \psi_2(p \lambda_2) \quad (9.31)$$

where particle 1 has momentum $\mathbf{p} = (p, \theta, \phi)$ in polar coordinates and helicity λ_1 and particle 2 has $-\mathbf{p}$ and helicity λ_2 . Sometimes we will use the total helicity

$$\lambda = \lambda_1 - \lambda_2, \quad \mu = \lambda_3 - \lambda_4 \quad (9.32)$$

to denote the two-particle helicity state, i.e. $|\lambda\rangle = |\lambda_1 \lambda_2\rangle$, when there is no danger of confusion. Expanding the plane-wave state in terms of those that have angular momentum J , M is the center of discussion.

If rotational invariance holds, angular momentum is conserved and the scattering matrix elements can be expressed as

$$\langle J' M' \lambda_3 \lambda_4 | S | J M \lambda_1 \lambda_2 \rangle = \delta_{JJ'} \delta_{MM'} \langle \lambda_3 \lambda_4 | S^J | \lambda_1 \lambda_2 \rangle \quad (9.33)$$

where $|JM\lambda_1\lambda_2\rangle$ is a state of definite angular momentum constructed out of states containing particles of definite helicity. The essence of the partial-wave expansion is condensed in this equality.

We note that the final state, where the particle is scattered by angle θ , can be obtained by rotation from the initial state, where the momentum is along the z -axis ($\theta = \phi = 0$). The rotation in this case is around the y -axis by θ and then around the z -axis by ϕ :

$$|p\theta\phi; \lambda\rangle = e^{-iJ_z\phi} e^{-iJ_y\theta} |p00; \lambda\rangle \equiv R(\theta, \phi) |p00; \lambda\rangle \quad (9.34)$$

where $R(\theta, \phi)$ denotes the rotation operator. Conventionally the alternative form

$$R(\theta) = e^{-iJ_z\phi} e^{-iJ_y\theta} e^{iJ_z\phi} \quad (9.35)$$

is often used, which we will adopt here. It differs from Eq. (9.34) only in the phase. Now we define the rotation matrix by

$$\begin{aligned} \langle J'M'\lambda | R(\theta, \phi) | JM\lambda \rangle &= \delta_{JJ'} e^{-i(M'-M)\phi} \langle JM'\lambda | e^{-i\theta J_y} | JM\lambda \rangle \\ &= \delta_{JJ'} e^{-i(M'-M)\phi} d_{M'M}^J(\theta) \end{aligned} \quad (9.36)$$

The two states have the same helicity, as it is conserved by rotation. The rotation matrix $d_{M'M}^J(\theta)$ satisfies the orthogonality condition and is normalized by

$$\int_{-1}^1 d \cos \theta d_{M'M'}^J(\theta) d_{M''M'}^{J'}(\theta) = \delta_{JJ'} \frac{2}{2J+1} \quad (9.37)$$

General properties and expressions for $d_{M'M}^J(\theta)$ are given in the Appendix E. Multiplying Eq. (9.34) by $\langle JM; \lambda |$ from the left gives

$$\begin{aligned} \langle JM\lambda | p\theta\phi; \lambda \rangle &= \sum_{J'M'} \langle JM\lambda | R(\theta, \phi) | J'M' \rangle \langle J'M' | p00; \lambda \rangle \\ &= \sum_{M'} e^{i(M'-M)\phi} d_{M'M'}^J(\theta) \langle JM'\lambda | p00; \lambda \rangle \\ &= e^{i(\lambda-M)\phi} d_{M\lambda}^J(\theta) \langle JM\lambda | p00; \lambda \rangle \\ &\equiv N_J e^{i(\lambda-M)\phi} d_{M\lambda}^J(\theta) \end{aligned} \quad (9.38)$$

The penultimate equality follows because the momentum of the state $|p00; \lambda\rangle$ is along the z axis and its total helicity $\lambda = \lambda_1 - \lambda_2$ is exactly the same as $M = J_z$, hence

$$\langle JM'\lambda | p00; \lambda \rangle = \delta_{M'\lambda} N_J \quad (9.39)$$

N_J is calculated to be

$$N_J = \sqrt{\frac{2J+1}{4\pi}} \quad (9.40)$$

To derive N_J , we use the normalization conditions

$$\langle J' M' \lambda' | J M \lambda \rangle = \delta_{JJ'} \delta_{MM'} \delta_{\lambda' \lambda_1} \delta_{\lambda'_2 \lambda_2} \quad (9.41a)$$

$$\langle p' \theta' \phi'; \lambda' | p \theta \phi; \lambda \rangle = \delta(\cos \theta - \cos \theta') \delta(\phi - \phi') \delta_{\lambda'_1 \lambda_1} \delta_{\lambda'_2 \lambda_2} \quad (9.41b)$$

and apply them to

$$\begin{aligned} 1 &= \langle J M \lambda | J M \lambda \rangle = \sum_{\mu} \int d \cos \theta d \phi \langle J M \lambda | p \theta \phi; \mu \rangle \langle p \theta \phi; \mu | J M \lambda \rangle \\ &\stackrel{(9.38)(9.37)}{=} \frac{4\pi}{2J+1} |\langle J M \lambda | p 0 0; \lambda \rangle|^2 = \frac{4\pi}{2J+1} N_J^2 \end{aligned} \quad (9.41c)$$

Next we expand the state Eq. (9.34) in terms of angular momentum eigenstates:

$$|p \theta \phi \lambda\rangle = \sum_{J, M} |J M \lambda\rangle \langle J M \lambda | p \theta \phi; \lambda \rangle = \sum_{J M} N_J d_{M \lambda}^J e^{i(\lambda - M)\phi} |J M \lambda\rangle \quad (9.42)$$

Then the scattering matrix between the two states expressed in polar coordinates can be written

$$\begin{aligned} &\langle p_f \theta \phi; \mu | S | p_i 0 0; \lambda \rangle \\ &\stackrel{(9.33)}{=} \sum_{J M} \frac{2J+1}{4\pi} \left(e^{i(\mu - M)\phi} d_{M \mu}^J \right)^* \langle \mu | S^J(E) | \lambda \rangle \left(e^{i(\lambda - M)\phi} d_{M \lambda}^J \right) \Big|_{\theta=\phi=0} \\ &= \sum_{J M} \frac{2J+1}{4\pi} \langle \mu | S^J(E) | \lambda \rangle d_{\lambda \mu}^J(\theta) e^{i(\lambda - \mu)\phi} \\ &\langle \mu | S^J(E) | \lambda \rangle \equiv \langle \lambda_3 \lambda_4 | S^J(E) | \lambda_1 \lambda_2 \rangle, \quad \lambda = \lambda_1 - \lambda_2, \quad \mu = \lambda_3 - \lambda_4 \end{aligned} \quad (9.43)$$

To normalize the scattering amplitude, we compare the scattering matrix in linear momentum space and in polar coordinates in the CM system:

$$\begin{aligned} \langle p_3 p_4; \mu | S | p_1 p_2; \lambda \rangle &= \delta_{if} - (2\pi)^4 i \delta^4 \left(\sum p_i - \sum p_f \right) \mathcal{M}_{fi} \\ &= (2\pi)^4 \delta^4(P_i - P_f) \times (4\pi)^2 \sqrt{\frac{s}{p_i p_f}} \\ &\quad \times \langle p_f \theta \phi; \mu | S | p_i 0 0; \lambda \rangle \end{aligned} \quad (9.44)$$

where $s = (p_1 + p_2)^2$. The factor in the second line results from the normalization conditions Eq. (9.41). To prove it, set $S = 1$ and use

$$\begin{aligned}
 \langle p_3 p_4 | p_1 p_2 \rangle &= (2\pi)^6 (16 E_1 E_2 E_3 E_4)^{1/2} \delta^3(\mathbf{p}_1 - \mathbf{p}_3) \delta^3(\mathbf{p}_2 - \mathbf{p}_4) \\
 \delta^3(\mathbf{p}_1 - \mathbf{p}_3) \delta^3(\mathbf{p}_2 - \mathbf{p}_4) d^3 p_3 d^3 p_4 &= \delta^3(\mathbf{P}_i - \mathbf{P}_f) \delta^3(\mathbf{p}_i - \mathbf{p}_f) d^3 p_f d^3 p_f \\
 &= \delta^4(P_i - P_f) \delta(\cos \theta - \cos \theta') \delta(\phi - \phi') \left(\frac{\partial E_f}{\partial p_f} \right)^{-1} p_f^2 d^4 p_f d \cos \theta d \phi
 \end{aligned} \tag{9.45}$$

and symmetrize between the initial and the final state. We define the scattering matrix $T^J(E)$ by²⁾

$$\langle \lambda_3 \lambda_4 | S^J(E) | \lambda_1 \lambda_2 \rangle = \delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} + i \langle \lambda_3 \lambda_4 | T^J(E) | \lambda_1 \lambda_2 \rangle \tag{9.46}$$

Then using Eq. (9.44) and comparing with the cross section formula (6.90) for $d\sigma/d\Omega|_{\text{CM}}$, we obtain

$$\frac{d\sigma}{d\Omega} = |f_{\lambda_3 \lambda_4, \lambda_1 \lambda_2}|^2 \tag{9.47a}$$

$$f_{\lambda_3 \lambda_4, \lambda_1 \lambda_2} = \sqrt{\frac{p_f}{p_i}} \frac{\mathcal{M}_{fi}}{8\pi\sqrt{s}} \tag{9.47b}$$

$$= \frac{1}{2ip_i} \sum_J (2J+1) \langle \lambda_3 \lambda_4 | (S^J(E) - 1) | \lambda_1 \lambda_2 \rangle d_{\lambda\mu}^J(\theta) e^{i(\lambda-\mu)\phi} \tag{9.47c}$$

The equation is an extension of Eq. (9.26). In fact, referring to Eq. (E.21) in Appendix E,

$$d_{00}^J = P_J(\theta) \tag{9.48}$$

This agrees with Eq. (9.26) when the particles have no spin.

Unitarity of S-matrix

From the unitarity of the scattering matrix we have

$$S S^\dagger = S^\dagger S = 1 \tag{9.49}$$

Sandwiching it between an initial and final state gives

$$\delta_{fi} = \sum_n S_{nf}^* S_{ni} \tag{9.50}$$

In terms of the T matrix defined in Eq. (9.46)

$$i(T_{if}^* - T_{fi}) = \sum_n T_{nf}^* T_{ni} \tag{9.51}$$

2) $T^J = (S^J - 1)/2i$ is also often used in the literature.

If it is a helicity amplitude, time reversal invariance requires $S_{fi} = S_{if}$ [Eq. (9.121)],

$$2 \operatorname{Im}(T_{fi}) = \sum_n T_{nf}^* T_{ni} \quad (9.52)$$

For forward scattering ($i = f$),

$$2 \operatorname{Im}(T_{ii}) = \sum_n |T_{ni}|^2 \quad (9.53)$$

The right hand side (rhs) is proportional to the total cross section and the left hand side (lhs) to the imaginary part of the forward scattering amplitude. The relation is general in the sense that it is valid for inelastic as well as elastic scattering. Namely,

$$\sigma_{\text{TOT}} = k \operatorname{Im}[f(\theta = 0)] \quad (9.54)$$

To obtain the proportionality constant, we use Eq. (9.47) and

$$T_{fi} = -(2\pi)^4 \delta^4(P_i - P_f) \mathcal{M}_{fi} \quad (9.55)$$

Referring to the cross section formula in terms of $\mathcal{M}_{fi}$ [see (6.86)]

$$\begin{aligned} \sigma_{\text{TOT}} &= \sum_f (2\pi)^4 \delta^4(P_i - P_f) \frac{|\mathcal{M}_{fi}|^2}{2s\lambda(1, x_1, x_2)} = \frac{2 \operatorname{Im}(\mathcal{M}_{fi})}{4p\sqrt{s}} \\ &= \frac{4\pi}{p} \operatorname{Im}\{f(0)\} \end{aligned} \quad (9.56)$$

where we have used the relation $p_{\text{CM}} = \lambda(s, m_1^2, m_2^2)/(2\sqrt{s}) = \sqrt{s}\lambda(1, x_1, x_2)/2$ [Eq. (6.56)].

Problem 9.3

Using Eq. (9.47), calculate the total cross section for the elastic scattering

$$\sigma_{\text{TOT}} = \int d\Omega \sum_{\lambda_3 \lambda_4} |f_{\lambda_3 \lambda_4, \lambda_1 \lambda_2}|^2 \quad (9.57)$$

and then prove that

$$\sigma_{\text{TOT}}(\lambda_1, \lambda_2) = \frac{4\pi}{p} \operatorname{Im}(f_{\lambda_3 \lambda_4, \lambda_1 \lambda_2}) \quad (9.58)$$

assuming there is no inelastic scattering.

9.2

Discrete Symmetries

Discrete symmetries are not functions of a continuous parameter and no infinitesimal variation or differentiation is possible. Parity transformation, time reversal and charge conjugation to exchange particles and antiparticles are such examples.

9.2.1

Parity Transformation

General Properties Parity, or P operation for short, is often referred to as “mirroring” (reflection) of an object, but the exact definition is reversal of signs on all space coordinates:

$$\mathbf{x} \xrightarrow{P} -\mathbf{x} \quad (9.59)$$

Performing the parity transformation again brings the state back to the original one, that is $P^2 = 1$. Therefore the corresponding unitary transformation U_P satisfies

$$U_P U_P = 1 \quad \therefore \quad U_P = U_P^{-1} = U_P^\dagger \quad (9.60)$$

The parity transformation is at the same time unitary and hermitian, hence an observable by itself. Its eigenvalue is ± 1 . It is known that parity is violated maximally in the weak interaction, but that it is conserved in the strong and electromagnetic interactions, at least within experimental limits.

Momentum and Angular Momentum There are many observables that have their own parity. Classically, the momentum $\mathbf{p}$ is $m\mathbf{v} = m d\mathbf{x}/dt$ and quantum mechanically $-i\nabla$; it changes its sign under parity transformation. Orbital angular momentum, on the other hand, has positive parity since it is given by $\mathbf{x} \times \mathbf{p}$. The parity of the spin angular momentum $\mathbf{S}$ is not clear, but we can assume it has the same property as its orbital counterpart. The above examples show that there are two kinds of (three-dimensional space) vectors. Those that change their sign under the parity transformation are called polar vectors and the others, which do not, axial vectors. Similarly, a quantity that is rotationally invariant but changes its sign under the P operation is called a pseudoscalar. One example is the helicity operator $\boldsymbol{\sigma} \cdot \mathbf{p}/|\mathbf{p}|$, the spin component along the momentum direction.

Intrinsic Parity of a Particle When the Hamiltonian is a function of the distance $r = |\mathbf{x}|$

$$H = \frac{\mathbf{p}^2}{2m} + V(r) \quad (9.61)$$

it is invariant under parity operation, and if $\psi(\mathbf{x})$ is a solution then $\psi(-\mathbf{x})$ is also a solution:

$$\psi'(-\mathbf{x}) = U_P \psi(\mathbf{x}) = \pm \psi(\mathbf{x}) \quad (9.62)$$

Generally, a particle has its own intrinsic parity. When operated on a one particle state,

$$P|q\rangle = \pm \eta_P |Pq\rangle \quad (9.63)$$

Here, q is a quantum number that specifies the state, and η_P is a phase that accompanies the parity transformation. The phase degree of freedom is allowed because the expectation value of an observable is always sandwiched by a state and its conjugate. Usually, the phase is taken to be 1 for simplicity.

Parity of Fields The parity of the Dirac bilinears was already given in Sect. 4.3.3. The parity of the scalar, vector and tensor fields are the same as the corresponding Dirac bilinears. We reproduce in Table 9.2 the parity transformation properties of the fields.

Parity of the Photon Now let us consider the parity of the photon. Since the time component of a charged current ($\rho = j^0$) is a scalar with positive parity and the space component is a polar vector, the Maxwell equations

$$\nabla \cdot \mathbf{E} = q\rho, \quad \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \quad (9.64)$$

tell us that

$$\mathbf{E}(\mathbf{x}) \rightarrow \mathbf{E}'(-\mathbf{x}) = -\mathbf{E}(\mathbf{x}), \quad \mathbf{B}(\mathbf{x}) \rightarrow \mathbf{B}'(-\mathbf{x}) = \mathbf{B}(\mathbf{x}) \quad (9.65)$$

Since the electromagnetic potential is connected to the field by

$$\mathbf{E} = \nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A} \quad (9.66)$$

the transformation property of the potential is given by

$$\phi(\mathbf{x}) \rightarrow \phi'(-\mathbf{x}) = \phi(\mathbf{x}), \quad \mathbf{A}(\mathbf{x}) \rightarrow \mathbf{A}'(-\mathbf{x}) = -\mathbf{A}(\mathbf{x}) \quad (9.67)$$

This means the photon is a polar vector and has negative parity.

Table 9.2 Property of Dirac bilinears under parity transformation.

	$S(t, \mathbf{x})$	$P(t, \mathbf{x})$	$V^\mu(t, \mathbf{x})$	$A^\mu(t, \mathbf{x})$	$T^{\mu\nu}(t, \mathbf{x})$
P	$S(t, -\mathbf{x})$	$-P(t, -\mathbf{x})$	$V_\mu(t, -\mathbf{x})$	$-A_\mu(t, -\mathbf{x})$	$T_{\mu\nu}(t, -\mathbf{x})$

Note 1: $S = \bar{\psi}\psi$, $P = i\bar{\psi}\gamma^5\psi$, $V^\mu = \bar{\psi}\gamma^\mu\psi$, $A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi$, $T^{\mu\nu} = \bar{\psi}\sigma^{\mu\nu}\psi$

Note 2: $V^\mu = (V^0, \mathbf{V})$, $V_\mu = (V^0, -\mathbf{V})$

Parity of Many-Particle Systems

The quantum number of a many-particle system is additive for those corresponding to continuous symmetry and multiplicative for those corresponding to discrete symmetry, provided they are independent. The wave function of a many-particle system when there are no interactions between the particles is described by

$$\Psi(x_1, x_2, \dots, x_n) = \psi_1(x_1)\psi_2(x_2)\cdots\psi_n(x_n) \quad (9.68a)$$

$$\Psi' = U\Psi = \psi'_1\psi'_2\cdots\psi'_n = U\psi_1 U\psi_2\cdots U\psi_n(x_n) \quad (9.68b)$$

When $U = e^{iQ\alpha}$

$$\begin{aligned} e^{iQ\alpha}\Psi &= \prod e^{iQ_i\alpha}\psi_i = e^{i(Q_1+Q_2+\cdots)\alpha}\prod\psi_i \\ \therefore Q &= Q_1 + Q_2 + \cdots + Q_n \end{aligned} \quad (9.69)$$

When the transformation is discrete, U itself is converted to its quantum number, the rhs becomes $\prod\eta_i\Psi$ and

$$\eta = \eta_1\eta_2\cdots\eta_n \quad (9.70)$$

Parity of a Two-Body System

Consider the case where the intrinsic parity of spinless particles 1, 2 is P_1, P_2 and the wave function of the relative motion is $\Phi(x) = f(r)Y_{LM}(\theta, \phi)$, where Y_{LM} is a spherical harmonic function. Under the parity operation $\theta \rightarrow \pi - \theta$ and $\phi \rightarrow \phi + \pi$. Using the property of the spherical harmonic function $Y_{LM}(\pi - \theta, \phi + \pi) = (-1)^L Y_{LM}(\theta, \phi)$,

$$U\Psi(x_1, x_2) = U[\phi_1(x_1)\Phi(x_1 - x_2)\psi_2(x_2)] = P_1 P_2 (-1)^L \Psi \quad (9.71)$$

The parity of a three-body system can be determined similarly as $P_1 P_2 P_3 (-1)^{\ell+L}$, where ℓ is the relative angular momentum between the particles 1, 2 and L is that of the particle 3 relative to the CM of the 1 + 2 system. Now consider a reaction $a + b \rightarrow c + d$, which has relative angular momentum ℓ, ℓ' before and after the reaction, respectively:

$$\begin{aligned} \langle cd|S|ab\rangle &= \langle cd|P^{-1}SP^{-1}P|ab\rangle = P_a P_b (-1)^\ell P_c P_d (-1)^{\ell'} \langle cd|S|ab\rangle \\ \therefore \{1 - P_a P_b (-1)^\ell P_c P_d (-1)^{\ell'}\} \langle cd|S|ab\rangle &= 0 \end{aligned} \quad (9.72)$$

If $\langle cd|S|ab\rangle \neq 0$ and $[P, S] = 0$, the parity before and after the reaction is conserved:

$$P_a P_b (-1)^\ell = P_c P_d (-1)^{\ell'} \quad (9.73)$$

When the parity of the particles a, b, c is known, that of d can be determined using the above equation. We shall determine the parity of the pion in the next section this way. Equation (9.73) does not hold if $\langle cd|S|ab\rangle = 0$ by some other selection rules. Therefore, there are as many ambiguities in the determination of the parity

assignment as the number of selection rules. For instance, to determine the parity of π^0 , one can use a process $p \rightarrow p + \pi^0$ or $n \rightarrow n + \pi^0$, but to determine that of π^+ one cannot use $p \rightarrow p + \pi^+$, because it is forbidden by charge conservation. The process $p \rightarrow n + \pi^+$ is allowed, but one needs to know the relative parity of n to p . Similarly one cannot determine the parity of K^+ using the process $p \rightarrow n + K^+$ as it is forbidden by strangeness conservation. $p \rightarrow \Lambda + K^+$ is allowed, then again the parity of K^+ and Λ cannot be determined independently. The usual assumption is that all the quarks u, d, s , etc. have relatively positive parities. Since all the hadrons are made of quarks, the parities of the hadrons can be determined in principle once the relative angular momentum among the quarks is specified. The parity of the familiar baryons $(p, n, \Lambda, \dots)$, therefore, is assumed to be positive.

Parity Transformation of Helicity States*

The parity transformation property of the helicity amplitude is a bit complicated, because it does not use the orbital angular momentum explicitly. Those who are not interested in the complication of the derivation may skip this part and use only the relevant formula Eq. (9.87). The following arguments follow closely those of Jacob and Wick [217].

To determine the parity of a partial wave state $|JM\lambda\rangle$ in the helicity formalism, we must go back to the beginning and evaluate the effect on the helicity states Eq. (9.31)

$$|p00\lambda_1\lambda_2\rangle = \psi_1(p\lambda_1)\psi_2(p\lambda_2) \quad (9.74)$$

where particles 1 and 2 are moving in $+z$ and $-z$ directions, respectively. ψ_1 is obtained by Lorentz boost from its rest frame. Since the eigenvalues of the parity operation are only phases, we need to fix the relative phases of various states beforehand. They can be determined from a requirement that $\psi(0\lambda)$ satisfies the usual angular momentum formula

$$J_{\pm}|s\lambda\rangle = (J_x \pm J_y)|s\lambda\rangle = [(s \mp \lambda)(s \pm \lambda + 1)]^{1/2}|s, \lambda \pm 1\rangle \quad (9.75)$$

The parity operation does not change the spin direction and at rest $J_z = \lambda$

$$P\psi(0\lambda) = \eta\psi(0\lambda) \quad (9.76)$$

where η is the particle's intrinsic parity. It is convenient to define a mirror operation with respect to the xz -plane:

$$Y \equiv e^{-i\pi J_y} P \quad (9.77)$$

Using the relations

$$e^{-i\theta J_y}|JM\lambda\rangle = \sum_{M'} d_{M'M}^J(\theta)|JM'\lambda\rangle \quad (9.78a)$$

$$d_{M'M}^J(\pi) = (-1)^{j+M'} \delta_{M',-M} \quad (9.78b)$$

we obtain

$$Y\psi_1(p\lambda_1) = \eta_1(-1)^{s_1-\lambda_1}\psi_1(p, -\lambda_1) \quad (9.79)$$

Similarly

$$Y\psi_2(p\lambda_2) = \eta_2(-1)^{s_2+\lambda_2}\psi_2(p, -\lambda_2) \quad (9.80)$$

The reason for $-\lambda_2$ in the exponent is that the particle's momentum is $-p$, hence $M = J_z = -\lambda$. Combining the two equations

$$Y|p00\lambda_1\lambda_2\rangle = \eta_1\eta_2(-1)^{s_1+s_2-\lambda_1+\lambda_2}|p00, -\lambda_1, -\lambda_2\rangle \quad (9.81a)$$

$$\therefore P|p00\lambda_1\lambda_2\rangle = \eta_1\eta_2(-1)^{s_1+s_2-\lambda_1+\lambda_2}e^{i\pi J_y}|p00, -\lambda_1, -\lambda_2\rangle \quad (9.81b)$$

Now, what we want is the phase that appears in

$$P|JM\lambda_1\lambda_2\rangle = e^{i\alpha}|JM, -\lambda_1, -\lambda_2\rangle \quad (9.82)$$

Since the parity and the angular momentum operators commute, the phase α does not depend on M . Therefore, without loss of generality, we can set $M = \lambda = \lambda_1 - \lambda_2$. For the same reason, we can confine our argument to the case $\theta = \phi = 0$. Expanding the rhs in Eq. (9.81b) in partial waves

$$\begin{aligned} e^{i\pi J_y}|p00, -\lambda_1, -\lambda_2\rangle &= \sum_{JM\mu_1\mu_2} \sum_{J'M'\nu_1\nu_2} |JM\mu_1\mu_2\rangle \\ &\times \langle JM\mu_1\mu_2|e^{i\pi J_y}|J'M'\nu_1\nu_2\rangle \langle J'M'\nu_1\nu_2|p00, -\lambda_1, -\lambda_2\rangle \\ &\stackrel{(9.39)}{=} \sum_{JM} |JM, -\lambda_1, -\lambda_2\rangle N_J d_{M, -\lambda}^J(-\pi) \\ &= \sum_J N_J (-1)^{\lambda-J} |J\lambda, -\lambda_1 - \lambda_2\rangle \end{aligned} \quad (9.83)$$

Inserting this into Eq. (9.81b) gives

$$P|p00\lambda_1\lambda_2\rangle = \sum_J N_J \eta_1\eta_2(-1)^{s_1+s_2-J} |J\lambda, -\lambda_1, -\lambda_2\rangle \quad (9.84)$$

On the other hand, expanding the lhs of Eq. (9.81b) directly using Eq. (9.42) gives

$$\begin{aligned} P|p00\lambda_1\lambda_2\rangle &= \sum_{JM} P|JM\lambda_1\lambda_2\rangle \langle JM\lambda_1\lambda_2|p00\lambda_1\lambda_2\rangle \\ &= \sum_J N_J P|J\lambda\lambda_1\lambda_2\rangle \end{aligned} \quad (9.85)$$

Comparing Eq. (9.84) and Eq. (9.85)

$$P|J\lambda\lambda_1\lambda_2\rangle = \eta_1\eta_2(-1)^{s_1+s_2-J}|J\lambda, -\lambda_1, -\lambda_2\rangle \quad (9.86)$$

Noting $s_1 + s_2 - J$ is an integer, we finally obtain

$$P|J\lambda\lambda_1\lambda_2\rangle = \eta_1\eta_2(-1)^{J-s_1-s_2}|J\lambda, -\lambda_1, -\lambda_2\rangle \quad (9.87)$$

When the parity is conserved, applying Eq. (9.87) to the scattering matrix,

$$\begin{aligned} \langle -\lambda_3, -\lambda_4 | S^J | -\lambda_1, -\lambda_2 \rangle &= \eta_S \langle \lambda_3 \lambda_4 | S^J | \lambda_1 \lambda_2 \rangle \\ \eta_S &= \eta_1 \eta_2 \eta_3 \eta_4 (-1)^{s_3+s_4-s_1-s_2} \end{aligned} \quad (9.88)$$

Using Eq. (9.47) and $d_{nm}^J(\pi - \theta) = (-1)^{J+n} d_{n,-m}^J(\theta)$, a similar formula for the scattering amplitude can be obtained:

$$\langle -\lambda_3, -\lambda_4 | f(\theta, \phi) | -\lambda_1, -\lambda_2 \rangle = \eta_S \langle \lambda_3 \lambda_4 | f(\theta, \pi - \phi) | \lambda_1 \lambda_2 \rangle \quad (9.89)$$

Parity Violation in the Weak Interaction

We will describe the detailed dynamics of the weak interaction in Chap. 15. Here, we describe only the essence of parity violation. The momentum $\mathbf{p}$ is an observable with negative parity, but it does not mean that parity is violated, since particles with $-\mathbf{p}$ exist equally. However, if the S-matrix contains an observable that has negative parity, it means parity is violated in the scattering. For instance,

$$\mathbf{J} \cdot \mathbf{p} = J p \cos \theta \quad (9.90)$$

is such an observable. In the strong magnetic field at ultra-low temperatures, a nucleus can be polarized along the magnetic field. An asymmetry in the angular distribution of the decay particles from the nuclei means that parity is violated in the decay. Here, the angle of the particle is defined relative to the magnetic field, i.e. the spin orientation of the parent nuclei and the asymmetry is relative to a plane perpendicular to the polarization axis, i.e. whether θ is $\geq \pi/2$. The experiment of Wu et al. [392] that proved parity violation in the weak interaction was determined this way.

In π - p scattering where the initial proton is unpolarized, the final proton can be polarized perpendicular to the plane of scattering, namely $\boldsymbol{\sigma} \cdot \mathbf{k}_1 \times \mathbf{k}_2 \neq 0$; where $\mathbf{k}_1, \mathbf{k}_2$ are the pion momenta before and after scattering, respectively. This is because the parity of $\boldsymbol{\sigma} \cdot \mathbf{k}_1 \times \mathbf{k}_2$ is positive. However, if the recoiled proton is longitudinally polarized, i.e. polarized along its momentum, which means $\boldsymbol{\sigma} \cdot \mathbf{p} \neq 0$, conservation of parity is violated.

The origin of the parity-violating transition can be traced back to the Hamiltonian. If parity is conserved in an interaction, the parity operator commutes with the Hamiltonian. Since the scattering matrix is made from the Hamiltonian, it commutes with the S-matrix, too. In order for the S-matrix to contain the parity-

violating term, the Hamiltonian must contain a parity-violating component, too:

$$H = H_0 + H_{\text{PV}} \quad (9.91a)$$

$$P H_0 P^{-1} = H_0, \quad P H_{\text{PV}} P^{-1} = -H_{\text{PV}} \quad (9.91b)$$

where H_0 and H_{PV} are a scalar and a pseudoscalar, respectively. H_{PV} is the origin of the parity-violating transition. We shall see how the parity-violating observable is related to H_{PV} .

Let us pause to consider the meaning of the extra parity-violating term in the Lagrangian. There are observables that change sign by the parity transformation P . The momentum $\mathbf{p}$ and helicity $\boldsymbol{\sigma} \cdot \mathbf{p}$ are examples. Their existence alone does not mean parity violation. We claim that parity is violated if a phenomenon exhibits a different behavior in the mirror world than in ours. This means the dynamical motion in the mirror world is different for a given initial state. Namely, it happens if the transition amplitude includes a term that behaves differently in the mirror world, in other words, if it includes odd-parity terms such as $\boldsymbol{\sigma} \cdot \mathbf{p}$. As the transition amplitude is a matrix element of the Hamiltonian for given initial and final states, the Hamiltonian itself has to include an odd-parity term to induce the parity-violating phenomenon. In mathematical language, it simply means $[P, H] \neq 0$. Calculations of the transition amplitude will show that if C' is the coefficient of the odd-parity term in the Hamiltonian, the parity-violating observables always appear multiplied by C' .

The weak interaction is known to violate parity. Beta decay, $(A, Z) \rightarrow (A, Z + 1) + e^- + \bar{\nu}_e$ or in the quark model $d \rightarrow u + e^- + \bar{\nu}_e$, is mediated by a charged force carrier, the $W^\mp$ vector boson, but in the low-energy limit, the interaction Hamiltonian is well described by the four-Fermi interaction, which can be written as

$$H_{\text{INT}} = (\bar{u} \gamma_\mu d) \{ \bar{e} (C_V \gamma^\mu + C'_V \gamma^5) \nu_e \} \\ + (\bar{u} \gamma_\mu \gamma^5 d) \{ \bar{e} (C_A \gamma^\mu \gamma^5 + C'_A \gamma^\mu) \nu_e \} + \text{h.c.} \quad (9.92)$$

where u, d, ν_e, e^- stand for the Dirac fields and h.c. means the hermitian conjugate of the preceding term. Referring to Table 9.2, we see that terms with C_V, C_A , which contain $V_\mu V^\mu, A_\mu A^\mu$, are parity conserving, while those with C'_V, C'_A , which contain $V_\mu A^\mu, A_\mu V^\mu$, are parity violating. Namely, C_V, C_A and C'_V, C'_A are the strengths of the parity-conserving and parity-violating terms, respectively. Since the observed rate is proportional to the square of the amplitude, the effect of the parity violation appears as the interference term. The total rate Γ is proportional to

$|H_0|^2 + |H_{PV}|^2$. By preparing suitable initial and final states of the nucleus, it can be arranged to pick up only the C_A , C'_A part.³⁾ Then the angular distribution of the electron is shown to have the form

$$\frac{d\Gamma}{d\Omega} \sim 1 + \alpha \frac{(\mathbf{J} \cdot \mathbf{p}_e)}{m_e} = 1 + \alpha P \nu \cos \theta \quad (9.93a)$$

$$\alpha = \frac{2 \operatorname{Re}(C_A^* C'_A)}{|C_A|^2 + |C'_A|^2} \quad (9.93b)$$

P is the polarization of the parent nucleus. This shows clearly that the origin of the parity-violating observable $\mathbf{J} \cdot \mathbf{p}_e$ is the parity-violating Hamiltonian H_{PV} . The effect is maximal when $|C_A| = |C'_A|$, and experimental data have shown this is the case ($C'_A \simeq -C_A$). Experimental data have fixed the strengths as well as the phases of the coupling constants ($C'_V = -C_V$, $C'_A = -C_A$, $C_A = C_V$) and excluded other types (i.e. S, P, T) of quadrilinear forms. The Hamiltonian for beta decay has been shown to be

$$H_{\text{INT}} = \frac{G_F}{\sqrt{2}} \{\bar{u}\gamma_\mu(1 - \gamma^5)d\} \{\bar{e}\gamma^\mu(1 - \gamma^5)\nu_e\} \quad (9.94)$$

Here, C_V was replaced by the universal Fermi coupling constant $G_F/\sqrt{2}$, $G_F = 10^{-5} \times m_p^2$, where m_p is the mass of the proton. This is called the V–A interaction and was the standard phenomenological Hamiltonian before the advent of the Standard Model.

9.2.2

Time Reversal

Time Reversal in Quantum Mechanics

Time reversal of a process means to make it proceed backward in time. In a macro-world, when a glass is dropped on the floor, the reverse process is that the broken pieces fly back to where the glass was broken and amalgamate to shape the original glass. Everybody knows this does not happen. But in the micro-world, Newton's equation of motion under the influence of a force $F(\mathbf{x})$ that does not depend on time

$$m \frac{d^2 \mathbf{x}}{dt^2} = F(\mathbf{x}) \quad (9.95)$$

is invariant under the time reversal operation $t \rightarrow -t$, and this is indeed realized. Physical objects that are dependent on time, for example the velocity $\mathbf{v} = d\mathbf{x}/dt$, the momentum $\mathbf{p} = m\mathbf{v}$, the angular momentum $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ change their sign under time reversal and are suitable observables to test the hypothesis. Take the typical example of π – p scattering:

$$\pi(\mathbf{p}_1) + p(\mathbf{p}_2) \rightarrow \pi(\mathbf{p}_3) + p(\mathbf{p}_4) \quad (9.96)$$

3) If the so-called Gamow–Teller nuclear transitions with $|\Delta J| = 1$ and no parity change are selected, the vector interaction does not contribute ($C_V = C'_V = 0$).

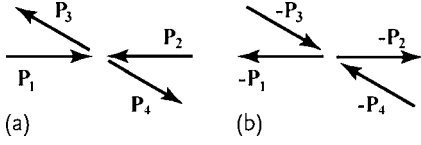


Figure 9.1 Time reversal. (a) Particles having momenta p_1 and p_2 enter, and exit as two particles with momenta p_3 and p_4 . (b) Time-reversed reaction of (a). Particles with momenta $-p_3$ and $-p_4$ enter, and exit as particles with momenta $-p_1$ and $-p_2$.

The time-reversed process is a reaction where the outgoing particles reverse their momenta, go back to where they were scattered, undergo the reverse interaction and recover the initial states with their momenta reversed (Fig. 9.1):

$$\pi(-p_3) + p(-p_4) \rightarrow \pi(-p_1) + p(-p_2) \quad (9.97)$$

Namely, their positions are unchanged, but their momenta reversed and the initial and final states are interchanged. Expressing this in terms of the transition amplitude

$$\langle T\phi(\mathbf{p}_f) | T\psi(\mathbf{p}_i) \rangle = \langle \psi'(\mathbf{p}'_i) | \phi'(\mathbf{p}'_f) \rangle = \langle \phi(-\mathbf{p}_f) | \psi(-\mathbf{p}_i) \rangle^* \quad (9.98)$$

The time-reversal operation seems to include the complex conjugate operation in addition to ordinary unitary transformation. Let us see if the notion is valid in quantum mechanics. Consider a Schrödinger equation

$$i \frac{\partial \psi(t)}{\partial t} = H \psi(t) \quad (9.99)$$

and reverse the time direction of the wave function (H is independent of time)

$$t \rightarrow t' = -t, \quad \psi(t) \rightarrow \psi'(t') = T\psi(t) \quad (9.100)$$

Then

$$i \frac{\partial \psi'(t')}{\partial t'} = H' \psi'(t'), \quad H' = T H T^{-1} \quad (9.101)$$

If $H' = H$,

$$-i \frac{\partial (T\psi)}{\partial t} = H(T\psi) \quad (9.102)$$

This shows that in order for the transformed wave function $T\psi$ to satisfy the same equation as Eq. (9.99), requiring $T H T^{-1} = H$ is not enough; it is also necessary to take the complex conjugate of both sides. The process is known as Wigner's time reversal. In other words, Eq. (9.98) is a necessary condition for the time reversal operation in quantum mechanics. Under time reversal, the wave function is transformed to

$$\psi(t) \xrightarrow{T} \psi'(t') = T\psi(t) = \psi^*(-t) \quad (9.103)$$

When ψ represents a plane wave with momentum $\mathbf{p}$

$$\psi(t, \mathbf{x}; \mathbf{p}) \sim e^{i\mathbf{p} \cdot \mathbf{x} - iEt} \xrightarrow{T} e^{i\mathbf{p} \cdot \mathbf{x} - iEt'}^*|_{t'=-t} = e^{-i\mathbf{p} \cdot \mathbf{x} - iEt} \quad (9.104)$$

time reversal reverses the momentum of the state to $-\mathbf{p}$ as expected. In general, a transformation of the form

$$\begin{aligned} \psi(t) &= a\phi_1(t) + b\phi_2(t) \\ \rightarrow \psi' &= a^*\phi_1'(t') + b^*\phi_2'(t') = a^*\phi_1^*(-t) + b^*\phi_2^*(-t) \end{aligned} \quad (9.105)$$

is called an antiunitary transformation. This means the T operation is expressed as a product of a unitary transformation U and complex conjugate operation K . Time-reversal invariance means the observables do not change under the operation Eq. (9.98). As the observables are expressed as the square of the transition amplitude, their antiunitarity does not produce any contradiction.

In fact, the necessity as well as consistency of complex conjugation can be shown in many examples. For instance, in classical mechanics $\mathbf{p} = m d\mathbf{x}/dt \xrightarrow{T} -\mathbf{p}$, but in quantum mechanics $\mathbf{p}$ is a space differential operator and has no t dependence. By complex conjugation it obtains the right sign. Besides, the quantum condition

$$[x_i, p_j] = i\delta_{ij} \quad (9.106)$$

and the commutator of the angular momentum

$$[L_i, L_j] = i\epsilon_{ijk}L_k \quad (9.107)$$

are not invariant for the change $\mathbf{p} \rightarrow -\mathbf{p}$, $\mathbf{L} \rightarrow -\mathbf{L}$ but are invariant after an additional complex conjugate operation. In conclusion, the time-reversal operation requires, in addition to flipping the time, the complex conjugate to be taken of all c -numbers in the equation.

Problem 9.4

Show the transformation properties under the time-reversal operation.

Electric field	$\mathbf{E}(t)$	$\rightarrow \mathbf{E}'(-t)$	$= \mathbf{E}(t)$	(9.108)
Magnetic field	$\mathbf{B}(t)$	$\rightarrow \mathbf{B}'(-t)$	$= -\mathbf{B}(t)$	
Potential	$A^\mu = (\phi, \mathbf{A})$	$\rightarrow A^{\mu'}(t')$	$= (\phi'(-t), \mathbf{A}'(-t))$	
			$= (\phi(t), -\mathbf{A}(t)) = A_\mu(t)$	

Time Reversal of the S-Matrix

The S -matrix is defined as $\lim_{t' \rightarrow -\infty}^{t \rightarrow \infty} U(t, t')$ [Eqs. (6.11), (6.12)]. Since $T U(t) T^{-1} = U(-t)$, it is obvious that $T S T^{-1} = S^\dagger$. Another way of seeing this is to use

$$\langle \text{out} : q_f | q_i : \text{in} \rangle = \langle \text{out} : q_f | S | q_i : \text{out} \rangle = \langle \text{in} : q_f | S | q_i : \text{in} \rangle \quad (9.109)$$

From the first expression, we obtain

$$\begin{aligned}\langle \text{out} : q_f | q_i ; \text{in} \rangle &= \langle \text{out} : q_f | T^{-1} T | q_i ; \text{in} \rangle = \langle \text{in} : T(q_f) | T(q_i) : \text{out} \rangle^* \\ &= \langle \text{out} : T(q_i) | T(q_f) ; \text{in} \rangle = \langle \text{out} : T(q_i) | S | T(q_f) ; \text{out} \rangle\end{aligned}\quad (9.110a)$$

where we have used $T|\text{in}\rangle = |\text{out}\rangle$, $T|\text{out}\rangle = |\text{in}\rangle$ to go to the third equality. From the last expression in Eq. (9.109), we obtain

$$\begin{aligned}\langle \text{in} : q_f | S | q_i ; \text{in} \rangle &= \langle \text{in} : q_f | T^{-1} T S T^{-1} T | q_i ; \text{in} \rangle \\ &= \langle \text{out} : T(q_f) | (T S T^{-1}) | T(q_i) : \text{out} \rangle^* \\ &= \langle \text{out} : T(q_i) | (T S T^{-1})^\dagger | T(q_f) : \text{out} \rangle\end{aligned}\quad (9.110b)$$

Comparing the last expressions of Eq. (9.110a) and Eq. (9.110b), we conclude

$$T S T^{-1} = S^\dagger \quad (9.111)$$

Time Reversal of Partial Waves

As the spin operator has no correspondence in classical mechanics, we are not sure how it transforms under the T operation. Let us assume its transformation property is the same as that of the orbital angular momentum, then $s \sim \mathbf{x} \times \mathbf{p}$ changes its sign, too. Therefore

$$T J_i T^{-1} = -J_i, \quad T J_\pm T^{-1} = -J_\mp \quad (9.112a)$$

$$J_z T |J M\rangle = -T J_z |J M\rangle = -M T |J M\rangle$$

$$\therefore T |J M\rangle = \eta(J, M) |J, -M\rangle \quad (9.112b)$$

The second equality of the first line follows from complex conjugation of i . To determine the phase η , we use the conventional relation among the angular momentum eigenstates:

$$\begin{aligned}T J_- |J M\rangle &= [(J + M)(J - M + 1)]^{1/2} T |J M - 1\rangle \\ &= [(J + M)(J - M + 1)]^{1/2} \eta(J, M - 1) |J, -M + 1\rangle \\ &= -J_+ T |J, M\rangle = -\eta(J, M) J_+ |J, -M\rangle \\ &= -\eta(J, M) [(J + M)(J - M + 1)]^{1/2} |J, -M + 1\rangle \\ \therefore \eta(J, M) &= -\eta(J, M - 1)\end{aligned}\quad (9.113a)$$

This equation means

$$\eta(J, M) = \eta(J) (-1)^M \quad (9.114)$$

$\eta(J)$ is a factor independent of M , and can be chosen considering the rotational

invariance of the S-matrix and the antiunitarity of the T reversal. By choosing

$$\eta(J) = (-1)^J \quad (9.115)$$

we can fix the phase as

$$T|JM\rangle = (-1)^{J-M}|J, -M\rangle \quad (9.116)$$

The phase of the helicity state can be determined similarly:

$$T|JM\lambda\rangle = (-1)^{J-M}|J, -M\lambda\rangle \quad (9.117)$$

To prove the equality, we use the fact that the helicity does not change its sign under T and that the T-transformed one-particle state $|p\lambda\rangle$ is $\eta|-\mathbf{p}\lambda\rangle$ which is equivalent to 180° rotation on the y axis. Namely

$$T|p00\lambda_1\lambda_2\rangle = \varepsilon e^{-i\pi J_y}|p00\lambda_1\lambda_2\rangle \quad (9.118)$$

where ε is a phase factor. The rest of the argument uses the same logic as we derived the parity of the state in Eq. (9.87).

Problem 9.5

Prove that the T-operated helicity eigenfunctions for a spin 1/2 particle given by Eq. (4.14)

$$\chi_+ = \begin{bmatrix} \cos \frac{\theta}{2} e^{-i\phi/2} \\ \sin \frac{\theta}{2} e^{i\phi/2} \end{bmatrix}, \quad \chi_- = \begin{bmatrix} -\sin \frac{\theta}{2} e^{-i\phi/2} \\ \cos \frac{\theta}{2} e^{i\phi/2} \end{bmatrix} \quad (9.119)$$

are given by

$$T\chi_\pm = -i\sigma_2\chi_\pm^* = (-1)^{S-M}\chi_\mp \quad (9.120)$$

Finally, as to the helicity scattering amplitude, using the antiunitarity of the T transformation and Eq. (9.117) we find

$$\langle\lambda_3\lambda_4|S^J|\lambda_1\lambda_2\rangle = \langle\lambda_1\lambda_2|S^J|\lambda_3\lambda_4\rangle \quad (9.121)$$

Time Reversal of Fields

Scalar and Vector Fields We define the T reversal of a scalar field that obeys the Klein–Gordon equation by

$$\varphi(t, \mathbf{x}) \xrightarrow{T} \varphi'(x') = T\varphi T^{-1} = \varphi'(-t, \mathbf{x}) \quad (9.122)$$

Since the free-field Lagrangian is bilinear in form in φ and $\varphi^\dagger$, both choices

$$\varphi' = \varphi \quad \text{and} \quad \varphi' = \varphi^\dagger \quad (9.123)$$

- 4) The conventional spherical harmonic function Y_{LM} is defined with $\eta(J) = 1$. To be consistent with our definition, it has to be changed to $Y_{LM} \rightarrow i^L Y_{LM}$, but this is not a problem in this book.

are possible. But here we impose a constraint by requiring that the wave function (denoted as $\psi = \text{c-number}$) should respect the rule we saw in quantum mechanics. As the field φ can be expanded in terms of plane waves as

$$\varphi = \sum_k \frac{1}{\sqrt{2\omega}} [a_k e^{-ik \cdot x} + b_k^\dagger e^{ik \cdot x}] \quad (9.124)$$

the wave function $\psi(t, \mathbf{x})$ describing a particle with momentum $\mathbf{p}$ can be extracted by

$$\psi(t, \mathbf{x}) = \langle 0 | \varphi(t, \mathbf{x}) | p \rangle \quad (9.125)$$

Similarly, the T-reversed state should be able to be extracted by

$$\psi'(-t, \mathbf{x}) = \langle T0 | \varphi(t, \mathbf{x}) | Tp \rangle \quad (9.126)$$

But for a c-number wave function, it obeys the rule of Eqs. (9.98) and (9.103)

$$\begin{aligned} \psi'(-t, \mathbf{x}) &= \psi(-t, \mathbf{x})^* \stackrel{\text{Eq. (9.98)}}{=} \langle 0 | \varphi(-t, \mathbf{x}) | p \rangle^* \\ &= \langle T0 | T\{\varphi(-t, \mathbf{x}) p\} \rangle = \langle T0 | T\varphi(-t, \mathbf{x}) T^{-1} | Tp \rangle \end{aligned} \quad (9.127)$$

Comparing Eqs. (9.126) and (9.127), we can choose the T-transformed field as

$$T\varphi(t, \mathbf{x}) T^{-1} = \varphi(-t, \mathbf{x}) \quad (9.128)$$

subject to the additional constraint of taking the complex conjugate of all the c-numbers.⁵⁾ Referring to the expansion formula of φ , the transformation properties of the annihilation and creation operators are

$$a_k \rightarrow Ta_k T^{-1} = a_{-k}, \quad b_k^\dagger \rightarrow Tb_k^\dagger T^{-1} = b_{-k}^\dagger \quad (9.129)$$

The transformation is consistent with the expression

$$|k\rangle = \sqrt{2\omega} a_k^\dagger |0\rangle \xrightarrow{T} \sqrt{2\omega} a_{-k}^\dagger |0\rangle = |-k\rangle \quad (9.130)$$

Under T reversal, the current operator of the scalar field

$$j^\mu = iq(\varphi^\dagger \partial^\mu \varphi - \partial^\mu \varphi^\dagger \varphi) \quad (9.131)$$

changes its argument $t \rightarrow -t$ and $i \rightarrow -i$. Therefore the transformation is

$$\begin{aligned} Tj^0(t, \mathbf{x}) T^{-1} &= j^0(-t, \mathbf{x}), \quad T\mathbf{j}(t, \mathbf{x}) T^{-1} = -\mathbf{j}(-t, \mathbf{x}) \\ \text{i.e. } j^\mu(t, \mathbf{x}) &\xrightarrow{T} j_\mu(-t, \mathbf{x}) \end{aligned} \quad (9.132)$$

As the vector field obeys Proca's equation

$$\partial_\mu f^{\mu\nu} + m^2 V^\nu = j^\nu, \quad f^{\mu\nu} = \partial^\mu V^\nu - \partial^\nu V^\mu \quad (9.133)$$

the consistency argument tells us that

$$V^\mu(t, \mathbf{x}) \xrightarrow{T} V_\mu(-t, \mathbf{x}) \quad (9.134)$$

5) As a matter of fact, an alternative definition of the reversal accompanied by the hermitian conjugation operation $\varphi(t, \mathbf{x}) \rightarrow \varphi^\dagger(-t, \mathbf{x})$ exists, but we will not elaborate on it in this book.

Dirac Field As the component of angular momentum changes its sign under time reversal, rearrangement of the spin component is necessary for the Dirac field. What we have to do is to find a transformation such that

$$\psi(x) \xrightarrow{T} T\psi(x)T^{-1} = \psi'(t', x) = B\psi(-t, x) \quad (9.135)$$

satisfies the same Dirac equation. Here, B is a c-number 4×4 matrix that operates on spin indices. Consider the Dirac equation for the wave field

$$(\gamma^\mu i\partial_\mu - m)\psi(t, x) = 0 \xrightarrow{T} (\gamma'^\mu i\partial'_\mu - m)^*\psi'(t', x) = 0 \quad (9.136)$$

where $\partial'_\mu = (-\partial_0, \nabla)$. Substituting Eq. (9.135) in Eq. (9.136) and multiplying B^{-1} from the left, we obtain

$$\begin{aligned} B^{-1}[\gamma^{0*}(-i)(-\partial_0) + \gamma^* \cdot (-i\nabla) - m]B\psi(-t, x) \\ = [B^{-1}\gamma^{0*}B(i\partial_0) + B^{-1}\gamma^*B \cdot (-i\nabla) - m]\psi(-t, x) = 0 \end{aligned} \quad (9.137)$$

Therefore, if B which satisfies the following relations exists

$$B^{-1}\gamma^{0*}B = \gamma^0, \quad B^{-1}\gamma^{k*}B = -\gamma^k \quad (9.138)$$

the time reversed field ψ satisfies the same equation and T transformation invariance holds. Considering $\gamma^{2*} = -\gamma^2$, $\gamma^{\mu*} = \gamma^\mu$ ($\mu \neq 2$), we can adopt as B

$$B = i\gamma^1\gamma^3 = -i\gamma^5C = -\Sigma_2 = \begin{bmatrix} -\sigma_2 & 0 \\ 0 & -\sigma_2 \end{bmatrix} \quad (9.139)$$

Then

$$\psi'(t') = T\psi(t)T^{-1} = B\psi(-t) = i\gamma^1\gamma^3\psi(-t) \quad (9.140a)$$

$$\bar{\psi}'(t') = T\bar{\psi}(t)T^{-1} = \bar{\psi}(-t)B^{-1} = \bar{\psi}(-t)i\gamma^1\gamma^3. \quad (9.140b)$$

The matrix B has the property that

$$\begin{aligned} B &= B^\dagger = B^{-1} = -B^* \\ B^{-1}\gamma^{\mu*}B &= \gamma_\mu \end{aligned} \quad (9.141)$$

Therefore, the transformation property of the Dirac bilinear vector is given by

$$\begin{aligned} \bar{\psi}_1(t)\gamma^\mu\psi_2(t) &\xrightarrow{T} \bar{\psi}'_1(-t)\gamma^{\mu*}\psi'_2(-t) = \bar{\psi}_1B^{-1}\gamma^{\mu*}B\psi_2 \\ &= \bar{\psi}_1(-t)\gamma_\mu\psi_2(-t) \end{aligned} \quad (9.142a)$$

$$\bar{\psi}_1(t)\gamma^\mu\gamma^5\psi_2(t) \xrightarrow{T} \bar{\psi}_1(-t)\gamma_\mu\gamma^5\psi_2(-t) \quad (9.142b)$$

The transformation properties of other Dirac bilinears are given similarly and we list them in Table 9.3.

Table 9.3 Properties of Dirac bilinears under T transformation.

	$S(t, x)$	$P(t, x)$	$V^\mu(t, x)$	$A^\mu(t, x)$	$T^{\mu\nu}(t, x)$
T	$S(-t, x)$	$-P(-t, x)$	$V_\mu(-t, x)$	$A_\mu(-t, x)$	$-T_{\mu\nu}(-t, x)$

Note 1: $S = \bar{\psi}\psi$, $P = i\bar{\psi}\gamma^5\psi$, $V^\mu = \bar{\psi}\gamma^\mu\psi$, $A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi$, $T^{\mu\nu} = \bar{\psi}\sigma^{\mu\nu}\psi$

Note 2: $V^\mu = (V^0, \mathbf{V})$, $V_\mu = (V^0, -\mathbf{V})$

Experimental Tests

Principle of Detailed Balance: The amplitude of a scattering process from a initial state $|i\rangle$ to a final state $|f\rangle$ and its inverse $f \rightarrow i$ are related by Eq. (9.121). The cross section of a process $A + B \rightarrow C + D$, assuming both A and B are unpolarized is expressed as

$$\frac{d\sigma}{d\Omega}(AB \rightarrow CD) = \frac{1}{(2S_A + 1)(2S_B + 1)} \sum_{\lambda} |f_{\lambda_C \lambda_D, \lambda_A \lambda_B}|^2 \quad (9.143)$$

On the other hand the cross section of the inverse process is, again assuming both C and D are unpolarized,

$$\frac{d\sigma}{d\Omega}(CD \rightarrow AB) = \frac{1}{(2S_C + 1)(2S_D + 1)} \sum_{\lambda} |f_{\lambda_A \lambda_B, \lambda_C \lambda_D}|^2 \quad (9.144)$$

As Eqs. (9.121) and (9.47) mean

$$p_{AB}^2 \sum_{\lambda} |f_{\lambda_C \lambda_D, \lambda_A \lambda_B}|^2 = p_{CD}^2 \sum_{\lambda} |f_{\lambda_A \lambda_B, \lambda_C \lambda_D}|^2 \quad (9.145)$$

where p_{AB} and p_{CD} are the momenta of particles in CM of AB and CD, respectively. we obtain

$$\frac{d\sigma(AB \rightarrow CD)}{d\sigma(CD \rightarrow AB)} = \frac{p_{CD}^2 (2S_C + 1)(2S_D + 1)}{p_{AB}^2 (2S_A + 1)(2S_B + 1)} \quad (9.146)$$

The equation is referred to as the “principle of detailed balance”.

Strong Interaction As an example of the application of the principle of detailed balance to test time-reversal invariance, we show both cross sections of the reactions $p + {}^{27}\text{Al} \rightleftharpoons \alpha + {}^{24}\text{Mg}$ in Fig. 9.2. The data show T-reversal invariance is preserved in the strong interaction to better than 10^{-3} .

Weak Interaction

There is evidence that small CP violation exists in the kaon and B-meson decays (see Chap. 16), hence the same amount of T violation is expected to exist as long as CPT invariance holds, which is discussed in Sect. 9.3.3. In fact, a T-violation effect consistent with CPT invariance has been observed in the neutral K meson

decays [107], which will be described in detail in Chap. 16. With this one exception, there is no other evidence of T violation. As it offers some of the most sensitive tests for models of new physics, we will discuss some of the experimental work below.

Observables that violate T invariance can be constructed as a triple product of T-odd variables like $\boldsymbol{\sigma} \cdot \mathbf{p}_1 \times \mathbf{p}_2$, $\mathbf{p}_1 \cdot \mathbf{p}_2 \times \mathbf{p}_3$. In the first example, $\boldsymbol{\sigma}$ can be prepared by a polarized beam, polarized targets or spin of scattered/decayed particles.

$K^+ \rightarrow \pi^0 + \mu^+ + \nu_\mu$: In the decay $K^+ \rightarrow \pi^0 + \mu^+ + \nu_\mu$, the transverse component of muon polarization relative to the decay plane determined by $\mathbf{p}_\mu \times \mathbf{p}_\pi$ (which is $\boldsymbol{\sigma}_\mu \cdot \mathbf{p}_\mu \times \mathbf{p}_\pi$) was determined to be $1.7 \pm 2.5 \times 10^{-3}$ [225, 226]. The imaginary (i.e. T-violating) part of the decay amplitude was also determined to be $\text{Im}(\xi) = -0.006 \pm 0.008$ (see Sect. 15.6.3 for the definition of ξ).

$n \rightarrow p + e^- + \bar{\nu}_e$: Another example is the triple correlation D of the neutron polarization and the momenta of electron and antineutrino in the beta decay $n \rightarrow p e^- \bar{\nu}_e$. It is defined as

$$dW \propto 1 + D \mathbf{P}_n \cdot \mathbf{p}_e \times \mathbf{p}_{\bar{\nu}} \quad (9.147a)$$

where $\mathbf{P}_n$ is the polarization of the neutron. The experimental value is given as $D = -4 \pm 6 \times 10^{-4}$ [350]. The data can be used to define the imaginary part of the coupling constant. The relative phase ϕ_{AV} is related to D by

$$\lambda = \left| \frac{g_A}{g_V} \right|, \quad \phi_{AV} = \text{Arg} \left[\frac{g_A}{g_V} \right] \quad (9.147b)$$

$$\sin \phi_{AV} = D \frac{(1 + 3\lambda^2)}{2\lambda} \quad (9.147c)$$

where g_A, g_V are axial and vector coupling constants of the weak interaction [C_V and C_A of Eq. (9.92)]. ϕ_{AV} is given as $180.06^\circ \pm 0.07^\circ$.

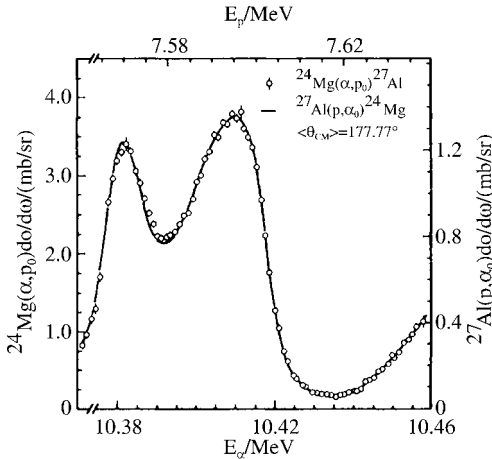


Figure 9.2 Test of detailed balance and time reversibility in the reaction $^{27}\text{Al} + p \rightleftharpoons ^{24}\text{Mg} + \alpha$. The intensity of the T-violating effect ($\xi = |f_{\text{T-violating}}/f_{\text{T-inv}}|^2$) is smaller than 5×10^{-4} [66].

The Electric Dipole Moment of the Neutron

The definition of the electric dipole moment (EDM) is classically

$$\mathbf{d} = \int \rho(\mathbf{x}) \mathbf{x} d^3x \quad (9.148)$$

which is nonzero if the charge distribution within an object is polarized, in other words, not distributed evenly. When the object is a particle like the neutron, regardless of whether of finite or point size, its only attribute that has directionality is the total spin $\boldsymbol{\sigma}$. If the particle has a finite $\mathbf{d}$, it has to be proportional to $\boldsymbol{\sigma}$. While the transformation property of $\boldsymbol{\sigma}$ under P, T is $+, -$, respectively, that of $\mathbf{d}$ is $-, +$, as can be seen from Eq. (9.148). Therefore the existence of the EDM of a particle violates both P and T. As the neutron is a neutral composite of quarks, its EDM can be a sensitive test of T-reversal invariance in the strong interaction sector as well as the weak interaction. The interaction of the magnetic and electric dipole moments with the electromagnetic field is given by

$$H_{\text{INT}} = -\boldsymbol{\mu} \cdot \mathbf{B} - \mathbf{d} \cdot \mathbf{E} \quad (9.149)$$

Its relativistic version can be expressed as

$$H_{\text{INT}} = -q \bar{\psi} \gamma^\mu \psi A_\mu + \frac{i}{2} \bar{\psi} \gamma^5 \sigma^{\mu\nu} \psi F_{\mu\nu} \quad (9.150)$$

Problem 9.6

Show that in the nonrelativistic limit, Eq. (9.150) reduces to Eq. (9.149).

For an $s = 1/2$ particle, depending on the magnetic quantum number $m = \pm 1/2$, the energy level splits in the electromagnetic field. When an oscillating field is injected, a resonance occurs corresponding to the Larmor frequency

$$\Delta E = 2\mu B \pm 2dE = h(\nu \pm \Delta\nu) \quad (9.151)$$

Here $\pm$ corresponds to the directions of the magnetic and electric fields being either parallel or antiparallel to each other. Therefore from the deviation $\Delta\nu$ of the resonance frequency when the electric field is applied, one can determine the strength of the EDM. What has to be carefully arranged experimentally is the parallel alignment of the electric field relative to the magnetic field, because if there exists a perpendicular component $\mathbf{E}_\perp$ a magnetic field of $\gamma \mathbf{v} \times \mathbf{E}_\perp$ for a moving system with velocity $\mathbf{v}$ is induced and gives a false signal.

An example material used for the experiment is the ultra cold neutron (UCN) of $T \sim 0.002$ K, which has energy $\sim 2 \times 10^{-7}$ eV and velocity $v \sim 6 \text{ ms}^{-1}$. This is the energy the neutron obtains in the earth's gravitational field when it drops 2 m. The de Broglie wavelength is very long ($\lambda \sim 670 \text{ \AA}$), and consequently the neutron interferes coherently with material, allowing the index of refraction n to be defined:

$$n = \left[1 - \frac{\lambda^2 N a_{\text{coh}}}{\pi} \pm \frac{\mu B}{M v^2/2} \right]^{1/2} \quad (9.152)$$

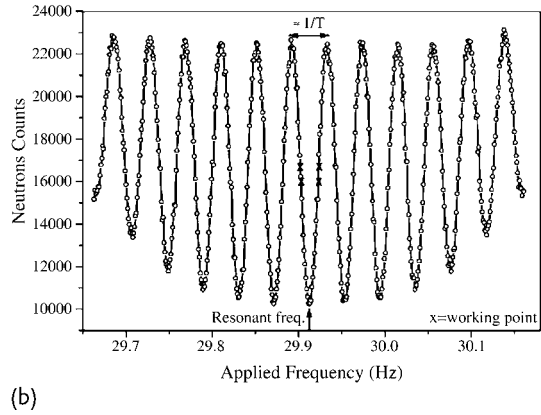
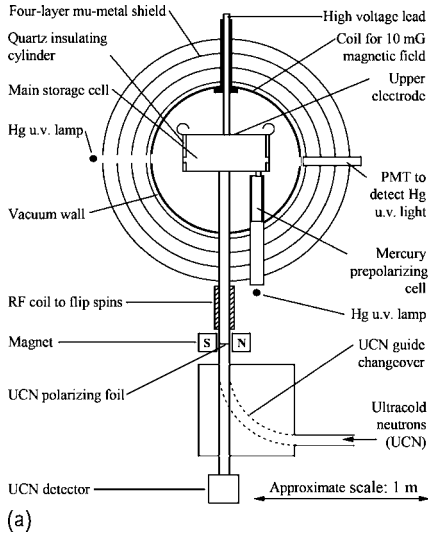


Figure 9.3 (a) The neutron EDM experimental apparatus of the RALySussex experiment at ILL [37, 198]. The bottle is a 20 liter storage cell composed of a hollow upright quartz cylinder closed at each end by aluminum electrodes that are coated with a thin layer of carbon. A highly uniform $1 \mu\text{T}$ magnetic field B_0 parallel to the axis of the bottle is generated by a coil and the electric field (E_0) is generated by applying high voltage between the electrodes. The storage volume is situated within

four layers of mu metal, giving a shielding factor of about 10 000 against external magnetic fluctuations. (b) A magnetic resonance plot showing the UCNs with spin-up count after the spin precessing magnetic field has been applied. The peak (valley) corresponds to the maximum (minimum) transmission through the analyzing foil. The measured points to detect the EDM effect are marked as four crosses. The corresponding pattern for spin down is inverted but otherwise identical.

where N is the number of nuclei per unit volume, a_{coh} is the coherent forward scattering amplitude and $Mv^2/2$ is the kinetic energy of the neutron. The $\pm$ sign depends on whether the magnetic moment and the field are parallel or antiparallel. Inserting the actual values, we obtain a total reflection angle $\sim 5^\circ$ for $v \sim 80 \text{ m}^{-1}$, but for a velocity $v < 6 \text{ m}^{-1}$ the angle exceeds 90° and total reflection is obtained at any angle. Namely, the neutron, if it is slow enough, can be transported through a bent tube just like light through an optical fiber and also can be confined in a bottle for a long duration of time (referred to as a magnetic bottle or a neutron bottle). That the refractive index differs depending on the polarization is used to separate neutrons of different polarization.

Thermal neutrons emerging from the deuterium moderator of a reactor are transported through a curved nickel pipe and further decelerated by a totally reflecting turbine before arriving at the entrance of the apparatus. Figure 9.3 shows the apparatus for the experiment [37, 198].

A magnetized iron–cobalt foil of $1 \mu\text{m}$ thickness is used to block neutrons of one spin orientation using the principle of total reflection stated above. The neutron bottle, a cylindrical 20-liter trap within a $1 \mu\text{T}$ uniform magnetic field B_0 , is able to

store neutrons for more than 100 s. The electric field, of approximately 10 kV/cm, was generated by applying high voltage (HV) to the electrode. After application of a resonant oscillating magnetic field (~ 30 Hz) perpendicular to $\mathbf{B}_0$, which turns the neutron spin perpendicular to the magnetic field and makes it precess freely, the shutter is opened, and the neutrons leave the bottle, dropping and passing through the iron–nickel foil again, which this time acts as an analyzer of the polarization. Only those that remain in the initial spin state can reach the detector, which is located below the UCN polarizing foil. The ratio of the two spin states is determined by the exact value of the applied frequency, as shown in Eq. (9.151).

Four points, marked with crosses, slightly off resonance, are measured where the slope is steepest. The existence of the EDM should appear as a frequency shift when the electric field is applied. The measurement did not detect any shift, giving the upper limit of the dipole moment of the neutron as

$$d_n < 2.9 \times 10^{-26} \text{ e cm} \quad (9.153)$$

Let us pause and think of the sensitivity that the measurement represents. Crudely speaking, the neutron is spatially spread to the size of the pion Compton wavelength ($\sim 10^{-13}$ cm). If the T violation effect is due to the weak interaction, it is probably reasonable to assume that the relative strength of emission and reabsorption of the W boson is $\sim (g_w^2/m_w^2)/(g_s^2/m_\pi^2)$. Here, g_w , g_s are the strength of the weak and strong interactions. If we assume $g_w \sim g_s$, following the spirit of the unified theory, then the strength of the EDM induced by the weak interaction is expected to be roughly of the order $10^{-13} \times (m_\pi/m_w)^2 \sim 10^{-19} \text{ e cm}$. Since the experimental value is 6 orders of magnitude smaller than the expectation, we can consider the EDM of the neutron a good test bench for T-reversal invariance. At present, the expected value of the Standard Model using the Kobayashi–Maskawa model is 10^{-33} – 10^{-34} e cm . The probability of detecting the finite value of the EDM is small as long as the Standard Model is correct. However, there are models that predict a value just below the present experimental upper limit [2, 129], and an improved experimental measurement is desired.

9.3

Internal Symmetries

9.3.1

U(1) Gauge Symmetry

Conserved Charge

We have already introduced the most fundamental result of the internal symmetry, the conserved charge current Eq. (5.38), in Chap. 5:

$$J^\mu = i q [\varphi^\dagger \partial^\mu \varphi - \partial^\mu \varphi^\dagger \varphi] \quad (9.154)$$

This was the result of the Lagrangian's symmetry, namely invariance under the phase transformation

$$\varphi \rightarrow e^{-iq\alpha}\varphi, \quad \varphi^\dagger \rightarrow \varphi^\dagger e^{iq\alpha} \quad (9.155)$$

which generally holds when the Lagrangian is bilinear in complex fields, including Dirac fields. The transformation Eq. (9.155) has nothing to do with the space-time coordinates and is an example of internal symmetry, referred to as $U(1)$ gauge symmetry of the first kind.⁶ The space integral of the time component of Eq. (9.154) is a conserved quantity and generically called a charge operator:

$$Q = -iq \int d^3x \sum_r [\pi_r(x)\varphi_r(x) - \pi_r^\dagger(x)\varphi_r^\dagger(x)] \quad (9.156a)$$

$$= iq \int d^3x [\varphi^\dagger \dot{\varphi} - \dot{\varphi}^\dagger \varphi] \quad (9.156b)$$

The first expression is a more general form, while the second one is specific to the complex Klein–Gordon field. It is easy to show that Q satisfies the commutation relations

$$[Q, \varphi] = -q\varphi, \quad [Q, \varphi^\dagger] = q\varphi^\dagger \quad (9.157)$$

This means that φ and $\varphi^\dagger$ are operators to decrease and increase the charge of the state by q . Then using the BCH formulae Eq. (9.15),

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2!}[A, [A, B]] + \cdots + \frac{1}{n!}[A, [A, [A \cdots [A, B]]]] + \cdots \quad (9.158)$$

we can show that

$$e^{i\alpha Q} \varphi e^{-i\alpha Q} = e^{-iq\alpha} \varphi, \quad e^{i\alpha Q} \varphi^\dagger e^{-i\alpha Q} = e^{+iq\alpha} \varphi^\dagger \quad (9.159)$$

Namely, the charge operator is the generator of the gauge transformation. In summary, we rephrase Noether's theorem in quantum field theory. If there is a continuous symmetry, i.e. a unitary transformation that keeps the Lagrangian invariant, a corresponding conserved (Noether) current exists. The space integral of its 0th component is a constant of time and is also the generator of the symmetry transformation.

9.3.2

Charge Conjugation

Charge conjugation (CC for short or C transformation) is an operation to exchange a particle and its antiparticle, or equivalently an operation to change the sign of

- 6) It is also called a global symmetry. When the phase is a function of space-time, i.e. $\alpha = \alpha(x)$, it is called a gauge transformation of the second kind or a local gauge transformation, which plays a key role in generating the known fundamental forces (See Chapter 18).
- 7) Here we use the word "charge" in a general sense. It includes not just electric charge, but also strangeness, hypercharge, etc.

the charge Q .⁷⁾ Examples of the C transformation are exchanges of an electron and positron $e^- \leftrightarrow e^+$, proton and antiproton $p \leftrightarrow \bar{p}$, and $\pi^+ \leftrightarrow \pi^-$. The photon and the neutral pion are not changed since they are charge neutral and the particles are their own antiparticles. The charge conjugation operation deals with variables in internal space, but because of the CPT theorem, it is inseparably connected with PT, space-time symmetry. This is because the antiparticle can be considered mathematically as a particle with negative energy-momentum traveling backward in time.

Charge Conjugation of the Field Operators

We learned in Chap. 4 that the charge conjugation operation generally involves taking the complex (hermitian for the operator) conjugate of the wave function or the field. This resulted from the fact that the interaction with the electromagnetic field is obtained by the gauge principle, which means replacement of the derivative by its covariant derivative

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i q A_\mu \quad (9.160)$$

Charge conjugation means essentially changing the sign of the charge, and the above form suggests it involves complex conjugation. Let us see if this is true.

We start by considering the charge eigenstate

$$Q|q, \mathbf{p}, s_z\rangle = q|q, \mathbf{p}, s_z\rangle \quad (9.161)$$

By definition, charge conjugation means

$$C|q, \mathbf{p}, s_z\rangle = \eta_C |-q, \mathbf{p}, s_z\rangle \quad (9.162)$$

η_C is a phase factor. Then

$$\begin{aligned} CQ|q, \mathbf{p}, s_z\rangle &= \eta_C Q|-q, \mathbf{p}, s_z\rangle = -\eta_C q|-q, \mathbf{p}, s_z\rangle \\ CQ|q, \mathbf{p}, s_z\rangle &= qC|q, \mathbf{p}, s_z\rangle = q\eta_C |-q, \mathbf{p}, s_z\rangle \end{aligned}$$

Therefore

$$CQ = -QC \quad \text{or} \quad CQC^{-1} = -Q \quad (9.163)$$

As a particle state is constructed from the creation operator $a^\dagger$ and its antiparticle from $b^\dagger$, we define the charge conjugation operator C with the properties

$$Ca_q C^{-1} = \eta_C b_q, \quad Cb_q^\dagger C^{-1} = \eta_C a_q^\dagger \quad (9.164a)$$

$$Cb_q C^{-1} = \eta_C^\dagger a_q, \quad Ca_q^\dagger C^{-1} = \eta_C^\dagger b_q^\dagger \quad (9.164b)$$

$$CC^\dagger = C^\dagger C = 1, \quad \eta_C \eta_C^\dagger = 1 \quad (9.164c)$$

Applying the charge conjugation twice brings the field back to its original form, hence its eigenvalue is ± 1 and is referred to as C parity. The definition Eq. (9.164)

of charge conjugation means the exchange $\varphi \leftrightarrow \varphi^\dagger$ for the Klein–Gordon field, as we guessed from a simple argument. Namely

$$C\varphi C^{-1} = \eta_C \varphi^\dagger, \quad C\varphi^\dagger C^{-1} = \eta_C^\dagger \varphi \quad (9.165)$$

The expression for the charge operator Eq. (9.156) shows clearly that the interchange $\varphi \rightleftharpoons \varphi^\dagger$ changes the sign of the charge operator. With the above property, it can easily be shown that the charge operator defined by Eq. (9.156) satisfies Eq. (9.163).

Problem 9.7

Prove

$$C = \exp \left[i \frac{\pi}{2} \sum_k (b_k^\dagger - a_k^\dagger)(b_k - a_k) \right] \quad (9.166)$$

has the required properties for the charge conjugation operator.

Charge Conjugation of the Dirac Field

For a field with spin, an additional operation is required to flip the spin component. For instance, for the Dirac field,

$$\begin{aligned} C\psi(x)C^{-1} &= \eta_C C' \bar{\psi}^T(x) = \eta_C C' \gamma^0 \psi^*(x) \\ C\bar{\psi}(x)C^{-1} &= -\eta_C^* \psi^T(x) C'^{-1} \end{aligned} \quad (9.167)$$

where C' on the rhs is a 4×4 matrix that acts on the spin components as given in Eq. (4.101). Using Eq. (9.167) and the property of the C' matrix [see Eq. (4.98)],

$$C'^{-1} \gamma^\mu C' = -(\gamma^\mu)^T \quad (9.168)$$

the bilinear form of the Dirac operators is changed to

$$\begin{aligned} \bar{\psi}_2 \gamma^\mu \psi_1 &\xrightarrow{C} \{-\eta_C^* \psi_2^T C'^{-1}\} \gamma^\mu \{\eta_C C' \bar{\psi}_1^T\} \\ &= \psi_2^T (\gamma^\mu)^T \bar{\psi}_1^T = -(\bar{\psi}_1 \gamma^\mu \psi_2)^T \\ &= -\bar{\psi}_1 \gamma^\mu \psi_2 = -(\bar{\psi}_2 \gamma^\mu \psi_1)^\dagger \end{aligned} \quad (9.169)$$

The minus sign in the last equality of the second line comes from the anticommutativity of the field and the removal of the transpose in the last equality is allowed because it is a 1×1 matrix. By setting $\psi_2 = \psi_1 = \psi$, we have a C-transformed current

$$j_C^\mu = C(q\bar{\psi}\gamma^\mu\psi)C^{-1} = -q\bar{\psi}\gamma^\mu\psi = -j^\mu \quad (9.170)$$

In quantum mechanics, the change of the sign is a part of the definition and had to be put in by hand, but it is automatic in field theory. Transformation properties of other types of bilinear Dirac fields can also be obtained by using Eq. (9.168), see Table 9.4.

Table 9.4 Property of Dirac bilinears under C transformation.

	$S(t, x)$	$P(t, x)$	$V^\mu(t, x)$	$A^\mu(t, x)$	$T^{\mu\nu}(t, x)$
C	$S^\dagger(t, x)$	$P^\dagger(t, x)$	$-V^{\mu\dagger}(t, x)$	$A^{\mu\dagger}(t, x)$	$-T^{\mu\nu\dagger}(t, x)$

Note 1: $S = \bar{\psi}\psi$, $P = i\bar{\psi}\gamma^5\psi$, $V^\mu = \bar{\psi}\gamma^\mu\psi$, $A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi$, $T^{\mu\nu} = \bar{\psi}\sigma^{\mu\nu}\psi$

Note 2: $V^\mu = (V^0, \mathbf{V})$, $V_\mu = (V^0, -\mathbf{V})$

Charge Conjugation of Vector Fields

As the vector and the scalar field satisfy the same Klein–Gordon equation, the charge conjugation operation transforms them as follows:

$$\varphi \rightarrow C\varphi C^{-1} = \eta_C \varphi^\dagger, \quad V^\mu \rightarrow C V^\mu C^{-1} = \eta_C V^{\mu\dagger} \quad (9.171)$$

The phase factor η_C cannot be determined for the free field. But if the field interacts, for instance, with the Dirac field, a Lorentz-invariant interaction Lagrangian has the form

$$\mathcal{L}_{\text{INT}} = f \bar{\psi}_1 \psi_2 \varphi + g \bar{\psi}_1 \gamma^\mu \psi_2 V_\mu + (\text{h.c.}) \quad (9.172)$$

The second term (h.c.) is the hermitian conjugate of the first term, and by its presence the Lagrangian becomes hermitian. The requirement of C invariance for the Lagrangian fixes the phase. Using Table 9.4 for the transformation of the Dirac bilinears, we obtain

$$\varphi \xrightarrow{C} \varphi^\dagger, \quad V^\mu \xrightarrow{C} -V^{\mu\dagger} \quad (9.173a)$$

$$f = f^*, \quad g = g^* \quad (9.173b)$$

As we noted earlier, by C operation, in addition to the electric charge, the strangeness, baryon number, lepton number and all other additive quantum numbers change their sign, and therefore the eigenstate of the charge conjugation has to be truly neutral in the sense that all the quantum numbers have to be zero.

C Parity of the Photon The photon is a massless vector boson and is described by the Maxwell equation.

$$\partial_\mu \partial^\mu A^\nu = q j^\nu \quad (9.174)$$

The interaction Lagrangian of the electromagnetic interaction is given by $\mathcal{L}_{\text{INT}} = -q j^\mu A_\mu$. Since the electric current changes its sign by C transformation, A_μ must also change sign to keep the Lagrangian invariant. Therefore, the C parity of the photon is -1 . This can also be derived from Eq. (9.173) when applied to a real vector field. The C parity of an n -photon system is $(-1)^n$. This is independent of both the orbital and the spin angular momentum.

Experimental Tests

C in the Strong Interaction: Let us consider a process

$$\bar{p} + p \rightarrow \pi^+(\pi^0) + X \quad \xleftrightarrow{C} \quad p + \bar{p} \rightarrow \pi^-(\pi^0) + \bar{X} \quad (9.175)$$

If the strong interaction conserves the charge conjugation symmetry, the angular distribution measured relative to the incoming antiprotons should exhibit a symmetry between θ and $\pi - \theta$, as shown in Fig. 9.4 and in the following equations. The energy spectrum should be identical in both cases, too:

$$\frac{d\sigma}{d\Omega}(\pi^+; \theta) = \frac{d\sigma}{d\Omega}(\pi^-; \pi - \theta) \quad (9.176a)$$

$$\frac{d\sigma}{d\Omega}(\pi^0; \theta) = \frac{d\sigma}{d\Omega}(\pi^0; \pi - \theta) \quad (9.176b)$$

$$\frac{d\sigma}{dE}(\pi^+) = \frac{d\sigma}{dE}(\pi^-) \quad (9.176c)$$

These relations have been confirmed experimentally within the errors.

C Parity of the Neutral Two-Particle System: A two-particle system consisting of a particle and its antiparticle is neutral and can be a C eigenstate. We consider them as identical particles in different states. If the spin and orbital angular momentum of the system are L, S , the C parity of the system is expressed as

$$C = (-1)^{L+S} \quad (9.177)$$

Proof: If they are fermions ($f\bar{f} = e^-e^+, p\bar{p}$, etc.) they change sign by a particle exchange because of Fermi statistics. The exchange consists of that of the charge (C), space coordinates and spin coordinates. Therefore

$$-1 = C(-1)^L(-1)^{S+1} \quad (9.178)$$

Here, we have used the fact that the spin wave function is symmetric when $S = 1$ and antisymmetric when $S = 0$. The case for the boson system can be proved similarly. $\square$

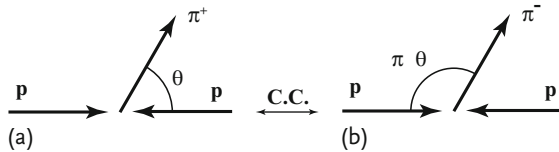


Figure 9.4 (a) $\bar{p} + p \rightarrow \pi^+ + X$, (b) $p + \bar{p} \rightarrow \pi^- + \bar{X}$. If C is conserved, $d\sigma(\bar{p} + p \rightarrow \pi^+) = d\sigma(p + \bar{p} \rightarrow \pi^-)$, where $d\sigma$ stands for either $d\sigma/d\Omega$ or $d\sigma/dE$. If the pion angle is measured relative to the incident antiprotons, $d\sigma/d\Omega(\theta) = d\sigma/d\Omega(\pi - \theta)$.

Problem 9.8

The ρ^0 meson is a vector particle with $J^P = 1^-$ and decays into $\pi^+ + \pi^-$. Prove it cannot decay into $2\pi^0$.

C in Electromagnetic Interactions: The most stringent test of C symmetry in electromagnetic interactions comes from the nonexistence of the $\pi^0 \rightarrow 3\gamma$ decay [311]:

$$\frac{\Gamma(\pi^0 \rightarrow 3\gamma)}{\Gamma_{\text{total}}} < 3.1 \times 10^{-8} \quad 90\% \text{CL} \quad (\text{CL} = \text{confidence level}) \quad (9.179)$$

The following decays are known to occur via the electromagnetic interaction through their strength (the decay rate). Asymmetry tests are not as stringent as that of decay branching ratios, but we list them here as another test. As $\eta(548)$ is a neutral scalar meson having 0^- and mass of 548 MeV, it changes to itself by C operation and so do π^0 and γ . Since π^+ and π^- are interchanged, the energy spectrum of the decays

$$\eta \rightarrow \pi^+ + \pi^- + \pi^0 \quad (9.180a)$$

$$\eta \rightarrow \pi^+ + \pi^- + \gamma \quad (9.180b)$$

should be identical. Experimentally, the asymmetry is less than 0.1%.

As π^0 , η decay into 2γ , their C parity is positive and they cannot decay into 3γ . Conversely, if a particle decays into 3γ or $\pi^0\gamma$, $\eta\gamma$ its C parity is negative.

Problem 9.9

Positronium is a bound state of an electron and a positron connected by the Coulomb force. For the $L = 0$ ground state, there are two states, with $S = 0, 1$. Show that the $S = 0$ state decays into 2γ and the $S = 1$ state into 3γ , but not vice versa.

C Violation in Weak Interactions: Next we consider weak interactions, for example

$$\mu^\mp \rightarrow e^\mp + \nu + \bar{\nu} \quad (9.181)$$

If C symmetry is respected, the helicity of the electron and that of the positron should be the same provided other conditions are equal. The helicity can be measured first by making the electron emit photons by bremsstrahlung and then letting the photons pass through magnetized iron. The transmissivity of the photon depends on its polarization, which in turn depends on the electron helicity. The result has shown that $h(e^-) \simeq -1$, $h(e^+) \simeq +1$ [111, 267]. Therefore charge conjugation symmetry is almost 100% broken in weak interactions. The origin lies in the Hamiltonian. The parity-violating weak interaction Hamiltonian was given in Eq. (9.92). If we take into account the C transformation property of the axial vector (see Table 9.4) we have

$$\bar{\psi}_2 \gamma^\mu \gamma^5 \psi_1 \xrightarrow{C} \bar{\psi}_1 \gamma^\mu \gamma^5 \psi_2 \quad (9.182)$$

This and Eq. (9.173) mean under C transformation

$$\begin{aligned} C_V &\rightarrow C_V^*, & C_V' &\rightarrow -C_V'^* \\ C_A &\rightarrow C_A^*, & C_A' &\rightarrow -C_A'^* \end{aligned} \quad (9.183)$$

Namely, if C invariance holds, C_V, C_A have to be real and C_V', C_A' have to be purely imaginary. But the decay asymmetry of a polarized nucleus that was discussed following Eq. (9.93) has shown $C_A' = -C_A = \text{real}$. Therefore, C is also maximally violated. But the combined CP symmetry does not change the V-A Hamiltonian of Eq. (9.94) and is conserved.

The chirality operator $(1 - \gamma^5)$ in front of the lepton field operator in Eq. (9.94) is the origin of $h(e^-) \simeq h(\nu_e) = -1$, $h(e^+) \simeq h(\bar{\nu}_e) = +1$ as we discussed in Sect. 4.3.5. The latter is typical of the weak interaction where the weak boson W^μ couples only to the left-handed field (i.e. $(1 - \gamma^5)\psi$). This will be discussed in detail in Chap. 15.

The V-A interaction Eq. (9.94) respects T invariance as well as CP invariance. In principle, CP invariance and T invariance are independent. However, the combined CPT invariance is rooted deeply in the structure of quantum field theory and is the subject of the next section.

As long as CPT invariance holds, CP violation means T violation. There is evidence that small CP violation exists in the kaon and B-meson decays, hence the same amount of T violation is expected to exist. In fact, a T-violation effect consistent with CPT invariance has been observed in the neutral K meson decays [107], which will be described in detail in Chap. 16.

9.3.3

CPT Theorem

CPT is a combined transformation of C, P and T. All the coordinates are inverted $(t, \mathbf{x}) \rightarrow (-t, -\mathbf{x})$ and the operation is antiunitary because of the T transformation. All c-numbers are changed to their complex conjugates. Writing the combined operator as Θ , its action is summarized as

$$\begin{aligned} (t, \mathbf{x}) &\xrightarrow{\text{CPT}} (-t, -\mathbf{x}) \\ \text{c-numbers} &\rightarrow (\text{c-numbers})^* \\ \langle \text{out} | H | \text{in} \rangle &\rightarrow \langle \Theta \text{in} | H | \Theta \text{out} \rangle \\ \Theta |q, \mathbf{p}, \sigma\rangle &\rightarrow |\bar{q}, \mathbf{p}, -\sigma\rangle \end{aligned} \quad (9.184)$$

The CPT combined transformation property of the Dirac bilinears is given by

$$\bar{\psi}_1 \Gamma \psi_2 \xrightarrow{\text{CPT}} \begin{cases} \bar{\psi}_2 \Gamma \psi_1 & \Gamma = S, P, T \\ -\bar{\psi}_2 \Gamma \psi_1 & \Gamma = V, A \end{cases} \quad (9.185)$$

independent of the order of the C, P, T operations. If we take the hermitian conjugate of the bilinears, the original form is recovered except for the sign of the coordinates. Transformation properties of the scalar, pseudoscalar, vector fields, etc.

are shown to be the same as the corresponding Dirac bilinears. It follows that any Lorentz scalar or second-rank tensor made of any number of fields are brought back to their original form by a combined operation of (CPT + hermitian conjugate). Therefore the Lagrangian density (a scalar) and the Hamiltonian density (a second rank tensor) transform under CPT

$$\mathcal{L}(t, \mathbf{x}) \xrightarrow{\text{CPT}} \mathcal{L}^\dagger(-t, -\mathbf{x}) \quad (9.186a)$$

$$\mathcal{H}(t, \mathbf{x}) \xrightarrow{\text{CPT}} \mathcal{H}^\dagger(-t, -\mathbf{x}) \quad (9.186b)$$

Since $\mathcal{L}$ and $\mathcal{H}$ are hermitian operators and the action is their integral over all space-time, we conclude that the equations of motion are invariant under CPT transformation. The Hamiltonian $H = \int d^3x \mathcal{H}$ changes the sign of time t , but as H is a conserved quantity and does not depend on time, the transformed Hamiltonian is the same as before it is transformed. As the scattering matrix is composed of the Hamiltonian, it is invariant, too. Unlike individual symmetry of C, P or T, the combined CPT invariance theorem can be derived with a few fundamental assumptions, such as the validity of Lorentz invariance, the local quantum field theory and the hermitian Hamiltonian. The violation of CPT means violation of either Lorentz invariance or quantum mechanics.⁸⁾ Theoretically it has firm foundations. Experimental evidence of CPT invariance is very strong, too.

We list a few examples of the CPT predictions:

- (1) The mass of a particle and its antiparticle is the same.

Proof: Let H be the CPT-invariant Hamiltonian. As the mass is an eigenstate of the Hamiltonian in the particles rest frame

$$\begin{aligned} m &= \langle q, \sigma | H | q, \sigma \rangle = \langle q, \sigma | \Theta^{-1} \Theta H \Theta^{-1} \Theta | q, \sigma \rangle = \langle \Theta q, \sigma | H | \Theta q, \sigma \rangle^* \\ &= \langle \bar{q}, -\sigma | H | \bar{q}, -\sigma \rangle \end{aligned} \quad (9.187)$$

The spin orientation of $|\bar{q}, -\sigma\rangle$ is different from $|q, \sigma\rangle$, but the mass does not depend on the spin orientation (see Poincaré group specification for a particle state in Sect. 3.6). Therefore the masses of the particle and antiparticle are the same. $\square$

- (2) The magnetic moment of a particle and its antiparticle is the same but its sign is reversed.

- (3) The lifetime of a particle and its antiparticle is the same.

Proof: Writing the S-matrix as

$$S_{\beta\alpha} = 1 + i2\pi\delta(M_\alpha - E_f)\langle\beta; \text{out}|T|\alpha; \text{out}\rangle \quad (9.188)$$

8) It is pointed out that CPT might be violated in quantum gravity [130, 210].

Table 9.5 Summary of transformation properties under C, P and T.

	$S(t, x)$	$P(t, x)$	$V^\mu(t, x)$	$A^\mu(t, x)$	$T^{\mu\nu}(t, x)$
P	$S(t, -x)$	$-P(t, -x)$	$V_\mu(t, -x)$	$-A_\mu(t, -x)$	$T_{\mu\nu}(t, -x)$
C	$S^\dagger(t, x)$	$P^\dagger(t, x)$	$-V^{\mu\dagger}(t, x)$	$A^{\mu\dagger}(t, x)$	$-T^{\mu\nu\dagger}(t, x)$
T	$S(-t, x)$	$-P(-t, x)$	$V_\mu(-t, x)$	$A_\mu(-t, x)$	$-T_{\mu\nu}(-t, x)$
CP	$S^\dagger(t, -x)$	$-P^\dagger(t, -x)$	$-V_\mu^\dagger(t, -x)$	$-A_\mu^\dagger(t, -x)$	$-T_{\mu\nu}^\dagger(t, -x)$
CPT	$S^\dagger(-t, -x)$	$P^\dagger(-t, -x)$	$-V^{\mu\dagger}(-t, -x)$	$-A^{\mu\dagger}(-t, -x)$	$T^{\mu\nu\dagger}(-t, -x)$

Note 1: $S = \bar{\psi}\psi$, $P = \bar{\psi}\gamma^5\psi$, $V^\mu = \bar{\psi}\gamma^\mu\psi$, $A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi$, $T^{\mu\nu} = \bar{\psi}\sigma^{\mu\nu}\psi$

Note 2: The differential operator ∂_μ transforms exactly the same as the vector field V_μ except it does not change sign under C-operation.

the total decay rate of α is given by

$$\begin{aligned}
\Gamma(\alpha, q_\alpha) &= 2\pi \sum_{\beta, q_\beta} \delta(M_\alpha - E_f) |\langle \beta, q_\beta; \text{out} | T | \alpha, q_\alpha; \text{out} \rangle|^2 \\
&= 2\pi \sum_{\beta, q_\beta} \delta(M_\alpha - E_f) |\langle \beta, q_\beta; \text{out} | \Theta^{-1} \Theta T \Theta^{-1} \Theta | \alpha, q_\alpha; \text{out} \rangle^*|^2 \\
&= 2\pi \sum_{\bar{\beta}, \bar{q}_\beta} \delta(M_{\bar{\alpha}} - E_f) |\langle \bar{\beta}, \bar{q}_\beta; \text{in} | T | \bar{\alpha}, \bar{q}_\alpha; \text{in} \rangle^*|^2 \\
&= \Gamma(\bar{\alpha}, \bar{q}_\alpha)
\end{aligned} \tag{9.189}$$

where, q_α, q_β denote quantum numbers other than particle species. We have used the fact that for one particle state, $|\alpha, q_\alpha; \text{in}\rangle = |\alpha, q_\alpha; \text{out}\rangle$, $M_\alpha = M_{\bar{\alpha}}$ and that both “in” and “out” states form complete sets of states, i.e.

$$\sum_{\beta, q_\beta} |\beta, q_\beta; \text{in}\rangle \langle \beta, q_\beta; \text{in}| = \sum_{\beta, q_\beta} |\beta, q_\beta; \text{out}\rangle \langle \beta, q_\beta; \text{out}| = 1 \tag{9.190}$$

□

Finally, in Table 9.5 we give a summary list of the transformation properties of C, P and T.

9.3.4

$SU(2)$ (Isospin) Symmetry

Isospin Multiplets

Among hadrons there are many small groups in which members having different electric charge share some common properties. For instance, in the following example members have almost the same mass values (given in units of MeV) within

groups.

$\begin{bmatrix} p \\ n \end{bmatrix}$	938.272 939.565	$\begin{bmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{bmatrix}$	139.570 134.977 139.570	$\begin{bmatrix} K^+ \\ K^0 \end{bmatrix}$	493.68 497.65	$\begin{bmatrix} \Sigma^+ \\ \Sigma^0 \\ \Sigma^- \end{bmatrix}$	1189.37 1192.64 1197.45
$\delta m/m$	0.14 %		3.3 %		0.8 %		0.33 %
(9.191)							

In addition, it is known that the strength of the interaction is almost the same within the accuracy of a few percent. Because they have different electric charges, the difference can be ascribed to the electromagnetic interaction, which is weaker by $\sim O(\alpha) = 1/137$. Therefore, if we neglect the small mass difference, there is no distinction among the members of the groups as far as the strong interaction is concerned. They constitute multiplets but are degenerate. In analogy to spin multiplets, which are degenerate under a central force, we call them “isospin multiplets”. Just as the degeneracy of the spin multiplets is resolved by a magnetic field (the Zeeman effect), that of the isospin multiplets is resolved by turning on the electromagnetic force. We conceive of an abstract space (referred to as internal space in contrast to external or real space) and consider the strong interaction as a central force with the electromagnetic force violating the rotational invariance just like the magnetic field in external space. If we identify the isospin as the equivalent of spin in real space, the mathematical structure of isospin is exactly the same as that of spin, which is the origin of the name.

Charge Independence

Historically, the concept of isospin was first proposed by Heisenberg in 1932 to consider the proton and the neutron as two different states of the same particle (nucleon) when he recognized the fact that the energy levels of mirror nuclei are very similar. Consider the potentials between the nucleons V . When

$$V_{pp} = V_{nn} \quad \text{charge symmetry} \quad (9.192)$$

we call it charge symmetry and when

$$V_{pp} = V_{nn} = V_{pn} \quad \text{charge independence} \quad (9.193)$$

we call it charge independence. To formulate the mathematics of the charge independence, we consider transformation of the proton wave function ψ_p and that of the neutron ψ_n and define the doublet function ψ by

$$\psi \equiv \begin{bmatrix} \psi_p \\ \psi_n \end{bmatrix} \rightarrow \psi' = \begin{bmatrix} \psi_p' \\ \psi_n' \end{bmatrix} = \begin{bmatrix} \alpha\psi_p + \beta\psi_n \\ \gamma\psi_p + \delta\psi_n \end{bmatrix} = U\psi \quad (9.194)$$

Charge independence means the expectation value of the potential does not change under the transformation U . Conservation of probability requires U to be a unitary matrix

$$U U^\dagger = U^\dagger U = 1 \quad (9.195)$$

Under this restriction, U should have the form

$$U = e^{i\phi} \begin{bmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{bmatrix}, \quad |\alpha|^2 + |\beta|^2 = 1 \quad (9.196)$$

If we further require U to be unimodular ($\det U = 1$), we have $\phi = 0$. Generally, any transformation that keeps the norm of N complex numbers

$$\psi^\dagger \psi = |\psi_1|^2 + |\psi_2|^2 + \cdots + |\psi_N|^2 \quad (9.197)$$

invariant makes a group that is called a unitary group $U(N)$, and when $\det U = 1$ it is called a special unitary group $SU(N)$. An $SU(N)$ matrix can be expressed as

$$U = \exp \left[-\frac{i}{2} \sum_{i=1}^{N^2-1} \lambda_i \theta_i \right] \quad (9.198)$$

where λ_i are traceless hermitian matrices (Appendix G) called generators of the $SU(N)$ transformation. For $N = 2$, there are three traceless hermitian matrices and we can adopt the Pauli matrices for them:

$$\tau_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \tau_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \tau_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (9.199)$$

Since $\mathbf{I} = \boldsymbol{\tau}/2$ satisfies the same commutation relation

$$[I_i, I_j] = i\epsilon_{ijk} I_k \quad (9.200)$$

as the angular momentum, the $SU(2)$ transformation is mathematically equivalent (holomorphic) to the rotational operation. The difference is that the operand is not a spin state in real space, but a multiplet consisting of particles of roughly the same mass as those in Eq. (9.191). Namely, the act of rotation is performed not in a real space but in a kind of abstract space (referred to as isospin space) and the state vector ψ denotes p when its direction is upward and n when downward. Such symmetry in the internal space is called an internal symmetry. External symmetries refer to those in a real space such as spin and parity. If we denote the difference of the states like p or n by a subscript r as $\phi_r(x)$, the internal symmetry operation acts on r , while the external symmetry acts on the space-time coordinate x and legs (components) of Lorentz tensors. They are generally independent operations, but sometimes connected like CPT.

As the mathematics of $SU(2)$ is the same as that of angular momentum, it is convenient to express various physical quantities using the same terminology as used in space rotation. For instance, p and n make a doublet; we say the nucleon has isospin $I = 1/2$. The three kinds of pions $\pi^\pm, \pi^0$ constitute a triplet, or a vector in isospace, and have $I = 1$. The term “charge independence” of the nuclear force means the interaction between the nucleon and the pion is rotationally invariant in isospin space, namely the interaction Hamiltonian is an isoscalar.

Conserved Isospin Current

We consider the proton and the neutron as the same particle, but they have different projections in the abstract isospin space and constitute a doublet, called an isospinor:

$$\psi = \begin{bmatrix} p \\ n \end{bmatrix} \quad (9.201)$$

Here we have adopted the notation p, n to denote ψ_p, ψ_n . The Lagrangian density is written as

$$\begin{aligned} \mathcal{L} &= \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \\ &= \bar{p}(i\gamma^\mu \partial_\mu - m)p + \bar{n}(i\gamma^\mu \partial_\mu - m)n \end{aligned} \quad (9.202)$$

They should have the same mass because of the isospin symmetry. Note, p is a spinor with four components in real space, and ψ a two-component spinor in isospin space. Therefore, ψ has 8 independent components. In exactly the same manner as a rotation in real space, a rotation in isospin space can be carried out by using a unitary operator

$$U_I = e^{-i \sum_i \alpha_j I_j} = e^{-i \boldsymbol{\alpha} \cdot \mathbf{I}} \quad (9.203)$$

where $\mathbf{I} = (I_1, I_2, I_3)$ satisfies Eq. (9.200). Let us consider transformations defined as

$$\psi \rightarrow \psi' = e^{-\frac{i}{2} \tau_j \alpha_j} \psi \equiv e^{-\frac{i}{2} \boldsymbol{\tau} \cdot \boldsymbol{\alpha}} \psi, \quad (9.204a)$$

and

$$\bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} e^{\frac{i}{2} \boldsymbol{\tau} \cdot \boldsymbol{\alpha}} \quad \text{where} \quad \bar{\psi} = (\bar{p}, \bar{n}) \quad (9.204b)$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ is a set of three independent but constant variables. Since τ_i, α_i are independent of space-time, the transformation does not change any space-time structure of the spinors p, n , and the Lagrangian density is kept invariant under the continuous isospin rotation. Equations (9.204) have exactly the same form as the $U(1)$ phase transformation of Eq. (9.155) and is called the (global) $SU(2)$ gauge transformation. Now consider an infinitesimal rotation

$$\begin{aligned} \psi &\rightarrow \psi + \delta\psi, & \bar{\psi} &\rightarrow \bar{\psi} + \delta\bar{\psi} \\ \delta\psi &= -\frac{i}{2} \tau^i \psi \varepsilon_i = -\frac{i}{2} [\tau^i]_{ab} \psi_b \varepsilon_i, & \delta\bar{\psi} &= \frac{i}{2} \bar{\psi} \tau^i \varepsilon_i = \frac{i}{2} \bar{\psi}_a [\tau^i]_{ab} \varepsilon_i \end{aligned} \quad (9.205)$$

Applying Noether's formula Eq. (5.32) to the transformation, we obtain

$$\begin{aligned} J^\mu &= \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_r)} \delta \phi_r \Big|_{\phi_r \equiv \psi, \bar{\psi}} = \frac{\delta \mathcal{L}}{\delta(\partial_\mu \psi)} \delta\psi + \delta\bar{\psi} \frac{\delta \mathcal{L}}{\delta(\partial_\mu \bar{\psi})} \\ &= \left[-\frac{\delta \mathcal{L}}{\delta(\partial_\mu \psi_a)} \frac{i}{2} [\tau^i]_{ab} \psi_b + \frac{i}{2} \bar{\psi}_a [\tau^i]_{ab} \frac{\delta \mathcal{L}}{\delta(\partial_\mu \bar{\psi}_b)} \right] \varepsilon_i \\ &= \frac{1}{2} \bar{\psi} \gamma^\mu \boldsymbol{\tau}^i \psi \varepsilon_i \end{aligned} \quad (9.206)$$

The second term in the second line vanishes because there is no $\partial_\mu \bar{\psi}$ in the Lagrangian. Since ε_i is an arbitrary constant, Eq. (9.206) defines a conserved current referred to as isospin current:

$$(J^\mu)_j = \bar{\psi} \gamma^\mu \frac{\tau_j}{2} \psi \quad (9.207)$$

It is easy to check that an isospin operator defined by

$$\hat{I}_j = \int d^3x \bar{\psi} \gamma^0 \frac{\tau_j}{2} \psi \quad (9.208)$$

satisfies the relations

$$[I_j, \psi] = -I_j \psi, \quad [I_j, \psi^\dagger] = I_j \psi^\dagger, \quad e^{i\alpha \cdot \hat{I}} \psi e^{-i\alpha \cdot \hat{I}} = e^{-i\alpha \cdot I} \psi \quad (9.209)$$

Therefore the operator $\hat{I}$ is the generator of the isospin rotation which appeared in Eq. (9.203). The isospin current differs from the Dirac current only in the presence of the isospin operator $\boldsymbol{\tau}/2$ sandwiched between the isospinors $\bar{\psi}$ and ψ .

In a similar vein, when we have three identical real Klein–Gordon fields $\phi = (\phi_1, \phi_2, \phi_3)$, we can also establish the rotational invariance of $\phi \cdot \phi = \phi_1^2 + \phi_2^2 + \phi_3^2$ and that of the Lagrangian, which can be expressed as

$$\mathcal{L} = \partial_\mu \phi^\dagger \cdot \partial^\mu \phi - m^2 \phi^\dagger \cdot \phi \quad (9.210)$$

leading to the existence of conserved isospin angular momentum, for which ϕ is a $I = 1$ triplet. Since the pions have spin 0 and constitute a triplet of almost the same mass, we can identify them with the field ϕ .⁹⁾ In analogy to Eq. (9.208), it can be expressed as

$$I_\pi = -i \int d^3x \pi_a [t]_{ab} \phi_b \quad (9.211a)$$

$$t^1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad t^2 = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix}, \quad t^3 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (9.211b)$$

However, to identify $(|\pi^+\rangle, |\pi^0\rangle, |\pi^-\rangle)$ as the $T_3 = (+1, 0, -1)$ state, it is more convenient to use a slightly different representation:

$$t_1' = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad t_2' = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad (9.212)$$

$$t_3' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

9) Note, the pion field is a $I = 1$ triplet representation of the $SU(2)$ group. If the fundamental representation with $I = 1/2$ does not exist, we may have to consider the pion as the fundamental representation of the $SO(3)$ group, as π_j 's are real fields.

Problem 9.10

Prove that I, I' , satisfy

$$[I_i, I_j] = i\varepsilon_{ijk} I_k, \quad [I_{\pi i}, I_{\pi j}] = i\varepsilon_{ijk} I_{\pi k} \quad (9.213a)$$

$$[I_i, \psi_j] = i\varepsilon_{ijk} \psi_k, \quad [I_{\pi i}, \phi_j] = i\varepsilon_{ijk} \phi_k \quad (9.213b)$$

Problem 9.11

Prove

$$\left[I_{\pi 3}, \frac{\phi_1 \pm i\phi_2}{\sqrt{2}} \right] = \pm \frac{\phi_1 \pm i\phi_2}{\sqrt{2}}, \quad [I_{\pi 3}, \phi_3] = 0 \quad (9.214)$$

This equation can be used to define

$$\frac{\phi_1 \pm i\phi_2}{\sqrt{2}} |0\rangle = \alpha_p |\pi^\pm\rangle = \mp |\pi^\pm\rangle, \quad \phi_3 = |\pi^0\rangle \quad (9.215)$$

The phase α_p was chosen to produce

$$I_{\pi \pm} |I = 1, I_3 = 0\rangle = \sqrt{2} |1, \pm 1\rangle \quad (9.216)$$

which is a special case of the usual convention adopted from the angular momentum operator:

$$I_{\pi \pm} |I, I_3\rangle = [(I \mp I_{\pi 3})(I \pm I_3 + 1)]^{1/2} |I, I_3 \pm 1\rangle \quad (9.217)$$

The real nucleon and the pion do not have exactly the same mass within the multiplet. The isospin rotation is not an exact symmetry, but as long as the difference is small we can treat the isospin as a conserved quantity.

Isospin of the Antiparticle

In order to find the isospin of the antiparticle, we apply the charge conjugation operation to Eq. (9.194). Then the field represents the antiparticle. As the C operation contains complex conjugation, the transformation matrix becomes

$$\bar{p}' = \alpha^* \bar{p} + \beta^* \bar{n} \quad (9.218a)$$

$$\bar{n}' = -\beta \bar{p} + \alpha \bar{n} \quad (9.218b)$$

here, $\bar{p}, \bar{n}$ denote the fields of the antiparticles p, n . In general, if the symmetry transformation matrix is U , that of the antiparticle is given by its complex conju-

gate U^* . Rearranging Eq. (9.218), we obtain

$$(-\bar{n}') = \alpha (-\bar{n}) + \beta \bar{p} \quad (9.219a)$$

$$\bar{p}' = -\beta^* (-\bar{n}) + \alpha^* \bar{p} \quad (9.219b)$$

which means

$$\begin{bmatrix} -\bar{n}' \\ \bar{p}' \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{bmatrix} \begin{bmatrix} -\bar{n} \\ \bar{p} \end{bmatrix} = U \begin{bmatrix} -\bar{n} \\ \bar{p} \end{bmatrix} \quad (9.220)$$

The equation means that if (p, n) is an isodoublet, $(-\bar{n}, \bar{p})$ is also an isodoublet that transforms in the same way.

Charge Symmetry We mentioned charge symmetry in Eq. (9.192). It can be defined more explicitly as the rotation in isospin space by π on the y -axis. For a doublet its operation is

$$e^{-i\pi I_2} = e^{-i(\pi/2)\tau_2} = -i\tau_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (9.221)$$

10

Path Integral: Basics

The path integral is now widely used in many fields wherever quantum mechanical treatments are necessary, including quantum field theory and statistical mechanics. In solving the simple problems that are treated in this book, it is hardly a more powerful tool than the ordinary canonical formalism. In fact, when ambiguities arise in the path integral formulation, they are usually resolved by comparing it with the canonical quantization formalism or Schrödinger equation. But because of its powerful concept and widespread use, acquaintance with the method is indispensable. Feynman's path integral expresses a particle's dynamical propagation as a coherent sum of an infinite number of amplitudes over all possible paths. The path that is dictated by the classical equation of motion is simply the most probable, thus the method yields important insight into the connection between the classical and quantum mechanical approaches. It has the remarkable feature of not using operators in Hilbert space. Only ordinary c-numbers appear in the formula, although at the expense of an infinite number of integrations. What we deal with primarily is the propagator or the transition amplitude.

The purpose of this chapter is to give a basic idea of the path integral. Applications to field theory are given in the next chapter.

10.1

Introduction

Since the underlying ideas and mathematics are very different from those of the standard canonical quantization, we go back to the basics of quantum mechanics and review some of their features to pave the way to the path integral.¹⁾

10.1.1

Bra and Ket

We define a state or Dirac's ket $|q\rangle$ as the eigenstate of a position operator $\hat{q}$, namely

$$\hat{q}|q\rangle = q|q\rangle \quad (10.1)$$

1) In this chapter we write $\hbar$ and c explicitly.

where q is the coordinate of the position; it is a c-number. In this book we use a “hat”, e.g. $\hat{A}$, to denote an operator with eigenvalue A . For any ket vector, we can define its conjugate, bra vector and a scalar (inner) product of a bra $\langle\phi|$ and a ket $|\psi\rangle$ as a pure number with the property

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^* \quad (10.2)$$

The position eigenstates define a complete orthonormal basis. Namely they satisfy

$$\langle q|q'\rangle = \delta(q - q') : \quad \text{orthonormality} \quad (10.3a)$$

$$\int dq |q\rangle\langle q| = 1 : \quad \text{completeness} \quad (10.3b)$$

Similarly, eigenstates of the momentum operator $\hat{p}$ also satisfy the orthonormality and completeness conditions

$$\begin{aligned} \hat{p}|p\rangle &= p|p\rangle \\ \langle p|p'\rangle &= 2\pi\hbar \delta(p - p') \\ \int \frac{dp}{2\pi\hbar} |p\rangle\langle p| &= 1 \end{aligned} \quad (10.4)$$

The position and momentum operators satisfy the commutation relation

$$[\hat{q}, \hat{p}] = i\hbar \quad (10.5)$$

10.1.2

Translational Operator

If we define a translational operator by

$$\hat{U}(a)|q\rangle = |q + a\rangle \quad (10.6)$$

it satisfies the relation

$$\hat{q}\hat{U}(a)|q\rangle = \hat{q}|q + a\rangle = (q + a)|q + a\rangle = \{\hat{U}(a)\hat{q} + a\hat{U}(a)\}|q\rangle \quad (10.7)$$

Since this holds for any $|q\rangle$ that are a complete orthonormal basis,

$$[\hat{q}, \hat{U}(a)] = a\hat{U}(a) \quad (10.8)$$

We learned in the previous chapter that $\hat{U}(a)$ can be expressed as

$$\hat{U}(a) = e^{-ia\hat{p}} \quad (10.9)$$

It is straightforward to show

$$\hat{U}(a)^{-1} \hat{q} \hat{U}(a) = e^{ia\hat{p}} \hat{q} e^{-ia\hat{p}} = \hat{q} + a \quad (10.10a)$$

$$\hat{U}(a) \hat{U}(b) = \hat{U}(a + b) \quad (10.10b)$$

It follows that an operator $O(\hat{q}, \hat{p})$ that can be Taylor expanded in powers of $\hat{p}$ and $\hat{q}$ is transformed by $\hat{U}(a)$ as

$$\hat{U}(a)^{-1} O(\hat{q}, \hat{p}) \hat{U}(a) = O(\hat{q} + a, \hat{p}) \quad (10.11)$$

Using the translation operator, any eigenvector of the $\hat{q}$ is expressed as

$$|q\rangle = \hat{U}(q - q_0)|q_0\rangle \quad (10.12)$$

where q_0 is the origin of the coordinate. When translational invariance holds, the origin can be placed anywhere. Without loss of generality we may put $q_0 = 0$ and the position eigenvector is expressed as

$$|q\rangle = \hat{U}(q)|0\rangle_q \quad (10.13)$$

The subscript “ q ” is attached to remind us it is a position eigenvector. Let us see the effect of $\hat{p}$ acting on the position vector:

$$\begin{aligned} |q + dq\rangle &= \left(1 - dq \frac{i}{\hbar} \hat{p}\right) |q\rangle + O(dq^2) \\ \therefore \hat{p}|q\rangle &= \lim_{dq \rightarrow 0} i\hbar \frac{|q + dq\rangle - |q\rangle}{dq} = i\hbar \frac{\partial}{\partial q} |q\rangle \end{aligned} \quad (10.14)$$

Multiplying by $\langle p|$ on both sides, we obtain

$$\langle p|\hat{p}|q\rangle = p \langle p|q\rangle = i\hbar \frac{\partial}{\partial q} \langle p|q\rangle \quad (10.15)$$

This has a solution

$$\langle p|q\rangle = N \exp\left(-\frac{i}{\hbar} p q\right) \quad (10.16)$$

To determine the constant N , we use Eq. (10.4). Inserting a complete set in the intermediate states

$$\begin{aligned} \langle p'|p\rangle &= \int dq \langle p'|q\rangle \langle q|p\rangle = \int dq N e^{-\frac{i}{\hbar} q p'} N^* e^{\frac{i}{\hbar} q p} \\ &= 2\pi\hbar \delta(p' - p) |N|^2 \\ \therefore |N| &= 1 \end{aligned} \quad (10.17)$$

Therefore, we have obtained the transformation amplitude:

$$\langle q|p\rangle = \langle p|q\rangle^* = \exp\left(\frac{i}{\hbar} q p\right) \quad (10.18)$$

Then it can easily be shown that

$$|q\rangle = \int \frac{dp}{2\pi\hbar} |p\rangle \langle p|q\rangle = \int \frac{dp}{2\pi\hbar} e^{-\frac{i}{\hbar}qp} |p\rangle \quad (10.19a)$$

$$|p\rangle = \int dq |q\rangle \langle q|p\rangle = \int dq e^{\frac{i}{\hbar}qp} |q\rangle \quad (10.19b)$$

Namely, the position basis vector and the momentum basis vector are connected to each other by Fourier transformation.

To find representations of $\hat{p}$ in the position basis, we use Eq. (10.14). Multiplying by $\langle q|$ on both sides, we obtain

$$\begin{aligned} \langle q|\hat{p}|q'\rangle &= \langle q|i\hbar \frac{\partial}{\partial q'}|q'\rangle = \left(i\hbar \frac{\partial}{\partial q'} \langle q|q'\rangle \right) = i\hbar \frac{\partial}{\partial q'} \delta(q - q') \\ &= -i\hbar \frac{\partial}{\partial q} \delta(q - q') \end{aligned} \quad (10.20)$$

Note: The operator $\hat{p}$ acts on the ket vector $|q\rangle$ and does nothing to ordinary c-numbers. Equation (10.15) is different from the one used in the Schrödinger equation, $p = -i\partial/\partial q$. This is because “bra” and “ket” are elements of abstract Hilbert space and the Schrödinger wave function is a representation of the ket vector $|\psi\rangle$ in the coordinate basis.²⁾ Defining the wave function as

$$\psi(q) = \langle q|\psi\rangle \quad (10.21)$$

The relations between the ket vectors

$$\hat{q}|q\rangle = q|q\rangle \quad (10.22a)$$

$$\hat{p}|q\rangle = i\hbar \frac{\partial}{\partial q} |q\rangle \quad (10.22b)$$

$$\langle q|q'\rangle = \delta(q - q') \quad \text{and} \quad 1 = \int dq |q\rangle \langle q| \quad (10.22c)$$

are converted to operations acting on the wave function:

$$\langle q|\hat{q} = q\langle q| \rightarrow \langle q|\hat{q}|\psi\rangle = q\langle q|\psi\rangle = q\psi(q) \quad (10.23a)$$

$$\langle q|\hat{p} = -i\hbar \frac{\partial}{\partial q} \langle q| \rightarrow \langle q|\hat{p}|\psi\rangle = -i\hbar \frac{\partial}{\partial q} \langle q|\psi\rangle = -i\hbar \frac{\partial}{\partial q} \psi(q) \quad (10.23b)$$

$$\langle \phi|\psi\rangle = \int dq \langle \phi|q\rangle \langle q|\psi\rangle = \int dq \phi(q)^* \psi(q) \quad (10.23c)$$

2) The situation is reminiscent of the vector A in Newton mechanics. Its length, direction and inner product are defined, but otherwise it is an abstract entity. A representation (A_1, A_2, A_3) can be obtained by expanding in terms of the Cartesian base vectors e_i . In the ket notation, it is written as

$|A\rangle = \sum |e_i\rangle \langle e_i|A\rangle$. The effect of an action on the basis vectors is opposite to that on the representation. For instance, the expression for an active rotation of a vector is mathematically equivalent to inverse rotation of the base vectors.

Using

$$\langle q|p\rangle = e^{ipq/\hbar} \quad (10.24)$$

$$\begin{aligned} \psi(q) &= \langle q|\psi\rangle = \langle q|\left(\int \frac{dp}{2\pi\hbar}|p\rangle\langle p|\right)|\psi\rangle = \int \frac{dp}{2\pi\hbar}\langle q|p\rangle\langle p|\psi\rangle \\ &= \int \frac{dp}{2\pi\hbar} e^{ipq/\hbar} \psi(p) \end{aligned} \quad (10.25a)$$

Similarly

$$\psi(p) = \int dq e^{-ipq/\hbar} \psi(q) \quad (10.25b)$$

Notice the sign of the exponent is different from that of the $|q\rangle$ and $|p\rangle$ Fourier transformation of Eq. (10.19).

10.2

Quantum Mechanical Equations

10.2.1

Schrödinger Equation

In terms of bras and kets, the Schrödinger equation is expressed as

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H(\hat{p}, \hat{q}) |\psi(t)\rangle \quad (10.26)$$

where the Hamiltonian H is time independent. Let us calculate the matrix element of the Hamiltonian operator in the position basis:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q}) \quad (10.27)$$

$$\begin{aligned} \langle q|\frac{\hat{p}^2}{2m}|q'\rangle &= \frac{1}{2m} \int d\gamma \langle q|\hat{p}|\gamma\rangle \langle \gamma|\hat{p}|q'\rangle \\ &= \frac{1}{2m} \int d\gamma \left[-i\hbar \frac{\partial}{\partial q} \delta(q-\gamma) \right] \left[-i\hbar \frac{\partial}{\partial \gamma} \delta(\gamma-q') \right] \\ &= -\frac{\hbar^2}{2m} \int d\gamma \frac{\partial}{\partial q} \delta(q-\gamma) \frac{\partial}{\partial \gamma} \delta(\gamma-q') \\ &= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial q^2} \delta(q-q') \\ \langle q|V(\hat{q})|q'\rangle &= V(q) \delta(q-q') \\ \therefore \langle q|\hat{H}|q'\rangle &= \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial q^2} + V(q) \right] \delta(q-q') \end{aligned} \quad (10.28)$$

In the q representation

$$i\hbar \frac{\partial}{\partial t} \langle q|\psi(t)\rangle = \int dq' \langle q|\hat{H}(\hat{p}, \hat{q})|q'\rangle \langle q'|\psi(t)\rangle \quad (10.29a)$$

Inserting Eq. (10.28)

$$i\hbar \frac{\partial}{\partial t} \psi(tq) = \hat{H}(q, -i\hbar \nabla) \psi(tq) \quad (10.29b)$$

This is the familiar Schrödinger equation one encounters in textbooks of quantum mechanics.

10.2.2

Propagators

The Schrödinger equation can be cast in an integral form. To do this we define a time evolution operator by

$$|\psi(t_1)\rangle = \begin{cases} \hat{K}(t_1, t_2) |\psi(t_2)\rangle & t_1 > t_2 \\ 0 & t_1 < t_2 \end{cases} \quad (10.30)$$

If we insert Eq. (10.30) into the Schrödinger equation (10.26) and integrate, we obtain

$$\hat{K}(t_1, t_2) = \theta(t_1 - t_2) e^{-\frac{i}{\hbar} \hat{H} t_1} e^{\frac{i}{\hbar} \hat{H} t_2} = \theta(t_1 - t_2) e^{-\frac{i}{\hbar} \hat{H} (t_1 - t_2)} \quad (10.31)$$

The last equation holds when the Hamiltonian is time independent. The time evolution operator satisfies

$$\left(i\hbar \frac{\partial}{\partial t_1} - \hat{H} \right) \hat{K}(t_1, t_2) = i\hbar \delta(t_1 - t_2) \quad (10.32)$$

The propagator is given by taking matrix elements in a given basis. In the coordinate basis defined by $\hat{q}|q\rangle = q|q\rangle$, it is written as

$$\langle q_1 | \hat{K}(t_1, t_2) | q_2 \rangle \equiv K(t_1 q_1; t_2 q_2) \quad (10.33)$$

Because $\hat{K}(t_1, t_2) \rightarrow 1$ as $t_1 \rightarrow t_2$, it satisfies the initial condition

$$\lim_{t_1 \rightarrow t_2} K(t_1 q_1; t_2 q_2) = \delta(q_1 - q_2) \quad (10.34)$$

Equation (10.32) shows that it obeys the equation

$$\left(i\hbar \frac{\partial}{\partial t_1} - \hat{H} \right) K(t_1 q_1; t_2 q_2) = i\hbar \delta(t_1 - t_2) \delta(q_1 - q_2) \quad (10.35)$$

which shows that it is the Green's function (or the propagator) for the time-dependent Schrödinger equation. We remind the reader that $\hat{K}(t_1, t_2)$ is an operator while $K(t_1 q_1; t_2 q_2)$ is a number.

The Schrödinger equation assumes all the time dependence is contained in the wave function and the operator is time independent. We can also take the attitude that all the time dependence is contained in the operator and the wave function is

time independent, which is called the Heisenberg picture. They are mathematically equivalent. The difference lies in the choice of base vectors. In the field theoretical treatment, the Heisenberg picture is generally more convenient and intuitively appealing, because equations of quantized fields are straightforward extensions of classical field equations.

So we will investigate the propagator in the Heisenberg picture. Attaching the subscripts S or H to differentiate objects in the two pictures

$$O_H(t) = e^{\frac{i}{\hbar}\hat{H}t} O_S e^{-\frac{i}{\hbar}\hat{H}t} \quad (10.36a)$$

$$|tq\rangle_H = e^{\frac{i}{\hbar}\hat{H}t}|q\rangle_S \quad (10.36b)$$

Note, in the Heisenberg picture

$$\hat{q}_H|tq\rangle_H = q|tq\rangle_H \quad (10.37)$$

Despite its apparent dependence on the time coordinate in the ket, it represents a stationary state. Therefore, it should not be interpreted as that the ket changes with time; rather it should be interpreted as that the dynamical development of an operator at time t is transformed to the state. The transition amplitude in the Heisenberg picture is given by

$$\begin{aligned} {}_H\langle t_1 q_1 | t_2 q_2 \rangle_H &= {}_S\langle q_1 | e^{-\frac{i}{\hbar}\hat{H}t_1} e^{\frac{i}{\hbar}\hat{H}t_2} | q_2 \rangle_S = {}_S\langle q_1 | \hat{K}(t_1, t_2) | q_2 \rangle_S \\ &= K(t_1 q_1; t_2 q_2) \end{aligned} \quad (10.38)$$

Equation (10.38) states that the propagator is a probability amplitude for finding the system at some point q_1 in the configuration space at time t_1 when it was originally at q_2 at time t_2 . Or stated differently, it is a matrix element of the time-ordered time evolution operator between the coordinate basis states in the Heisenberg picture.

As stated earlier, the state vector $|\psi\rangle_H = e^{\frac{i}{\hbar}\hat{H}t}|\psi(t)\rangle_S$ in the Heisenberg picture is stationary, hence the wave function $\langle q|\psi\rangle_H$ is stationary. But as Eq. (10.36b) shows, the coordinate basis is moving and ${}_H\langle tq|\psi\rangle_H$ is the same wave function as that appears in the Schrödinger equation:

$${}_H\langle tq|\psi\rangle_H = {}_S\langle q|\psi(t)\rangle_S = \psi(tq) \quad (10.39)$$

From now on, we omit the subscript H. Completeness of states enables us to write

$$\begin{aligned} \psi(t_f q_f) &= \langle t_f q_f | \psi \rangle = \int dq_i \langle t_f q_f | t_i q_i \rangle \langle t_i q_i | \psi \rangle \\ &= \int dq_i K(t_f q_f; t_i q_i) \psi(t_i q_i) \end{aligned} \quad (10.40)$$

This is Huygens' principle, that a wave at later time t_f is the sum of propagating waves from q_i at time t_i . The equation is very general and satisfies the causality condition ($K(t_f; t_i) = 0$ for $t_i > t_f$).

10.3

Feynman's Path Integral

10.3.1

Sum over History

To calculate the propagator, we first divide the continuous time interval into N small but finite and discrete intervals

$$t_j = t_0 + j \varepsilon, \quad \varepsilon = \frac{t_f - t_i}{N}, \quad t_0 = t_i, \quad t_N = t_f$$

$$\exp \left[-\frac{i}{\hbar} \hat{H}(t_f - t_i) \right] = \prod_{j=1}^N \exp \left[-\frac{i}{\hbar} \hat{H}(t_j - t_{j-1}) \right] \quad (10.41)$$

and insert complete sets of coordinate bases in the intermediate time point (see Fig. 10.1).

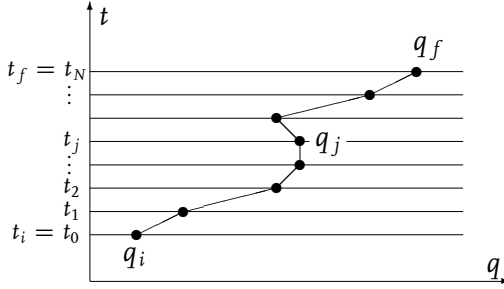


Figure 10.1 To perform the path integral, the time coordinate is discretized to N intervals between t_i and t_f . At each $t = t_j$, q_j is integrated from $-\infty$ to ∞ . q_i and q_f are fixed. $\int \mathcal{D}q \equiv \lim_{N \rightarrow \infty} \int dq_1 \cdots dq_{N-1}$.

$$\langle t_f q_f | t_i q_i \rangle = \int \cdots \int dq_1 dq_2 \cdots dq_{N-1} \langle t_f q_f | t_{N-1}, q_{N-1} \rangle$$

$$\cdots \langle t_j, q_j | t_{j-1}, q_{j-1} \rangle \cdots \langle t_1 q_1 | t_i q_i \rangle \quad (10.42)$$

Referring to Eq. (10.38), we can rewrite each intermediate inner product as

$$\langle t_j q_j | t_{j-1} q_{j-1} \rangle = \langle q_j | e^{-\frac{i}{\hbar} \hat{H}(t_j - t_{j-1})} | q_{j-1} \rangle = \langle q_j | e^{-\frac{i}{\hbar} \varepsilon \hat{H}} | q_{j-1} \rangle \quad (10.43)$$

If the Hamiltonian is of the form

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q}) \quad (10.44)$$

$$\begin{aligned}
\langle q_j | V(\hat{q}) | q_{j-1} \rangle &= V(q_{j-1}) \langle q_j | q_{j-1} \rangle = \int \frac{dp}{2\pi\hbar} \langle q_j | p \rangle \langle p | q_{j-1} \rangle V(q_{j-1}) \\
\langle q_j | \frac{\hat{p}^2}{2m} | q_{j-1} \rangle &= \int \frac{dp'}{2\pi\hbar} \frac{dp}{2\pi\hbar} \langle q_j | p' \rangle \langle p' | \frac{\hat{p}^2}{2m} | p \rangle \langle p | q_{j-1} \rangle \\
&= \int \frac{dp'}{2\pi\hbar} \frac{dp}{2\pi\hbar} e^{\frac{i}{\hbar} p' q_j} e^{-\frac{i}{\hbar} p q_{j-1}} (2\pi\hbar) \delta(p - p') \frac{p^2}{2m} \\
&= \int \frac{dp}{2\pi\hbar} \langle q_j | p \rangle \langle p | q_{j-1} \rangle \frac{p^2}{2m} \quad (10.45)
\end{aligned}$$

$$\therefore \langle q_j | H(\hat{p}, \hat{q}) | q_{j-1} \rangle = \int \frac{dp}{2\pi\hbar} \langle q_j | p \rangle \langle p | q_{j-1} \rangle H(p, q_{j-1}) \quad (10.46)$$

The Hamiltonian is no longer an operator but is written in terms of classical numbers. Generally, the Hamiltonian may include a polynomial of the product of operators $\hat{q}$, $\hat{p}$, as, for instance, the magnetic force on the charged particle. In such cases, a certain subtlety arises. If all $\hat{p}$'s stand to the left of the coordinate operators, we can insert the complete momentum basis to the left of the Hamiltonian and obtain

$$\begin{aligned}
\langle q_j | e^{-\frac{i}{\hbar} \varepsilon H(\hat{p}, \hat{q})} | q_{j-1} \rangle &\simeq \int \frac{dp}{2\pi\hbar} \langle q_j | p \rangle \langle p | \left[1 - \frac{i}{\hbar} \varepsilon H(\hat{p}, \hat{q}) \right] | q_{j-1} \rangle \\
&= \int \frac{dp}{2\pi\hbar} \langle q_j | p \rangle \langle p | q_{j-1} \rangle \left[1 - \frac{i}{\hbar} \varepsilon H(p, q_{j-1}) \right] \\
&\simeq \int \frac{dp}{2\pi\hbar} \langle q_j | p \rangle \langle p | q_{j-1} \rangle e^{-\frac{i}{\hbar} \varepsilon H(p, q_{j-1})} \quad (10.47)
\end{aligned}$$

which has the same form as Eq. (10.46). On the other hand, if all $\hat{p}$'s stand to the right of the coordinates,

$$\begin{aligned}
\langle q_j | e^{-\frac{i}{\hbar} \varepsilon H(\hat{p}, \hat{q})} | q_{j-1} \rangle &\simeq \int \frac{dp}{2\pi\hbar} \langle q_j | \left[1 - \frac{i}{\hbar} \varepsilon H(\hat{p}, \hat{q}) \right] | p \rangle \langle p | q_{j-1} \rangle \\
&= \int \frac{dp}{2\pi\hbar} \langle q_j | p \rangle \langle p | q_{j-1} \rangle \left[1 - \frac{i}{\hbar} \varepsilon H(p, q_j) \right] \\
&\simeq \int \frac{dp}{2\pi\hbar} \langle q_j | p \rangle \langle p | q_{j-1} \rangle e^{-\frac{i}{\hbar} \varepsilon H(p, q_j)} \quad (10.48)
\end{aligned}$$

As there is no well-defined way to order the product of operators, we take a midpoint.³⁾ Putting $\langle q_j | p \rangle = e^{ipq_j}$, we obtain

$$\langle t_j q_j | t_{j-1} q_{j-1} \rangle = \int \frac{dp}{2\pi\hbar} \exp \left[i p (q_j - q_{j-1}) - \frac{i}{\hbar} \varepsilon H \left(p, \frac{q_j + q_{j-1}}{2} \right) \right] \quad (10.49)$$

3) For a special ordering called Weyl ordering, where the ordering of operators is symmetrized, like $p^2 q = \frac{1}{3}(p^2 q + p q p + q p^2)$, the midpoint method is exact, and this is the conventional assumption. See [113, 255] for more detailed treatments.

Writing all the inner products similarly, we obtain

$$\begin{aligned} \langle t_f q_f | t_i q_i \rangle &= \lim_{\varepsilon \rightarrow 0} \int \cdots \int dq_1 \cdots dq_{N-1} \frac{dp_1}{2\pi\hbar} \cdots \frac{dp_N}{2\pi\hbar} \\ &\quad \times \exp \left[\frac{i}{\hbar} \sum_{j=1}^N \left\{ p_j (q_j - q_{j-1}) - \varepsilon H \left(p_j, \frac{q_j + q_{j-1}}{2} \right) \right\} \right] \end{aligned} \quad (10.50)$$

Note that there are N inner products, hence N integrals over the momenta, but there are only $N-1$ intermediate states, hence $N-1$ integrals over the coordinates. In the continuum limit of $\varepsilon \rightarrow 0$, the exponent can be rewritten as

$$\begin{aligned} &\lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \frac{i}{\hbar} \sum_{j=1}^N \left\{ p_j (q_j - q_{j-1}) - \varepsilon H \left(p_j, \frac{q_j + q_{j-1}}{2} \right) \right\} \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \frac{i}{\hbar} \varepsilon \sum_{j=1}^N \left\{ p_j \left(\frac{q_j - q_{j-1}}{\varepsilon} \right) - H \left(p_j, \frac{q_j + q_{j-1}}{2} \right) \right\} \quad (10.51) \\ &= \frac{i}{\hbar} \int_{t_i}^{t_f} dt [p \dot{q} - H(p, q)] = \frac{i}{\hbar} \int_{t_i}^{t_f} dt L = \frac{i}{\hbar} S[q] \end{aligned}$$

where $S[q]$ is the classical action. Therefore, the propagator may be symbolically written as

$$\langle t_f q_f | t_i q_i \rangle = \int \mathcal{D}q \mathcal{D} \left(\frac{p}{2\pi\hbar} \right) \exp \left[\frac{i}{\hbar} \int_{t_i}^{t_f} dt \{ p \dot{q} - H(p, q) \} \right] \quad (10.52)$$

This is the path integral expression for the propagator from $(t_i q_i)$ to $(t_f q_f)$. In the continuum limit q becomes a function of t , and the integral is a functional integral. Each function $q(t)$ and $p(t)$ defines a path in phase space. The integration is over all functions and is infinite dimensional. An outstanding characteristic of the path integral is that the quantum operators $\hat{q}$ and $\hat{p}$ do not appear, but only their representations expressed in ordinary numbers. It is not obvious, however, that the infinite-dimensional integrals of this type are well-defined mathematically, i.e. if they converge. We proceed assuming they do.

If the Hamiltonian has a special form, namely quadratic in the momenta, such as

$$H(p, q) = \frac{p^2}{2m} + V(q) \quad (10.53)$$

we can perform the p -integration:

$$\begin{aligned}
 & \int \frac{dp_j}{2\pi\hbar} \exp \left[\frac{i}{\hbar} \varepsilon \left(\frac{p_j(q_j - q_{j-1})}{\varepsilon} - \frac{p_j^2}{2m} \right) \right] \\
 &= \int \frac{dp_j}{2\pi\hbar} \exp \left[-\frac{i\varepsilon}{2m\hbar} \left\{ \left(p_j - \frac{m(q_j - q_{j-1})}{\varepsilon} \right)^2 - \left(\frac{m(q_j - q_{j-1})}{\varepsilon} \right)^2 \right\} \right] \\
 &= \frac{1}{2\pi\hbar} \left(\frac{2\pi m\hbar}{i\varepsilon} \right)^{1/2} \exp \left[\frac{im\varepsilon}{2\hbar} \left(\frac{q_j - q_{j-1}}{\varepsilon} \right)^2 \right] \\
 &= \left(\frac{m}{2\pi i\hbar\varepsilon} \right)^{1/2} \exp \left[\frac{im\varepsilon}{2\hbar} \left(\frac{q_j - q_{j-1}}{\varepsilon} \right)^2 \right] \quad 4)
 \end{aligned} \tag{10.54}$$

Substituting this expression into Eq. (10.50), we obtain

$$\begin{aligned}
 \langle t_f q_f | t_i q_i \rangle &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i\hbar\varepsilon} \right)^{N/2} \int \cdots \int dq_1 \cdots dq_{N-1} \\
 &\quad \times \exp \left[\frac{i}{\hbar} \varepsilon \sum_{j=1}^N \left\{ \frac{m}{2} \left(\frac{q_j - q_{j-1}}{\varepsilon} \right)^2 - V \left(\frac{q_j + q_{j-1}}{2} \right) \right\} \right] \\
 &= C \int \mathcal{D}q \exp \left[\frac{i}{\hbar} \int dt \left\{ \frac{1}{2} m \dot{q}^2 - V(q) \right\} \right] \\
 &= C \int \mathcal{D}q e^{\frac{i}{\hbar} S[q]}
 \end{aligned} \tag{10.56}$$

where C is a constant independent of the dynamics of the system and $S[q]$ is the action of the system. This is the Feynman path integral in an alternative form. Although it is proved only for a special form of the Hamiltonian, the use of the Lagrangian has many desirable features. The action is a Lorentz scalar and the symmetry structure of the system is very easy to grasp. Besides it has the classical path as the first approximation, so the transition from classical mechanics to quantum mechanics is transparent in this formalism. Therefore, we will use this as the starting point from which any equation of motion is derived. The propagator is expressed as a sum of all the path $[q(t)]$ configurations connecting $t_i q_i$ with $t_f q_f$ with the weight $e^{(i/\hbar)S[q]}$. Classically, the path is determined as the one to minimize the action, but, in quantum mechanics, any path connecting the end points is possible, though the action of a path far from the classical one generally gives a large

- 4) The Gaussian integral can be extended by analytically continuing $|A|$ to $|A|e^{i\theta}$ ($-\pi/2 \leq \theta \leq \pi/2$) in

$$\int dq e^{-|A|q^2} = \sqrt{\frac{\pi}{|A|}} \rightarrow \int_{-\infty}^{\infty} dq e^{\mp i|A|q^2} = \sqrt{\mp \frac{\pi i}{|A|}} \tag{10.55a}$$

This is a variation of the Fresnel integrals

$$\int_0^{\infty} dx \sin x^2 = \int_0^{\infty} dx \cos x^2 = \frac{1}{2} \sqrt{\frac{\pi}{2}} \tag{10.55b}$$

value relative to $\hbar$, hence a small weight because of the oscillatory nature of the factor $e^{(i/\hbar)S[q]}$. If there is more than one path, as in the simple and familiar two-slit experiment, both paths contribute and the interference is an important aspect of quantum mechanical treatment. If the Hamiltonian is not of the type Eq. (10.53), we may still use Eq. (10.56), but in that case we have to be careful in defining the path integral measure $\mathcal{D}q$.

10.3.2

Equivalence with the Schrödinger Equation

From the form of the path integral, it is hard to think there is any connection with the Schrödinger equation, but we have developed the formalism starting from the Schrödinger equation. Both forms are known to reproduce the quantum mechanical phenomena correctly. We show here that the path integral includes the Schrödinger equation. The latter is a differential equation, so we examine an infinitesimal form of the path integral. For a small time difference $t_f = t_i + \varepsilon$, the propagator in Eq. (10.56) is expressed as

$$\langle t_f = t + \varepsilon, q_f | t q_i \rangle = \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{1/2} \exp \left[\frac{i}{\hbar} \varepsilon \left\{ \frac{m}{2} \left(\frac{q_f - q_i}{\varepsilon} \right)^2 - V \left(\frac{q_f + q_i}{2} \right) \right\} \right] \quad (10.57)$$

As the propagator describes the propagation of a wave function and satisfies

$$\psi(t + \varepsilon, q) = \int_{-\infty}^{\infty} d\gamma K(t + \varepsilon, q; t\gamma) \psi(t, \gamma) \quad (10.58)$$

we can insert Eq. (10.57) in Eq. (10.58) and obtain

$$\psi(t + \varepsilon, q) = \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{1/2} \int d\gamma \exp \left[\frac{i}{\hbar} \varepsilon \left\{ \frac{m}{2} \left(\frac{q - \gamma}{\varepsilon} \right)^2 - V \left(\frac{q + \gamma}{2} \right) \right\} \right] \psi(t, \gamma) \quad (10.59)$$

Changing the integration variable to $z = \gamma - q$, we can rewrite the expression as

$$\begin{aligned} \psi(t + \varepsilon, q) &= \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{1/2} \int dz \exp \left(\frac{im}{2\hbar \varepsilon} z^2 \right) \exp \left[-\frac{i}{\hbar} \varepsilon V \left(q + \frac{z}{2} \right) \right] \\ &\quad \times \psi(t, q + z) \end{aligned} \quad (10.60)$$

As ε is a small number, the dominant contribution will come from the region where

$$\left| \frac{m}{2\hbar \varepsilon} z^2 \right| \lesssim 1 \quad (10.61)$$

In this region the first term is of the order of 1. We can expand the second and the third terms in ε . Note each power of z^n gives a factor of $\varepsilon^{n/2}$. Then

$$\begin{aligned}
 & \psi(t + \varepsilon, q) \\
 &= \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{1/2} \int dz \exp \left(\frac{i m}{2 \hbar \varepsilon} z^2 \right) \left[1 - \frac{i}{\hbar} \varepsilon V(q) + O(\varepsilon z) \right] \\
 & \quad \times \left[\psi(tq) + z \psi'(tq) + \frac{z^2}{2} \psi''(tq) + O(z^3) \right] \\
 &= \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{1/2} \int dz \exp \left(\frac{i m}{2 \hbar \varepsilon} z^2 \right) \left[\psi(tq) \left(1 - \frac{i \varepsilon}{\hbar} V(q) \right) \right. \\
 & \quad \left. + z \psi'(tq) + \frac{z^2}{2} \psi''(tq) + O(z^3, \varepsilon z) \right] \tag{10.62}
 \end{aligned}$$

Each term can easily be integrated and the results are

$$\int_{-\infty}^{\infty} dz \exp \left(\frac{i m}{2 \hbar \varepsilon} z^2 \right) = \left(\frac{2\pi i \hbar \varepsilon}{m} \right)^{1/2} \tag{10.63a}$$

$$\int_{-\infty}^{\infty} dz z \exp \left(\frac{i m}{2 \hbar \varepsilon} z^2 \right) = 0 \tag{10.63b}$$

$$\int_{-\infty}^{\infty} dz z^2 \exp \left(\frac{i m}{2 \hbar \varepsilon} z^2 \right) = \frac{i \hbar \varepsilon}{m} \left(\frac{2\pi i \hbar \varepsilon}{m} \right)^{1/2} \tag{10.63c}$$

Substituting the above equation back into Eq. (10.62), we obtain

$$\psi(t + \varepsilon, q) = \psi(tq) + \frac{i \hbar \varepsilon}{2m} \psi''(tq) - \frac{i \varepsilon}{\hbar} V(q) \psi(tq) + O(\varepsilon^2) \tag{10.64}$$

In the limit $\varepsilon \rightarrow 0$, we obtain

$$i \hbar \frac{\partial \psi(tq)}{\partial t} = \left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial q^2} + V(q) \right) \psi(tq) \tag{10.65}$$

Therefore the path integral contains the Schrödinger equation.

10.3.3

Functional Calculus

The path integral is a functional integral, which uses a type of mathematics readers may not have encountered. Perhaps it is useful to introduce basic operations of the functional and to evaluate simple examples for future applications. The simplest functional we know is the action functional defined by

$$S[q] = \int_{t_i}^{t_f} dt L(q, \dot{q}) \tag{10.66}$$

Unlike a function, which gives a value for a particular point q [or $q = (q_1, \dots, q_N)$], the functional gives a value for an entire trajectory $q = q(t)$ along which the integration is carried out. If the trajectory connecting $q_i = q(t_i)$ and $q_f = q(t_f)$

is different, the value of the functional is different. Thus the functional has the generic form

$$F[f] = \int dq F(f(q)) \quad (10.67)$$

Traditionally square brackets are used to denote the functional. Some people call $F(f(q))$ a functional, meaning a function of functions. Note that, unlike our definition Eq. (10.67), which maps a continuous set of variables to one number, it maps to another continuous set of numbers. The notion of derivatives is also used in the functional. We define the functional derivative by

$$\frac{\delta F[f(q)]}{\delta f(y)} = \lim_{\varepsilon \rightarrow 0} \frac{F[f(q) + \varepsilon \delta(q - y)] - F[f(q)]}{\varepsilon} \quad (10.68)$$

It follows from the above equation

$$\frac{\delta f(q)}{\delta f(y)} = \delta(q - y) \quad (10.69)$$

Problem 10.1

Prove that if

$$F[f] = \int dy F(f(y)) = \int dy \{f(y)\}^n \quad (10.70)$$

then $\frac{\delta F[f]}{\delta f(q)} = n \{f(q)\}^{n-1}$

Problem 10.2

Prove that

$$\frac{\delta S}{\delta q(t')} = -\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right)_{t=t'} + \frac{\partial L}{\partial q} \Big|_{t=t'} \quad (10.71)$$

where S is the action given by Eq. (10.66).

Problem 10.3

Define

$$F_x[f] = \int dy G(x, y) f(y) \quad (10.72)$$

Here, x is to be regarded as a parameter. Prove

$$\frac{\delta F_x[f]}{\delta f(z)} = G(x, z) \quad (10.73)$$

We can make a Taylor expansion of the functional, too:

$$\begin{aligned}
 S[q] = S[q_c + \eta] &= S[q_c] + \int dt \eta(t) \left. \frac{\delta S[q]}{\delta q(t)} \right|_{q=q_c} \\
 &+ \frac{1}{2!} \int dt_1 dt_2 \eta(t_1) \eta(t_2) \left. \frac{\delta^2 S[q]}{\delta q(t_1) \delta q(t_2)} \right|_{q=q_c} + \dots
 \end{aligned} \quad (10.74)$$

The form of the Taylor expansion Eq. (10.74) can be understood most easily if we divide the time integral into N segments (see Fig. 10.2):

$$\begin{aligned}
 S[q] &= \int_{t_i}^{t_f} dt L(q(t), \dots) = \sum_{j=1}^N \varepsilon_j L(q_j, \dots) \\
 &= F[q_0, q_1, q_2, \dots, q_{N-1}, q_N], \quad q_0, q_N \text{ fixed}
 \end{aligned} \quad (10.75)$$

In this form, the functional becomes an ordinary multivariable function:

$$(q_0, q_1, q_2, \dots, q_j, \dots, q_N)$$

Then the Taylor expansion of the functional can be carried out using the ordinary Taylor expansion:

$$\begin{aligned}
 F[q + \eta] &= F(q_0 + \eta_0, q_1 + \eta_1, \dots, q_N + \eta_N) \Big|_{\eta_0 = \eta_N = 0} \\
 &= F[q] + \sum_{n=0}^{\infty} \sum_{j_1=1}^N \dots \sum_{j_n=1}^N \frac{1}{n!} T_n(j_1, \dots, j_n) \eta_{j_1} \dots \eta_{j_n}
 \end{aligned} \quad (10.76)$$

where

$$T_n(j_1, \dots, j_n) = \left. \frac{\partial^n F[q]}{\partial q_{j_1} \dots \partial q_{j_n}} \right|_{\eta=0} \quad (10.77)$$

Going back to continuous variables

$$j_k \rightarrow t_k, \quad \eta_{j_k} (j_k = 1, \dots, N) \rightarrow \eta(t_k), \quad \sum_{j_k} \rightarrow \int dt_k, \quad t_i \leq t_k \leq t_f \quad (10.78)$$

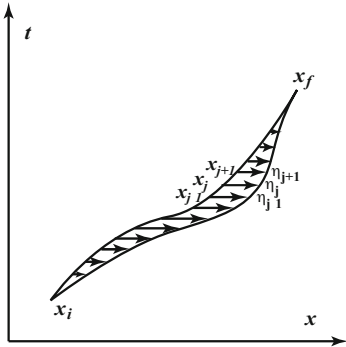


Figure 10.2 Functional differential path.

the second term in Eq. (10.76) with $n = 1$ becomes

$$\sum_k \frac{\partial F}{\partial q_k} \eta_k = \sum_k \varepsilon_k \left[\frac{1}{\varepsilon_k} \frac{\partial F}{\partial q_k} \right] \eta_k = \int dt' \eta(t') \frac{\delta S[q]}{\delta q(t')} \quad (10.79)$$

Note, in the second equality, $(1/\varepsilon)\partial F/\partial q_k$ is well behaved because of Eq. (10.75) and gives the second term in Eq. (10.74). Other higher order terms can be derived similarly.

10.4

Propagators: Simple Examples

Here we calculate the propagator in some simple examples to gain familiarity with the path integral method.

10.4.1

Free-Particle Propagator

The free particle is the simplest example. In fact, the propagator can be obtained easily without depending on the path integral. For the sake of later applications, we derive the result in two ways.

(1) Canonical Method The definition of the transition amplitude Eq. (10.38) gives

$$K_{\text{free}}(f, i) \equiv K_0(t_f q_f; t_i q_i) = \langle t_f q_f | t_i q_i \rangle_{H=H_0} = \langle q_f | e^{-\frac{i}{\hbar} H_0(t_f - t_i)} | q_i \rangle \quad (10.80)$$

The Hamiltonian of the free particle is given by

$$H_0 = \frac{\hat{p}^2}{2m} \quad (10.81)$$

It does not contain coordinate operators, therefore it can be diagonalized by simply inserting the momentum basis in the intermediate states:

$$\begin{aligned} K_0(t_f q_f; t_i q_i) &= \theta(t_f - t_i) \int \frac{dp'}{2\pi\hbar} \frac{dp}{2\pi\hbar} \langle q_f | p' \rangle \langle p' | e^{-\frac{i}{\hbar} H_0(t_f - t_i)} | p \rangle \langle p | q_i \rangle \\ &= \int \frac{dp}{2\pi\hbar} \int \frac{dp'}{2\pi\hbar} \exp \left[\frac{i}{\hbar} (p' q_f - p q_i) \right] \\ &\quad \times \left[\theta(t_f - t_i) 2\pi\hbar \delta(p - p') \exp \left(-\frac{i}{\hbar} \frac{p^2}{2m} (t_f - t_i) \right) \right] \\ &= \theta(t_f - t_i) \int \frac{dp}{2\pi\hbar} \exp \left[\frac{i}{\hbar} \left\{ p(q_f - q_i) - \frac{p^2}{2m} (t_f - t_i) \right\} \right] \\ &= \theta(t_f - t_i) \sqrt{\frac{m}{2\pi i \hbar (t_f - t_i)}} \exp \left[\frac{i}{\hbar} \frac{m(q_f - q_i)^2}{2(t_f - t_i)} \right] \end{aligned} \quad (10.82)$$

where we have used $\int dz e^{-Az^2} = \sqrt{\pi/A}$ and its analytical continuation to $A = |A|e^{i\theta}$, ($\pi/2 \leq \theta \leq -\pi/2$). It expresses the well known fact that a localized wave packet spreads with time.

It is often more convenient to work in momentum space. The boundary condition $K = 0$ for $t - t' < 0$ can be explicitly taken into account in the momentum representation. Using the integral representation for the step function

$$\theta(t) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{dz e^{izt}}{z - i\varepsilon} \quad (10.83)$$

and inserting it into Eq. (10.82), we obtain

$$K_0(tq : t'q') = \int \frac{dp}{2\pi\hbar} \frac{d\omega}{2\pi i} \frac{1}{\omega - i\varepsilon} \times \exp \left\{ \frac{i}{\hbar} \left[p(q - q') - \left(\frac{p^2}{2m} - \hbar\omega \right) (t - t') \right] \right\} \quad (10.84)$$

In this form, the energy no longer takes its on-shell value $p^2/2m$, but any value $E = p^2/2m - \hbar\omega$. So we use the energy variable to express the propagator

$$K_0(tq : t'q') = \int \frac{dp}{2\pi\hbar} \frac{dE}{2\pi\hbar} e^{\frac{i}{\hbar}[p(q-q') - E(t-t')]} \frac{i\hbar}{E - (p^2/2m) + i\varepsilon} \quad (10.85)$$

This is the momentum representation of the free propagator. In field theory, we will encounter its relativistic counterpart.

Note: Classically, the exponent of the free propagator can be expressed as $(i/\hbar)(pq/2)(p = m(q_f - q_i)/(t_f - t_i) = m\dot{q}, q = q_f - q_i)$, but quantum mechanically, the momentum and the position cannot be specified simultaneously. The action has to be written in terms of the initial and final coordinates and the elapsed time only, as expressed in the last equation of Eq. (10.82). Let us compare the result with the classical relation. The solution to the action

$$S[q] = \int_{t_i}^{t_f} dt \frac{m}{2} \dot{q}^2 \quad (10.86)$$

is given by solving the Euler–Lagrange equation, which is written as

$$m\ddot{q} = 0 \rightarrow \dot{q} = v = \text{constant} \quad (10.87)$$

Therefore, the classical action gives

$$S[q_{cl}] = \int_{t_i}^{t_f} dt \frac{1}{2} \dot{q}^2 = \frac{m}{2} v^2 (t_f - t_i) = \frac{m}{2} \frac{(q_f - q_i)^2}{t_f - t_i} \quad (10.88)$$

Thus the propagator is written as

$$K_{\text{free}}(f, i) = \left[\frac{m}{2\pi i\hbar(t_f - t_i)} \right]^{1/2} e^{\frac{i}{\hbar} S_{cl}} \quad (10.89)$$

This is a special example because the quantum solution is completely determined by the classical path.

(2) Gaussian Integral The second method is a brute-force way, but provides a good exercise for path integral calculation. Since the Hamiltonian contains the momentum in quadratic form, we can use Eq. (10.56) for the path integral calculation. The action is given by Eq. (10.86). Then the propagator is calculated to give

$$\begin{aligned}
 K_{\text{free}}(f, i) &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{N/2} \int dq_1 \cdots dq_{N-1} \exp \left[\frac{i}{\hbar} \varepsilon \sum_{j=1}^N \frac{m}{2} \left(\frac{q_j - q_{j-1}}{\varepsilon} \right)^2 \right] \\
 &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{N/2} \int dq_1 \cdots dq_{N-1} \exp \left[\frac{im}{2\hbar \varepsilon} \sum_{j=1}^N (q_j - q_{j-1})^2 \right]
 \end{aligned} \tag{10.90}$$

Using $\int dz e^{-Az^2} = \sqrt{\pi/A}$ with its analytical continuation to $A = |A|e^{i\theta}$ ($-\pi/2 \leq \theta \leq \pi/2$), we can evaluate the integral. Changing the integration value to

$$y_j = \left(\frac{m}{2i\hbar \varepsilon} \right)^{1/2} q_j \tag{10.91}$$

we have

$$\begin{aligned}
 K_{\text{free}}(f, i) &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{N/2} \left(\frac{2i\hbar \varepsilon}{m} \right)^{(N-1)/2} \\
 &\quad \times \int dy_1 \cdots dy_{N-1} \exp \left[- \sum_{j=1}^N (y_j - y_{j-1})^2 \right]
 \end{aligned} \tag{10.92}$$

Let us integrate the j th integral and see what relations exist between different j 's:

$$\begin{aligned}
 I_j &\equiv \int dy_j dy_{j+1} \exp \left[- \left\{ (y_j - y_{j-1})^2 + (y_{j+1} - y_j)^2 \right. \right. \\
 &\quad \left. \left. + (y_{j+2} - y_{j+1})^2 \right\} \right] \\
 &= \int dy_j dy_{j+1} \exp \left[- \left\{ 2 \left(y_j - \frac{y_{j+1} + y_{j-1}}{2} \right)^2 \right. \right. \\
 &\quad \left. \left. + \frac{1}{2} (y_{j+1} - y_{j-1})^2 + (y_{j+2} - y_{j+1})^2 \right\} \right]
 \end{aligned}$$

Integration over y_j will take I_j to the form

$$\begin{aligned}
 I_j &= \left(\frac{\pi}{2} \right)^{1/2} \int dy_{j+1} \exp \left[- \left\{ \frac{3}{2} \left(y_{j+1} - \frac{2y_{j+2} + y_{j-1}}{3} \right)^2 \right. \right. \\
 &\quad \left. \left. + \frac{1}{3} (y_{j+2} - y_{j-1})^2 \right\} \right] \\
 &= \left(\frac{\pi}{2} \right)^{1/2} \left(\frac{2\pi}{3} \right)^{1/2} \exp \left[- \frac{1}{3} (y_{j+2} - y_{j-1})^2 \right]
 \end{aligned} \tag{10.93}$$

From the above calculation, we see that repeating the integration j times starting from $j = 1$ gives

$$\begin{aligned} \int \cdots \int d\gamma_1 \cdots d\gamma_j &= \left(\frac{\pi}{2}\right)^{1/2} \left(\frac{2\pi}{3}\right)^{1/2} \cdots \left(\frac{j\pi}{j+1}\right)^{1/2} \\ &\quad \times \exp\left[-\frac{1}{j+1}(\gamma_{j+1} - \gamma_0)^2\right] \\ &= \left[\frac{(\pi)^j}{j+1}\right]^{1/2} \exp\left[-\frac{1}{j+1}(\gamma_{j+1} - \gamma_0)^2\right] \end{aligned} \quad (10.94)$$

Inserting Eq. (10.94) into Eq. (10.92), we obtain

$$\begin{aligned} K_{\text{free}}(f, i) &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon}\right)^{N/2} \left(\frac{2i\hbar\varepsilon}{m}\right)^{(N-1)/2} \\ &\quad \times \left[\left(\frac{(\pi)^{N-1}}{N}\right)^{1/2} \exp\left\{-\frac{1}{N}(\gamma_N - \gamma_0)^2\right\}\right] \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon N}\right)^{1/2} \exp\left[\frac{im}{2\hbar N \varepsilon}(q_N - q_0)^2\right] \\ &= \left[\frac{m}{2\pi i \hbar(t_f - t_i)}\right]^{1/2} \exp\left[\frac{i}{\hbar} \frac{m(q_f - q_i)^2}{2(t_f - t_i)}\right] \end{aligned} \quad (10.95)$$

where we have used $N\varepsilon = t_f - t_i$, $q_0 = q_i$, $q_N = q_f$ in the last equation. Note all the potentially singular terms have disappeared from the expression. In the limit $t_f \rightarrow t_i$

$$K_{\text{free}}(f, i) = \langle t_f q_f | t_i q_i \rangle \xrightarrow{t_f \rightarrow t_i} \delta(q_f - q_i) \quad (10.96)$$

which is nothing but the orthonormality relation in the coordinate basis.

10.4.2

Harmonic Oscillator

Next we consider the one-dimensional harmonic oscillator. This is an example that can also be solved exactly. In fact, the method and conclusion are valid for any action that contains variables up to quadratic terms, and thus have very general uses. In the following, we consider a harmonic oscillator interacting with an external source described by the Lagrangian

$$L = \frac{1}{2} m \dot{q}^2 - \frac{1}{2} m \omega^2 q^2 + Jq \quad (10.97)$$

The time-dependent external source $J(t)$ can be thought of as an electric field if the oscillator carries an electric charge. The well-known result for the free harmonic oscillator can be obtained from this system in the limit $J(t) \rightarrow 0$. Furthermore,

if the source is time independent, the problem can also be solved exactly. In this case the system is nothing but an oscillator with its equilibrium position displaced, because the Lagrangian can be rewritten to give

$$L = \frac{1}{2}m\dot{\gamma}^2 - \frac{1}{2}m\omega^2\gamma^2 + \frac{J^2}{2m\omega^2}, \quad \gamma = q - \frac{J}{m\omega^2} \quad (10.98)$$

This is the case, for instance, of a suspended spring under the effect of gravity.

The Euler–Lagrange equation for the action Eq. (10.97) is determined by the requirement

$$\left. \frac{\delta S[q]}{\delta q(t)} \right|_{q=q_{\text{cl}}} = 0 \quad (10.99)$$

which gives the equation of motion to determine the classical path $q_{\text{cl}}(t)$.

$$m\ddot{q}_{\text{cl}} + m\omega^2 q_{\text{cl}} = J \quad (10.100)$$

In quantum mechanics, $q(t)$ can take any path, but it is convenient to expand it around the classical path. Define the quantum fluctuation path $\eta(t)$ by

$$q(t) = q_{\text{cl}}(t) + \eta(t) \quad (10.101)$$

Because of its quadratic nature, the Taylor expansion of the action terminates at the second power of $\eta(t)$

$$\begin{aligned} S[q] &= S[q_{\text{cl}} + \eta] \\ &= S[q_{\text{cl}}] + \int dt \eta(t) \left. \frac{\delta S[q]}{\delta q(t)} \right|_{q=q_{\text{cl}}} \\ &\quad + \frac{1}{2!} \int dt_1 dt_2 \eta(t_1) \eta(t_2) \left. \frac{\delta^2 S[q]}{\delta q(t_1) \delta q(t_2)} \right|_{q=q_{\text{cl}}} \end{aligned} \quad (10.102)$$

The second term in Eq. (10.102) vanishes because the classical path is where the action is at its extremum. Then

$$S[q] = S[q_{\text{cl}}] + \frac{1}{2!} \int dt_1 dt_2 \eta(t_1) \eta(t_2) \left. \frac{\delta^2 S[q]}{\delta q(t_1) \delta q(t_2)} \right|_{q=q_{\text{cl}}} \quad (10.103)$$

Calculating the functional derivatives, this becomes

$$S[q] = S[q_{\text{cl}} + \eta] = S[q_{\text{cl}}] + \frac{m}{2} \int_{t_i}^{t_f} dt [\dot{\eta}(t)^2 - \omega^2 \eta^2] \quad (10.104)$$

The variable $\eta(t)$ represents the quantum fluctuations around the classical path. As the end points of the trajectories are fixed, they satisfy the boundary condition

$$\eta(t_i) = \eta(t_f) = 0 \quad (10.105)$$

As the harmonic oscillator is an important example, we calculate the propagator again in two ways.

(1) Matrix Method

Summing over all the paths is equivalent to summing over all possible fluctuations subject to constraints Eqs. (10.105). The propagator is expressed as

$$\begin{aligned} K_{\text{osci}}(f, i) &\equiv \langle t_f q_f | t_i q_i \rangle |_{\text{harmonic oscillator}} \\ &= C \int \mathcal{D}q \, e^{\frac{i}{\hbar} S[q]} = C e^{\frac{i}{\hbar} S[q_{\text{cl}}]} \int \mathcal{D}\eta \\ &\quad \times \exp \left[\frac{i}{\hbar} \left\{ \frac{m}{2} \int dt \{ \dot{\eta}(t)^2 - \omega^2 \eta^2 \} \right\} \right] \end{aligned} \quad (10.106a)$$

$$\begin{aligned} &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{N/2} e^{\frac{i}{\hbar} S[q_{\text{cl}}]} \int_{-\infty}^{\infty} d\eta_1 \cdots d\eta_{N-1} \\ &\quad \times \exp \left[\frac{im}{2\hbar} \varepsilon \sum_{j=1}^N \left\{ \left(\frac{\eta_j - \eta_{j-1}}{\varepsilon} \right)^2 - \omega^2 \left(\frac{\eta_j + \eta_{j-1}}{2} \right)^2 \right\} \right] \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{N/2} \left(\frac{2i\hbar\varepsilon}{m} \right)^{(N-1)/2} e^{\frac{i}{\hbar} S[q_{\text{cl}}]} \int d\gamma_1 \cdots d\gamma_{N-1} \\ &\quad \times \exp \left[- \sum_{j=1}^N \left\{ (\gamma_j - \gamma_{j-1})^2 - \varepsilon^2 \omega^2 \left(\frac{\gamma_j + \gamma_{j-1}}{2} \right)^2 \right\} \right] \end{aligned} \quad (10.106b)$$

We have the boundary condition $\gamma_0 = \gamma_N = 0$ because $\eta(t_i) = \eta(t_f) = 0$. The exponent is quadratic in γ and can be rewritten in matrix form:

$$\sum_{j=1}^N \left\{ (\gamma_j - \gamma_{j-1})^2 - \varepsilon^2 \omega^2 \left(\frac{\gamma_j + \gamma_{j-1}}{2} \right)^2 \right\} = \sum_{j,k=1}^{N-1} \gamma_j M_{jk} \gamma_k = \mathbf{\gamma}^T \mathbf{M} \mathbf{\gamma} \quad (10.107)$$

where

$$\mathbf{\gamma} = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_j \\ \vdots \\ \gamma_{N-1} \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} a & b & 0 & \cdots \\ b & a & b & \cdots \\ 0 & b & a & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (10.108a)$$

$$a = 2 \left(1 - \frac{\varepsilon^2 \omega^2}{4} \right), \quad b = - \left(1 + \frac{\varepsilon^2 \omega^2}{4} \right) \quad (10.108b)$$

$\mathbf{M}$ is a symmetric or more generally hermitian matrix and therefore can be diagonalized by a unitary matrix U . Then

$$\mathbf{y}^T \mathbf{M} \mathbf{y} = \mathbf{y}^T U^\dagger \mathbf{M}_D U \mathbf{y} = \mathbf{z}^\dagger \mathbf{M}_D \mathbf{z} = \sum_{j=1}^{N-1} \lambda_j z_j^2 \quad (10.109a)$$

where

$$\mathbf{z} = U \mathbf{y}, \quad \mathbf{M}_D = \begin{bmatrix} \lambda_1 & 0 & \cdots \\ 0 & \lambda_2 & \cdots \\ \vdots & \vdots & \\ & & \lambda_{N-1} \end{bmatrix} \quad (10.109b)$$

The Jacobian for the change of the variables from y_j to z_j is unity, because the transformation is unitary. The integration over all z_j gives

$$\left[\frac{\pi^{(N-1)}}{\prod_{j=1}^{N-1} \lambda_j} \right]^{1/2} = \left[\frac{\pi^{(N-1)}}{\det \mathbf{M}} \right]^{1/2} \quad (10.110)$$

Substituting Eq. (10.110) back into Eq. (10.106) gives

$$\begin{aligned} K_{\text{osci}}(f, i) &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{N/2} \left(\frac{2i\hbar \varepsilon}{m} \right)^{(N-1)/2} e^{\frac{i}{\hbar} S[q_d]} \\ &\quad \times \int d\gamma_1 \cdots d\gamma_{N-1} \exp[-\mathbf{y}^\dagger \mathbf{M} \mathbf{y}] \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty \\ N\varepsilon = t_f - t_i}} \left(\frac{m}{2\pi i \hbar \varepsilon \det \mathbf{M}} \right)^{1/2} e^{\frac{i}{\hbar} S[q_d]} \end{aligned} \quad (10.111)$$

The action has an interaction term, but the result is still governed by the classical path. This is because the Lagrangian contains only up to quadratic terms. In that case the path integral over the quantum fluctuation only changes the normalization factor, which appears as $\sqrt{1/\varepsilon \det \mathbf{M}}$. It is clear that the path integral can be defined only if the matrix $\mathbf{M}$ does not have any vanishing eigenvalues.

It remains to evaluate the determinant of the matrix. Denoting the determinant of a $j \times j$ submatrix as D_j , and inspecting Eq. (10.108a), we see it satisfies the following recursion relations:

$$D_{j+1} = a D_j - b^2 D_{j-1} \quad (10.112a)$$

$$D_{-1} = 0, \quad D_0 = 1, \quad D_1 = a \quad (10.112b)$$

We note that we can reformulate the recursion relations also in a simple matrix form as

$$\begin{bmatrix} D_{j+1} \\ D_j \end{bmatrix} = \begin{bmatrix} a & -b^2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} D_j \\ D_{j-1} \end{bmatrix} \quad (10.113)$$

Successive insertion gives

$$\begin{bmatrix} D_{N-1} \\ D_{N-2} \end{bmatrix} = \underbrace{\begin{bmatrix} a & -b^2 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} D_1 \\ D_0 \end{bmatrix}}_{(N-2) \text{ factors}} = \begin{bmatrix} a & -b^2 \\ 1 & 0 \end{bmatrix}^{N-2} \begin{bmatrix} a \\ 1 \end{bmatrix} \quad (10.114)$$

We can determine eigenvalues of the basic 2×2 matrix in Eq. (10.114) easily. From

$$\det \begin{vmatrix} a - \lambda & -b^2 \\ 1 & -\lambda \end{vmatrix} = 0 \quad (10.115)$$

we obtain

$$\lambda_{\pm} = \frac{1}{2}(a \pm \sqrt{a^2 - 4b^2}) \quad (10.116a)$$

$$a = \lambda_+ + \lambda_- \quad (10.116b)$$

A similarity matrix to diagonalize the basic matrix is also easily obtained:

$$S = \begin{bmatrix} c\lambda_+ & d\lambda_- \\ c & d \end{bmatrix}, \quad S^{-1} = \frac{1}{\lambda_+ - \lambda_-} \begin{bmatrix} 1/c & -\lambda_-/c \\ -1/d & \lambda_+/d \end{bmatrix} \quad (10.117)$$

Then

$$\begin{bmatrix} a & -b^2 \\ 1 & 0 \end{bmatrix} = S \begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix} S^{-1} \quad (10.118)$$

Inserting this in Eq. (10.114), we obtain

$$\begin{bmatrix} D_{N-1} \\ D_{N-2} \end{bmatrix} = S \begin{bmatrix} \lambda_+^{N-2} & 0 \\ 0 & \lambda_-^{N-2} \end{bmatrix} S^{-1} \begin{bmatrix} a \\ 1 \end{bmatrix} \quad (10.119)$$

Inserting Eq. (10.117) and Eq. (10.116) in Eq. (10.119)

$$D_{N-1} = \frac{\lambda_+^N - \lambda_-^N}{\lambda_+ - \lambda_-} \quad (10.120)$$

This formula is consistent with Eqs. (10.112). Using Eq. (10.116a) and Eq. (10.108b)

$$\begin{aligned} \lambda_+ - \lambda_- &= (a^2 - 4b^2)^{1/2} = \left[4 \left(1 - \frac{\varepsilon^2 \omega^2}{4} \right)^2 - 4 \left(1 + \frac{\varepsilon^2 \omega^2}{4} \right)^2 \right]^{1/2} \\ &= (-4\varepsilon^2 \omega^2)^{1/2} = 2i\varepsilon\omega \end{aligned} \quad (10.121a)$$

$$\begin{aligned} \lambda_+^N &= \left[\frac{a + \sqrt{a^2 - 4b^2}}{2} \right]^N = \left[\frac{2 \left(1 - \frac{\varepsilon^2 \omega^2}{4} \right) + 2i\varepsilon\omega}{2} \right]^N \\ &= [1 + i\varepsilon\omega + O(\varepsilon^2)]^N \end{aligned} \quad (10.121b)$$

$$\begin{aligned}\lambda_-^N &= \left[\frac{a - \sqrt{a^2 - 4b^2}}{2} \right]^N = \left[\frac{2 \left(1 - \frac{\varepsilon^2 \omega^2}{4} \right) - 2i\varepsilon\omega}{2} \right]^N \\ &= [1 - i\varepsilon\omega + O(\varepsilon^2)]^N\end{aligned}\quad (10.121c)$$

Substituting Eqs. (10.121) in Eq. (10.120), we obtain

$$\begin{aligned}\lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \varepsilon D_{N-1} &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \varepsilon \frac{\lambda_+^N - \lambda_-^N}{\lambda_+ - \lambda_-} \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \frac{\varepsilon}{2i\varepsilon\omega} [(1 + i\varepsilon\omega)^N - (1 - i\varepsilon\omega)^N] \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \frac{\varepsilon}{2i\varepsilon\omega} \left[\left(1 + \frac{i\omega T}{N} \right)^N - \left(1 - \frac{i\omega T}{N} \right)^N \right] \\ &= \frac{e^{i\omega T} - e^{-i\omega T}}{2i\omega} = \frac{\sin \omega T}{\omega}\end{aligned}\quad (10.122)$$

We obtain the propagator for the harmonic oscillator as

$$\begin{aligned}K_{\text{osci}}(f, i) &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty \\ N\varepsilon = t_f - t_i}} \left(\frac{m}{2\pi i\hbar \varepsilon \det M} \right)^{1/2} e^{\frac{i}{\hbar} S[q_{\text{cl}}]} \\ &= \left(\frac{m\omega}{2\pi i\hbar \sin \omega T} \right)^{1/2} e^{\frac{i}{\hbar} S[q_{\text{cl}}]}\end{aligned}\quad (10.123)$$

Note all the potentially singular terms have disappeared and the end result is finite.

Problem 10.4

Show that for $L = \frac{1}{2} m \dot{q}^2 - \frac{1}{2} m \omega^2 q^2$ with boundary condition $q(t_i) = q_i$, $q(t_f) = q_f$

$$e^{\frac{i}{\hbar} S[q_{\text{cl}}]} = \exp \left[\frac{i m \omega}{2\hbar \sin \omega T} \{ (q_f^2 + q_i^2) \cos \omega T - 2q_f q_i \} \right] \quad (10.124a)$$

$$\begin{aligned}K_{\text{osci}}(f, i) &= \left(\frac{m\omega}{2\pi i\hbar \sin \omega T} \right)^{1/2} \\ &\times \exp \left[\frac{i m \omega}{2\hbar \sin \omega T} \{ (q_f^2 + q_i^2) \cos \omega T - 2q_f q_i \} \right]\end{aligned}\quad (10.124b)$$

(2) Operator Method

It is a well-known fact that a linear operator⁵⁾ has matrix representations and thus the matrix method developed in the previous section can easily be extended to the

5) If an operator satisfies the relation

$$\hat{O}(a|\psi_1\rangle + b|\psi_2\rangle) = a\hat{O}|\psi_1\rangle + b\hat{O}|\psi_2\rangle \quad (10.125)$$

it is said to be linear.

case of operators. Here, we calculate the determinant of the hermitian operator explicitly, which has more applications later in field theory. The important part of the path integral for the harmonic oscillator has the form

$$I = \int \mathcal{D}\eta \exp \left[\frac{i}{\hbar} \int_{t_i}^{t_f} dt (\dot{\eta}^2 - \omega^2 \eta^2) \right] \quad (10.126)$$

By partial integration, it can be rewritten as

$$I = \int \mathcal{D}\eta \exp \left[\frac{i}{\hbar} \int_{t_i}^{t_f} dt \eta(t) \left\{ -\frac{d^2}{dt^2} - \omega^2 \right\} \eta(t) \right] \quad (10.127)$$

To handle the operator as a matrix we introduce two mathematical tricks.

1) First we have to make sure the integral is well behaved i.e. convergent. This is made possible when the function has a certain property so that we could analytically continue into the complex plane by rotating the time axis $t \rightarrow t e^{-i\theta}$. This is called Wick rotation. The technique is widely used and when we set the rotation angle $\theta = \pi/2$ or $t \rightarrow -i\tau$, and correspondingly $d^2/dt^2 \rightarrow -d^2/d\tau^2$, we have moved to the Euclidean space from the Minkowski space. Namely,

$$-S_E(x, \tau) = i S_M(x, -i\tau) \quad (10.128)$$

The sense of the rotation is completely fixed by the singularity structure of the theory. We shall discuss it further later at Sect. 10.6 when we define the generating functional. Then we have the Euclidean action

$$S_E = \int_{\tau_i}^{\tau_f} d\tau \eta(\tau) \left[-\frac{d^2}{d\tau^2} + \omega^2 \right] \eta(\tau) \equiv \int_{\tau_i}^{\tau_f} d\tau \eta(\tau) A \eta(\tau) \quad (10.129)$$

and the path integral

$$I = \int \mathcal{D}q e^{-\frac{1}{\hbar} S_E} \quad (10.130)$$

becomes a well-behaved integral as long as S_E is positive definite. After performing the calculation, we can go back to the original Minkowski space by setting τ back to $\tau = it$.

2) To show the matrix nature of the operator A , we start as usual from the discrete time interval.

If we divide the Euclidean time interval into N segments and put

$$\begin{aligned} \varepsilon = \varepsilon' &= \frac{\tau_f - \tau_i}{N}, \quad \eta_j = \eta(j\varepsilon), \quad \eta_k = \eta(k\varepsilon') \\ \frac{d^2}{d\tau^2} \eta(\tau) &= \frac{1}{\varepsilon^2} (\eta_{j+1} - 2\eta_j + \eta_{j-1}) \end{aligned} \quad (10.131)$$

The time integral is expressed as the continuous limit of the discrete sum

$$\begin{aligned} S_E &= - \sum_j^{N-1} \varepsilon \eta_j \left[\frac{1}{\varepsilon^2} (\eta_{j+1} - 2\eta_j + \eta_{j-1}) - \omega^2 \eta_j \right] \\ &= \sum_{j,k} \varepsilon \varepsilon' \eta_j A_{jk} \eta_k = \sum_{j,k} \varepsilon \varepsilon' \boldsymbol{\eta}^T A \boldsymbol{\eta} \end{aligned} \quad (10.132)$$

where

$$A_{jk} = -\frac{1}{\varepsilon'} \left[\frac{\delta_{j+1,k} - 2\delta_{j,k} + \delta_{j-1,k}}{\varepsilon^2} \right] + \omega^2 \frac{\delta_{j,k}}{\varepsilon'} \quad (10.133)$$

In the last line of Eq. (10.132) we used bold letters for $\boldsymbol{\eta}$ to emphasize its vector property $\boldsymbol{\eta} = (\eta_1, \dots, \eta_N)$. In the continuous limit, this gives

$$\begin{aligned} S_E &\equiv \int d\tau \boldsymbol{\eta}^T(\tau) A(\tau) \boldsymbol{\eta}(\tau) = \int d\tau d\tau' \boldsymbol{\eta}^T(\tau) A(\tau, \tau') \boldsymbol{\eta}(\tau') \\ A(\tau, \tau') &= \left(-\frac{d^2}{d\tau^2} + \omega^2 \right) \delta(\tau - \tau') \end{aligned} \quad (10.134)$$

where we have used the relation

$$\lim_{\varepsilon' \rightarrow 0} \frac{\delta_{j,k}}{\varepsilon'} = \delta(\tau - \tau') \quad (10.135)$$

In this form the operator is clearly seen to be a matrix whose elements are specified by two continuous variables. Referring to Eq. (10.133), $A(\tau, \tau')$ is a symmetric operator and in the proper basis satisfies the eigenequation

$$A\phi_j = \lambda_j \phi_j \quad (10.136)$$

As the eigenfunction constitutes a complete basis, the variable $\boldsymbol{\eta}$ can be expanded in terms of ϕ_j

$$\boldsymbol{\eta}(\tau) = \sum_j a_j \phi_j(\tau) \quad (10.137)$$

The boundary condition $\boldsymbol{\eta}(\tau_i) = \boldsymbol{\eta}(\tau_f) = 0$ translates to $\phi_j(\tau_i) = \phi_j(\tau_f) = 0$. The normalized eigenfunction is easily obtained as

$$\phi_j(\tau) = \sqrt{\frac{2}{(\tau_f - \tau_i)}} \sin \frac{j\pi}{(\tau_f - \tau_i)} (\tau - \tau_i) \quad j = 1, \dots, N-1 \quad (10.138a)$$

$$\lambda_j = \left(\frac{j\pi}{\tau_f - \tau_i} \right)^2 + \omega^2 \quad (10.138b)$$

We note that the eigenvalues are positive definite and ensure the convergence of the Euclidean integral. Note, also, that if we assume the time integral is discretized into N intervals as before, there are $N-1$ intermediate time points and there can be just as many $(N-1)$ independent coefficients a_j in the Fourier expansion in Eq. (10.137). Therefore it can be considered as an orthogonal transformation to diagonalize the matrix A . Equation (10.137) is still valid if $N \rightarrow \infty$ and remains an orthogonal transformation that keeps the value of the Jacobian at unity. When we

change the path integral variable from η to a_j , we obtain

$$\begin{aligned}
 \int d\tau (\boldsymbol{\eta}^T \mathbf{A} \boldsymbol{\eta}) &= \int d\tau d\tau' \sum_{j,k=1}^{\infty} a_j \phi_j(\tau) A(\tau, \tau') a_k \phi_k(\tau') \\
 &= \sum_{j=1}^{\infty} \lambda_j a_j^2 \\
 \therefore \int \mathcal{D}\eta e^{-\frac{1}{\hbar} \int d\tau \boldsymbol{\eta}^T \mathbf{A} \boldsymbol{\eta}} &= \prod_{j=1}^{\infty} \int da_j e^{-\frac{1}{\hbar} \lambda_j a_j^2} = \prod_{j=1}^{\infty} \left(\frac{\pi \hbar}{\lambda_j} \right)^{1/2} \\
 &= C (\det \mathbf{A})^{-1/2}
 \end{aligned} \tag{10.139}$$

The determinant of the hermitian operator is expressed formally as

$$\begin{aligned}
 \det \left(-\frac{d^2}{d\tau^2} + \omega^2 \right) &= \det \mathbf{A} = \prod_{j=1}^{\infty} \lambda_j \\
 &= \prod_{j=1}^{\infty} \left(\frac{j\pi}{\tau_f - \tau_i} \right)^2 \prod_{j=1}^{\infty} \left[1 + \left\{ \frac{\omega(\tau_f - \tau_i)}{j\pi} \right\}^2 \right] \\
 &= C' \frac{\sinh \omega(\tau_f - \tau_i)}{\omega(\tau_f - \tau_i)} \\
 C' &= \prod_{j=1}^{\infty} \left(\frac{j\pi}{\tau_f - \tau_i} \right)^2
 \end{aligned} \tag{10.140}$$

C, C' are divergent constants that do not contain dynamical variables and can be absorbed in the normalization constant. Analytically continuing this back to real time, we obtain

$$\det \left(-\frac{d^2}{d\tau^2} + \omega^2 \right) \xrightarrow{\tau=it} \det \left(\frac{d^2}{dt^2} + \omega^2 \right) = N'^2 \frac{\sin \omega(t_f - t_i)}{\omega(t_f - t_i)} \tag{10.142}$$

where N' is a normalization constant. It may be divergent, but as long as it is independent of dynamical variables, the actual value is irrelevant for physical outcomes, as will be clarified in later discussions. As a result we obtain

$$I = \int \mathcal{D}\eta \exp \left[\frac{i}{\hbar} \int_{t_i}^{t_f} dt \eta \left(-\frac{d^2}{dt^2} - \omega^2 \right) \eta \right] = N \left[\frac{\sin \omega(t_f - t_i)}{\omega(t_f - t_i)} \right]^{-1/2} \tag{10.143}$$

6) We used the formulae

$$\sin x = x \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2 \pi^2} \right), \quad \cos x = \prod_{n=1}^{\infty} \left(1 - \frac{4x^2}{(2n-1)^2 \pi^2} \right) \tag{10.141}$$

The constant N can be determined by noting that, in the limit $\omega \rightarrow 0$, the propagator should agree with that of the free particle. The final result is

$$K_{\text{osci}}(f, i) = \left[\frac{m\omega}{2\pi i \hbar \sin \omega(t_f - t_i)} \right]^{1/2} e^{\frac{i}{\hbar} S[q_{\text{cl}}]} \quad (10.144)$$

which agrees with what we obtained using the matrix method. Our result was obtained for the harmonic oscillator, but the result can be extended to general quadratic actions of the form

$$I = \int \mathcal{D}\eta \exp \left(\frac{i}{\hbar} \int_{t_i}^{t_f} dt \, \eta^T(t) A \eta(t) \right) = \frac{N}{\sqrt{\det A}} \quad (10.145)$$

So far, we have used only real variables. The path integral can be generalized to complex variables. We note

$$\frac{2\pi}{a} = \int dx e^{-ax^2/2} \int d\gamma e^{-a\gamma^2/2} = \int dx d\gamma e^{-a(x^2 + \gamma^2)/2} \quad (10.146)$$

Introducing a complex variable $z = (x + i\gamma)/\sqrt{2}$, we write

$$\frac{2\pi}{a} = \frac{1}{i} \int dz dz^* e^{-az^*z} \quad (10.147)$$

Note, z and z^* are independent variables. Generalizing to n complex variables and writing $\mathbf{z} = (z_1, z_2, \dots)$,

$$\int d^n z d^n z^* e^{-z^\dagger A z} = \int d^n z d^n z^* e^{-z_i^* A_{jk} z_k} = \frac{(2\pi i)^n}{\det A} \quad (10.148)$$

In this case, it goes without saying that the N -dimensional symmetric matrix should be replaced by the hermitian matrix. Generalization of the path integral formula Eq. (10.145) to complex variables is now straightforward. Introducing the complex variable $\varphi(t) = (\eta + i\xi)/\sqrt{2}$

$$I = \int \mathcal{D}\varphi^* \mathcal{D}\varphi \exp \left(\frac{i}{\hbar} \int_{t_i}^{t_f} dt \, \varphi^\dagger(t) A \varphi(t) \right) = \frac{N'}{\det A} \quad (10.149)$$

where A is a hermitian operator.

10.5

Scattering Matrix

In this section, we try to calculate the scattering amplitude when we have a scattering potential in the Hamiltonian. In general, the propagator is not exactly calculable and we rely on the perturbation expansion. This is valid when the potential or, more precisely, when the time integral of $V(tq)$ is small compared with $\hbar$.

10.5.1

Perturbation Expansion

The propagator is given by Eq. (10.56)

$$\begin{aligned}
 K(t_f q_f; t_i q_i) &= C \int \mathcal{D}q \exp \left[\frac{i}{\hbar} \int dt \left\{ \frac{1}{2} m \dot{q}^2 - V(q) \right\} \right] \\
 &= \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{N/2} \int \cdots \int dq_1 \cdots dq_{N-1} \exp \left[\frac{i}{\hbar} \varepsilon \sum_{j=1}^N \{T_j - V_j\} \right]
 \end{aligned} \quad (10.150)$$

where T_j and V_j are the kinetic and potential part of the Lagrangian at $t = t_j$. As the action in the path integral is a classical action, the Lagrangian can be separated into the kinetic and potential terms. We then expand the potential part

$$\exp \left(-\frac{i}{\hbar} \int_{t_{k-1}}^{t_k} dt V(tq) \right) = 1 - \frac{i}{\hbar} \varepsilon V(t_k, q_k) + \frac{1}{2!} \left(-\frac{i}{\hbar} \varepsilon V(t_k, q_k) \right)^2 \cdots \quad (10.151)$$

When Eq. (10.151) is substituted back into Eq. (10.150), we obtain a series expansion

$$K = K_0 + K_1 + K_2 + \cdots \quad (10.152)$$

The first term does not contain V and is the free propagator:

$$\begin{aligned}
 K_0 &= C \int \mathcal{D}q \exp \left[\frac{1}{2} m \dot{q}^2 \right] \\
 &\stackrel{\text{Eq. (10.82)}}{=} \theta(t_f - t_i) \left(\frac{m}{2\pi i \hbar (t_f - t_i)} \right)^{1/2} \exp \left[\frac{i m (q_f - q_i)^2}{2 \hbar (t_f - t_i)} \right]
 \end{aligned} \quad (10.153)$$

K_1 includes only one $V(t_k, q_k)$ and is expressed as

$$\begin{aligned}
 K_1 &= \frac{-i}{\hbar} \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \sum_{k=1}^{N-1} \varepsilon \int dq_k \\
 &\quad \times \left[C^{N-k} \int dq_{k+1} \cdots dq_{N-1} \exp \left\{ \frac{i m}{2 \hbar \varepsilon} \sum_{j=k}^{N-1} (q_j - q_{j-1})^2 \right\} \right] \\
 &\quad \times V(t_k, q_k) \left[C^k \int dq_1 \cdots dq_{k-1} \exp \left\{ \frac{i m}{2 \hbar \varepsilon} \sum_{j=1}^{k-1} (q_j - q_{j-1})^2 \right\} \right] \\
 &= -\frac{i}{\hbar} \int_{t_i}^{t_f} dt \int_{-\infty}^{\infty} dq K_0(t_f q_f; tq) V(tq) K_0(tq; t_i q_i)
 \end{aligned} \quad (10.154)$$

where $C = \sqrt{m/2\pi i \hbar \varepsilon}$. As $K(t_f q_f; tq)$ vanishes if $t > t_f$, and $K(tq; t_i q_i)$ if $t < t_i$, the integration over time can be extended to all values of t to give

$$K_1 = -\frac{i}{\hbar} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dq K_0(t_f q_f; tq) V(tq) K_0(tq; t_i q_i) \quad (10.155)$$

Similarly, we can prove that the second-order term is expressed as

$$K_2 = \left(-\frac{i}{\hbar}\right)^2 \int_{-\infty}^{\infty} dt_1 dt_2 \int_{-\infty}^{\infty} dq_1 dq_2 \quad (10.156)$$

$$\times [K_0(t_f q_f; t_2 q_2) V(t_2 q_2) K_0(t_2 q_2; t_1 q_1) V(t_1 q_1) K_0(t_1 q_1; t_i q_i)]$$

Analogous expressions hold for all K_j in the expansion Eq. (10.152). Note, there is no $1/2!$ in Eq. (10.156). The reason for this is that the two interactions occur at different times, but they are integrated over all space-time and they are indistinguishable. The effect is merely to compensate $1/2!$. Similarly, there are $j!$ different combinations that give identical contributions, and no $1/j!$ appears in the expression for K_j . Equation (10.152) is now expressed as

$$K(t_f q_f; t_i q_i)$$

$$= K_0(t_f q_f; t_i q_i) - \frac{i}{\hbar} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dq K_0(t_f q_f; t q) V(t q) K_0(t q; t_i q_i) + \cdots$$

$$= K_0(t_f q_f; t_i q_i) - \frac{i}{\hbar} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dq K_0(t_f q_f; t q) V(t q) K(t q; t_i q_i) \quad (10.157)$$

The first equation is the perturbation expansion of the propagator, called the Born series. The second equation is obtained by replacing K_0 in the second term by K , which is equivalent to including all the successive terms. It is exact and is an integral equation for the propagator.

We have elaborated on using the path integral method to derive the perturbation series. In fact, it can be derived directly from Eq. (10.35). Separating the potential from the Hamiltonian $H = H_0 + V$ it can be rewritten as

$$\left(i\hbar \frac{\partial}{\partial t_1} - H_0\right) K(t_1 q_1; t_2 q_2) = i\hbar \delta(t_1 - t_2) \delta(q_1 - q_2) + V(t_1 q_1) K(t_1 q_1; t_2 q_2) \quad (10.158)$$

Since K_0 satisfies the same equation with $V = 0$, the equation can be transformed to a integral equation that reproduces Eq. (10.157).

Using the propagator, we can now express the time development of a wave function. Inserting Eq. (10.157) in Eq. (10.40),

$$\psi(t_f q_f) = \int dq_i K(t_f q_f; t_i q_i) \psi(t_i q_i)$$

$$= \int dq_i K_0(t_f q_f; t_i q_i) \psi(t_i q_i)$$

$$- \frac{i}{\hbar} \int dt dq \int dq_i K_0(t_f q_f; t q) V(t q) K(t q; t_i q_i) \psi(t_i q_i)$$

$$= \int dq_i K_0(t_f q_f; t_i q_i) \psi(t_i q_i)$$

$$- \frac{i}{\hbar} \int dt dq K_0(t_f q_f; t q) V(t q) \psi(t q) \quad (10.159)$$

Note, the time sequence is such that $t_f > t > t_i$, but the integral can be made over the entire time region from $-\infty$ to $+\infty$, because the existence of the retarded propagator guarantees the vanishing of the integrand for $t > t_f$ or $t < t_i$. Here we have changed from one space dimension to three. This is an integral equation for the wave function. Now we assume that for $t \rightarrow -\infty$, ψ becomes a free plane wave function, which we denote as ϕ . Then, the first term of the rhs of Eq. (10.159) is also a plane wave, because the free propagation of a plane wave is also a plane wave:

$$\psi(t_f \mathbf{q}_f) = \phi(t_f \mathbf{q}_f) - \frac{i}{\hbar} \int dt d\mathbf{q} K_0(t_f \mathbf{q}_f; t \mathbf{q}) V(t \mathbf{q}) \psi(t \mathbf{q}) \quad (10.160)$$

As $\psi(t \mathbf{q})$ obeys the Schrödinger equation, namely

$$\left(i \frac{\partial}{\partial t} - H_0 \right) \psi(t \mathbf{q}) = V(t \mathbf{q}) \psi(t \mathbf{q}) \quad (10.161)$$

where H_0 is the free Hamiltonian and $\phi(t \mathbf{q})$ is the solution to the free equation ($V = 0$), the propagator must obey the equation

$$\left(i \frac{\partial}{\partial t} - H_0 \right) K_0(t \mathbf{q}; t' \mathbf{q}') = i \hbar \delta(t - t') \delta^3(\mathbf{q} - \mathbf{q}') \quad (10.162)$$

This is the equation for the Green's function of Eq. (10.161), which reproduces Eq. (10.35).

Problem 10.5

Using the free-particle propagator Eq. (10.82) or Eq. (10.95), prove that

$$\int d\mathbf{q}_i K_0(t_f \mathbf{q}_f; t_i \mathbf{q}_i) e^{\frac{i}{\hbar}(p \mathbf{q}_i - i E t_i)} = e^{\frac{i}{\hbar}(p \mathbf{q}_f - i E t_f)}, \quad E = \frac{p^2}{2m} \quad (10.163)$$

Extension to three dimensions is straightforward.

10.5.2

S-Matrix in the Path Integral⁷⁾

Now we want to apply our results to more practical situations, namely how to relate the wave function to experimental measurements. In a typical scattering experiment a bunch of particles is provided with a definite initial condition, say, fixed momentum $\mathbf{p}_i$, they illuminate a target and the probability that they are scattered to a specific final state, at momentum $\mathbf{p}_f$, is measured. The incident and scattered particles are considered free in the distant past $t \rightarrow -\infty$ or in the far future $t \rightarrow +\infty$. We represent the initial- and final-state wave functions of particles with

7) Compare the discussion here with that in Sects. 6.1–6.3.

momentum $\mathbf{p}_i$ and $\mathbf{p}_f$ by

$$\psi_{\text{in}}(t\mathbf{q}; \mathbf{p}_i) = N e^{\frac{i}{\hbar}(\mathbf{p}_i \cdot \mathbf{q} - E_i t)} = \langle t\mathbf{q} | \mathbf{p}_i \rangle \quad (10.164a)$$

$$\psi_{\text{out}}(t\mathbf{q}; \mathbf{p}_f) = N e^{\frac{i}{\hbar}(\mathbf{p}_f \cdot \mathbf{q} - E_f t)} = \langle t\mathbf{q} | \mathbf{p}_f \rangle \quad (10.164b)$$

One problem with the expressions is that the plane wave spreads over all space and time, including the scattering center $V(t\mathbf{q})$, so the particles can never be free. We could use a wave packet formalism to circumvent the problem but usually this involves complicated calculation. The conventional way around the problem is the adiabatic treatment, in which the potential is switched on and off slowly as a function of time. Namely, we replace V by $V e^{-\varepsilon|t|}$ so that $V = 0$ at $t = \pm\infty$.⁸⁾ Assuming that is the case, and with the initial conditions $\psi = \psi_{\text{in}}$, $V \rightarrow 0$ for $|t| \rightarrow \infty$, the wave function at time $t = t_f$ is expressed as

$$\begin{aligned} \psi^+(t_f \mathbf{q}_f) &= \int d\mathbf{q}_i K_0(t_f \mathbf{q}_f; t_i \mathbf{q}_i) \psi_{\text{in}}(t_i \mathbf{q}_i) \\ &\quad - \frac{i}{\hbar} \int dt d\mathbf{q} K_0(t_f \mathbf{q}_f; t \mathbf{q}) V(t \mathbf{q}) \psi^+(t \mathbf{q}) \end{aligned} \quad (10.165)$$

This is the same equation as Eq. (10.159) except $\psi(t_i, \mathbf{q}_f)$ in the first term is replaced by ψ_{in} and the implicit assumption that $|V| \rightarrow 0$ at $t \rightarrow -\infty$. The superscript on $\psi^+(t\mathbf{q})$ denotes that it corresponds to a wave that was free at $t \rightarrow -\infty$, which is the result of the retarded nature of the propagator such that $K_0(t\mathbf{q}; t'\mathbf{q}') = 0$ for $t < t'$. To eliminate ψ^+ on the rhs, we make successive iterations, using the first term as the first input to ψ^+ in the second term. The result is

$$\begin{aligned} \psi^+(t_f \mathbf{q}_f) &= \int d\mathbf{q}_i K_0(t_f \mathbf{q}_f; t_i \mathbf{q}_i) \psi_{\text{in}}(t_i \mathbf{q}_i) \\ &\quad - \frac{i}{\hbar} \int dt d\mathbf{q} d\mathbf{q}_i K_0(t_f \mathbf{q}_f; t \mathbf{q}) V(t \mathbf{q}) K_0(t \mathbf{q}; t_i \mathbf{q}_i) \psi_{\text{in}}(t_i \mathbf{q}_i) + \cdots \end{aligned} \quad (10.166)$$

8) This is equivalent to introducing a small negative imaginary part for the energy, i.e. $E \rightarrow E - i\varepsilon$, which in turn is Feynman's recipe for handling the singular point in the propagator.

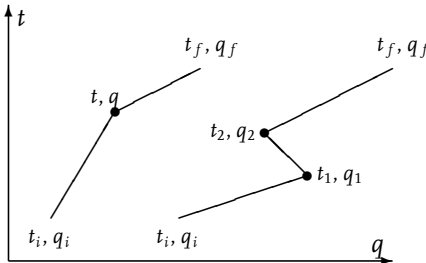


Figure 10.3 First- and second-order scattering diagram. Time runs from bottom to top.

We are interested in the transition amplitude from the state $|\psi_{\text{in}}\rangle$ at $t = -\infty$ to $|\psi_{\text{out}}\rangle$ at $t = \infty$, which is called the scattering amplitude. It is defined as

$$\begin{aligned}
 S_{fi} &= \langle \psi_{\text{out}} | \psi^+ \rangle_{t_f \rightarrow \infty} = \int d\mathbf{q}_f \langle \psi_{\text{out}} | t_f \mathbf{q}_f \rangle \langle t_f \mathbf{q}_f | \psi^+ \rangle \\
 &= \int d\mathbf{q}_f \psi_{\text{out}}^*(t_f \mathbf{q}_f) \psi^+(t_f \mathbf{q}_f) \\
 &= \int d\mathbf{q}_f d\mathbf{q}_i \psi_{\text{out}}^*(t_f \mathbf{q}_f) K_0(t_f \mathbf{q}_f; t_i \mathbf{q}_i) \psi_{\text{in}}(t_i \mathbf{q}_i) \\
 &\quad - \frac{i}{\hbar} \int dt d\mathbf{q}_f d\mathbf{q} d\mathbf{q}_i \psi_{\text{out}}^*(t_f \mathbf{q}_f) K_0(t_f \mathbf{q}_f; t \mathbf{q}) \\
 &\quad \times V(t \mathbf{q}) K_0(t \mathbf{q}; t_i \mathbf{q}_i) \psi_{\text{in}}(t_i \mathbf{q}_i) \\
 &\quad + \left(-\frac{i}{\hbar}\right)^2 \int dt_1 dt_2 d\mathbf{q}_f d\mathbf{q}_2 d\mathbf{q}_1 d\mathbf{q}_i \\
 &\quad \times \psi_{\text{out}}^*(t_f \mathbf{q}_f) K_0(t_f \mathbf{q}_f; t_2 \mathbf{q}_2) V(t_2, \mathbf{q}_2) K_0(t_2 \mathbf{q}_2; t_1 \mathbf{q}_1) \\
 &\quad \times V(t_1 \mathbf{q}_1) K_0(t_1 \mathbf{q}_1; t_i \mathbf{q}_i) \psi_{\text{in}}(t_i \mathbf{q}_i) + \dots
 \end{aligned} \tag{10.167}$$

where $\dots$ are higher order terms. The formula can be depicted as in Fig. 10.3.

1. Attach $\psi_{\text{in}}(t_i, \mathbf{q}_i)$ at point $(t_i, \mathbf{q}_i)$.
2. Draw a line from $(t_i, \mathbf{q}_i)$ to $(t, \mathbf{q})$. Attach $K_0(t_i \mathbf{q}_i; t, \mathbf{q})$ to the line.
3. Attach the vertex $\left[-\frac{i}{\hbar} V(t, \mathbf{q})\right]$ to the point $(t, \mathbf{q})$.
4. Draw a line from $(t, \mathbf{q})$ to $(t_f, \mathbf{q}_f)$. Attach $K_0(t \mathbf{q}; t_f, \mathbf{q}_f)$ to the line.
5. Attach $\psi_{\text{out}}(t_f, \mathbf{q}_f)$ at point $(t_f, \mathbf{q}_f)$.
6. Integrate over time t and over all the intermediate positions $\mathbf{q}_i, \mathbf{q}, \mathbf{q}_f$.

This completes the first-order scattering process. The corresponding picture is drawn in Fig. 10.3 (left). The second-order scattering process can be drawn in a similar manner. The ordering of the factors is such that in writing each factor one goes from the bottom upwards. In the nonrelativistic treatment, one does not go downwards (backwards in time) in Fig. 10.3. The correspondence is called the Feynman rules.⁹⁾ In nonrelativistic quantum mechanics, which we are dealing with at present, the pictorial representations shown in Fig. 10.3 are hardly necessary, but in quantum field theory they are a great aid to calculation.

As Problem 10.5 shows

$$\begin{aligned}
 \int d\mathbf{q}_i K_0(t_f \mathbf{q}_f; t_i \mathbf{q}_i) \psi_{\text{in}}(t_i \mathbf{q}_i) &= \psi_{\text{in}}(t_f \mathbf{q}_f) \\
 \int d\mathbf{q}_i \psi_{\text{out}}^*(t_f \mathbf{q}_f) K_0(t_f \mathbf{q}_f; t_i \mathbf{q}_i) &= \psi_{\text{out}}^*(t_i \mathbf{q}_i)
 \end{aligned} \tag{10.168}$$

9) In field theory, particles with negative energy can be formally interpreted as going backwards in time. The relativistic Feynman propagator includes both retarded and advanced propagation. One can go upwards or downwards freely in time and need not worry about the time sequence. The Feynman propagator takes care of it automatically (see discussion in Sect. 6.6.2).

Using Eq. (10.168), we can rewrite Eq. (10.167). The first term gives simply $\delta^3(\mathbf{p}_f - \mathbf{p}_i)$, which means no scattering. The interesting part is included in the second and following terms. Here, we write only up to second terms, which is the first-order term in the power expansion of the potential V :

$$\begin{aligned} S_{fi} &= \delta^3(\mathbf{p}_f - \mathbf{p}_i) - \frac{i}{\hbar} \int dt d\mathbf{q} \psi^*(t\mathbf{q}) V(t\mathbf{q}) \psi_{\text{in}}(t\mathbf{q}) \\ &= \delta^3(\mathbf{p}_f - \mathbf{p}_i) - \frac{i}{\hbar} N^2 \int dt d\mathbf{q} e^{-\frac{i}{\hbar}(\mathbf{p}_f \cdot \mathbf{q} - i E_f t)} V(t\mathbf{q}) e^{\frac{i}{\hbar}(\mathbf{p}_i \cdot \mathbf{q} - i E_i t)} \end{aligned} \quad (10.169)$$

where N is the normalization factor. For the time independent Coulomb potential, the time integration merely gives a δ function. Thus, the second term gives

$$S_{fi} - \delta_{fi} = -2\pi i \delta(E_f - E_i) N^2 \int e^{\frac{i}{\hbar}(\mathbf{p}_i - \mathbf{p}_f) \cdot \mathbf{q}} V(\mathbf{q}) d\mathbf{q} \quad (10.170)$$

This is the well-known Born approximation in quantum mechanics.

Problem 10.6

Prove Eqs. (10.169).

10.6

Generating Functional

10.6.1

Correlation Functions

So far we have considered the transition amplitude in the form of a path integral:

$$\langle t_f q_f | t_i q_i \rangle = N \int \mathcal{D}q e^{\frac{i}{\hbar} S[q]} \quad (10.171)$$

Later on, we shall need matrix elements of operator products such as

$$\begin{aligned} D_{12} &= \langle t_f q_f | \hat{q}(t_1) \hat{q}(t_2) | t_i q_i \rangle \\ &= \int \mathcal{D}q q_1 q_2 \exp \left[i \int_{t_i}^{t_f} dt L(q) \right] \quad t_f \geq t_1 \geq t_2 \geq t_i \end{aligned} \quad (10.172)$$

where the boundary conditions on the path integral are

$$q_f = q(t_f), \quad q_i = q(t_i), \quad q_1 = q(t_1), \quad q_2 = q(t_2) \quad (10.173)$$

We can show that it is related to the two-point vacuum correlation function

$$\langle \Omega | T[\hat{q}(t_1) \hat{q}(t_2)] | \Omega \rangle \quad (10.174)$$

where $|\Omega\rangle$ is the vacuum state in the Heisenberg picture. We add a hat, e.g. $\hat{q}$, when we want to distinguish operators from their representations q , which are ordinary numbers. To evaluate Eq. (10.172) we break up the functional integral as follows:

$$\int \mathcal{D}q = \int dq_1 \int dq_2 \int_{\substack{q(t_1)=q_1 \\ q(t_2)=q_2}} \mathcal{D}q \quad t_f \geq t_1 \geq t_2 \geq t_i \quad (10.175)$$

The main functional integral $\int \mathcal{D}q$ is now constrained at times t_1 and t_2 (in addition to the endpoints t_f and t_i). There $q_1 = q(t_1)$ and $q(t_2)$ are held constant. Their integration has to be done separately and is placed in front of the main functional integral. Then the numbers q_1, q_2 can be taken outside the main integral. The main integral is now divided into three transition amplitudes:

$$\begin{aligned} D_{12} &= \int dq_1 \int dq_2 q_1 q_2 \int_{\substack{q(t_1)=q_1 \\ q(t_2)=q_2}} \mathcal{D}q \exp \left[i \int_{t_i}^{t_f} dt L \right] \\ &= \int dq_1 \int dq_2 q_1 q_2 \langle t_f q_f | t_1 q_1 \rangle \langle t_1 q_1 | t_2 q_2 \rangle \langle t_2 q_2 | t_i q_i \rangle \\ &= \int dq_1 \int dq_2 q_1 q_2 \langle q_f | e^{-iH(t_f-t_1)} | q_1 \rangle \langle q_1 | e^{-iH(t_1-t_2)} | q_2 \rangle \\ &\quad \times \langle q_2 | e^{-iH(t_2-t_i)} | q_i \rangle \end{aligned} \quad (10.176)$$

We can turn the c-number q_1 into a Schrödinger operator using $q|q_1\rangle = \hat{q}_S|q_1\rangle$ [see Eq. (10.36b)]. The completeness relation $\int dq_1 |q_1\rangle \langle q_1| = 1$ allows us to eliminate the intermediate state $|q_1\rangle \langle q_1|$. The q_2 integration can be eliminated similarly, yielding the expression

$$D_{12} = \langle q_f | e^{-iH(t_f-t_1)} \hat{q}_S e^{-iH(t_1-t_2)} \hat{q}_S e^{-iH(t_2-t_i)} | q_i \rangle \quad (10.177)$$

The Schrödinger operator can be converted into the Heisenberg operator using the relation (10.36a). For the case $t_2 > t_1$, q_1 enters the path integral earlier than q_2 . Therefore the ordering of the operators $\hat{q}(t_1)\hat{q}(t_2)$ in D_{12} [Eq. (10.172)] should be reversed. Using a time-ordered product we can combine the two cases into one to give

$$D_{12} = \langle q_f | e^{-iH t_f} T[\hat{q}_H(t_1)\hat{q}_H(t_2)] e^{iH t_i} | t_i \rangle = \langle q_f t_f | T[\hat{q}_H(t_1)\hat{q}_H(t_2)] | t_i q_i \rangle \quad (10.178)$$

We see that the path integral has a built-in structure to give the time-ordered correlation function as its moments. A similar formula holds for arbitrary functions of operators:

$$\begin{aligned} &\langle t_f q_f | T[O_1(\hat{q}_H(t_1)) \cdots O_N(\hat{q}_H(t_N))] | t_i q_i \rangle \\ &= N \int \mathcal{D}q O_1(q(t_1)) \cdots O_N(q(t_N)) e^{\frac{i}{\hbar} S[q]} \end{aligned} \quad (10.179)$$

The salient feature of the transition amplitude is that we can extract the vacuum state by sending t_i, t_f to $\pm\infty$ in a slightly imaginary direction ($t \rightarrow t' = t e^{-i\delta}$).

This will extract the vacuum state $|\Omega\rangle$ from $|q_f\rangle$ and $|q_i\rangle$, because we can get rid of all the $n \neq 0$ states by making t_f, t_i large with a small negative imaginary part $t_f, t_i \rightarrow \pm\infty(1-i\delta)$:

$$\begin{aligned} & \lim_{t_i \rightarrow -\infty(1-i\delta)} e^{iHt_i} |q_i\rangle \\ &= \lim_{|t_i| \rightarrow \infty(1-i\delta)} \left[e^{-iE_0|t_i|} |\Omega\rangle \langle\Omega|q_i\rangle + \sum_{n \neq 0} e^{-iE_n|t_i|} |n\rangle \langle n|q_i\rangle \right] \\ &\Rightarrow \lim_{|t_i| \rightarrow \infty} |\Omega\rangle \left(\langle\Omega|q_i\rangle e^{-iE_0|t_i|(1-i\delta)} \right) \end{aligned} \quad (10.180a)$$

where $E_0 = \langle\Omega|H|\Omega\rangle$. Any state $|n\rangle$ for $n \neq 0$ has a larger damping factor than the $n = 0$ state and hence the ground state is the only surviving state after $|t_i|$ is made large. Similarly, it is shown that

$$\langle q_f | e^{-iHt_f} = \lim_{t_f \rightarrow \infty(1-i\delta)} (\langle q_f | \Omega \rangle e^{-iE_0 t_f(1-i\delta)} \langle \Omega | \quad (10.180b)$$

Using Eq. (10.180b) and (10.180a), the time-ordered correlation function is expressed as the vacuum expectation value:

$$\begin{aligned} & \lim_{\substack{t_i \rightarrow -\infty \\ t_f \rightarrow \infty}} \langle t_f q_f | T[\hat{q}(t_1)\hat{q}(q_2)] | t_i q_i \rangle \\ &= (\langle q_f | \Omega \rangle e^{-iE_0 \infty(1-i\delta)}) \langle \Omega | T[\hat{q}(t_1)\hat{q}(q_2)] | \Omega \rangle (\langle \Omega | q_i \rangle e^{-iE_0 \infty(1-i\delta)}) \end{aligned} \quad (10.181)$$

$$\therefore \langle \Omega | T[\hat{q}(t_1)\hat{q}(q_2)] | \Omega \rangle = \lim_{\substack{t_i \rightarrow -\infty \\ t_f \rightarrow \infty}} \frac{\langle t_f q_f | T[\hat{q}(t_1)\hat{q}(q_2)] | t_i q_i \rangle}{\langle q_f | \Omega \rangle \langle \Omega | q_i \rangle e^{-iE_0(t_f-t_i)(1-i\delta)}} \quad (10.182)$$

The rhs is independent of the endpoints and therefore the lhs must also be independent of the endpoints. The awkward extra phase and overlap factors can be eliminated by dividing by the same expression as D_{12} but without q_1 and q_2 . Our final expression for the n -point correlation function is

$$\langle \Omega | T[\hat{q}_H(t_1) \cdots \hat{q}_H(t_n)] | \Omega \rangle = \lim_{t_f, t_i \rightarrow \pm\infty(1-i\delta)} \frac{\int \mathcal{D}q \, q_1 \cdots q_n \exp \left[i \int_{t_i}^{t_f} dt L \right]}{\int \mathcal{D}q \exp \left[i \int_{t_i}^{t_f} dt L \right]} \quad (10.183)$$

10.6.2

Note on Imaginary Time

Tilting the time axis is a mathematical trick called the Wick rotation (Fig. 10.4a). The direction of the rotation is mandated by the requirement that there are no

singularities of the integrand encountered in the rotation of the time axis. This is the case for

$$t \rightarrow te^{-i\delta} \quad 0 \leq \delta < \pi \quad (10.184)$$

For larger angles, the rhs of Eq. (10.180) develops essential singularities for $t_i \rightarrow -\infty$ and $t_f \rightarrow \infty$. This rotation corresponds to an analytic continuation in the variable t . The real four-dimensional Minkowski space can be viewed as a subspace of a four-dimensional complex space (or of a five-dimensional space with three real space coordinates and one complex time coordinate). For $\delta = \pi/2$, this corresponds to $t = -i\tau$ (τ real) and the space becomes Euclidean. Physical results are obtained by evaluating the relevant expressions in Euclidean time because their oscillatory behavior is converted to an adequate damping factor and is well behaved. Then after the calculation, one can rotate the result ($\tau = it$) back to the real time axis (Minkowski time t) to obtain physically meaningful results. The method works well only if no singularities are encountered on the way.

We note that the end result of the Wick rotation and rotation back to Minkowski time is equivalent to addition of a small negative imaginary part in the energy. This produces the proper damping of the oscillating exponentials to make the integral converge. We know from our study of the Green's function in the complex energy plane (Fig. 10.4b), that the singularities have to occur at $\pm\omega \mp i\epsilon$ to incorporate the causal boundary conditions (see Sect. 6.6.2). Therefore, an analytic continuation from $\text{Re}(k^0)$ to $\text{Im}(k^0)$ is meaningful only if the rotation is counterclockwise. The two rotations are equivalent since we can represent $k^0 \rightarrow i\hbar\partial/\partial t$.

Furthermore, the Wick rotation also rationalizes the treatment of the scattering matrix in terms of plane waves, which are not local. The interaction is made to be turned on and off slowly without any energy transfer (i.e. adiabatic approximation) so that the notion of an asymptotically free field at large time is justified (see arguments in Sect. 10.5.2). All the reasonings stem from the constraint that the transition amplitude has to be finite at large time. In the following, we shall not always indicate it explicitly, but it should be understood as being done.

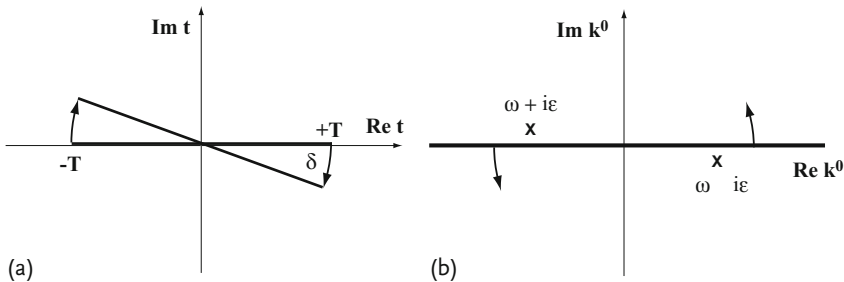


Figure 10.4 (a) Rotation of the time axis to ensure the convergence of the path integral. (b) Rotation in the complex energy plane to incorporate the causal boundary condition.

10.6.3

Correlation Functions as Functional Derivatives

Next, we show that we can generate various correlation functions by simply adding an appropriate external source. For instance, a charged particle generates an electromagnetic field, and its interaction is given by $A_\mu J^\mu$. The current J^μ plays the role of the source to generate the interactions. We define a modified action of the form

$$S[q, J] = S[q] + \int_{t_i}^{t_f} dt q(t) J(t) \quad (10.185)$$

and further define

$$Z[J] = \langle \Omega | \Omega \rangle_J = N \int \mathcal{D}q e^{\frac{i}{\hbar} S[q, J]} \quad (10.186)$$

where N is a normalization constant to make $Z[0] = 1$. We know how to rotate the time axis in the definition of $Z[J]$ to isolate the vacuum state. Equivalently, we may add a small negative imaginary part to the Hamiltonian. Adding $i\epsilon q^2$ to the Lagrangian will achieve this:

$$Z[J] = N \int \mathcal{D}q \exp \left[\frac{i}{\hbar} \int_{-\infty}^{\infty} dt (L + Jq + i\epsilon q^2) \right] \quad (10.187)$$

Then

$$\frac{\delta Z[J]}{\delta J(t_1)} = N \int \mathcal{D}q \frac{i}{\hbar} \frac{\delta S[q, J]}{\delta J(t_1)} e^{\frac{i}{\hbar} S[q, J]} = N \int \mathcal{D}q \frac{i}{\hbar} q(t_1) e^{\frac{i}{\hbar} S[q, J]} \quad (10.188)$$

Therefore,

$$\left. \frac{\delta Z[J]}{\delta J(t_1)} \right|_{J=0} = N \int \mathcal{D}q \frac{i}{\hbar} q(t_1) e^{\frac{i}{\hbar} S[q]} = \frac{i}{\hbar} N \langle \Omega | \hat{q}(t_1) | \Omega \rangle \quad (10.189)$$

Repeating the functional differentiation, we have

$$\begin{aligned} \left. \frac{\delta^2 Z[J]}{\delta J(t_1) \delta J(t_2)} \right|_{J=0} &= N \int \mathcal{D}q \left(\frac{i}{\hbar} \right)^2 q(t_1) q(t_2) e^{\frac{i}{\hbar} S[q]} \\ &= \left(\frac{i}{\hbar} \right)^2 N \langle \Omega | T[\hat{q}(t_1) \hat{q}(t_2)] | \Omega \rangle \end{aligned} \quad (10.190)$$

In general, we have

$$\left. \frac{\delta^n Z[J]}{\delta J(t_1) \cdots \delta J(t_n)} \right|_{J=0} = \left(\frac{i}{\hbar} \right)^n N \langle \Omega | T[\hat{q}(t_1) \cdots \hat{q}(t_n)] | \Omega \rangle \quad (10.191)$$

Then we have

$$Z[J] = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{i}{\hbar} \right)^n \int dt_1 \cdots dt_n J(t_1) \cdots J(t_n) \langle \Omega | T[\hat{q}(t_1) \cdots \hat{q}(t_n)] | \Omega \rangle \quad (10.192)$$

Namely, $Z[J]$ is a generating functional of n -point time-ordered correlation functions or the Green's functions in the vacuum. It is known as the vacuum functional or the generating functional of the vacuum Green's functions. Using Eqs. (10.192) and (10.182), the time-ordered n -point function is expressed as

$$\langle \Omega | T[\hat{q}(t_1) \cdots \hat{q}(t_n)] | \Omega \rangle = \left(\frac{\hbar}{i} \right)^n \left[\frac{1}{Z[J]} \frac{\delta^n Z[J]}{\delta J(t_1) \cdots \delta J(t_n)} \right]_{J=0} \quad (10.193)$$

The usefulness of Eq. (10.193) lies in the fact that any scattering amplitude can be converted to n -point vacuum expectation values using reduction formula (see Sect. 11.5). Therefore, if one knows all the vacuum Green's functions, one can construct the S-matrix of the theory and therefore solve the problem. In quantum field theory, these correlation functions play a central role.

It is convenient to introduce another related functional $W(J)$ defined by

$$Z[J] = e^{\frac{i}{\hbar} W[J]} \quad \text{or,} \quad W[J] = -i\hbar \ln Z[J] \quad (10.194)$$

Then taking the functional derivative, we obtain

$$\left. \frac{\delta W[J]}{\delta J(t_1)} \right|_{J=0} = \left[\frac{\hbar}{i} \frac{1}{Z[J]} \frac{\delta Z[J]}{\delta J(t_1)} \right]_{J=0} = \langle \Omega | \hat{q}(t_1) | \Omega \rangle \equiv \langle \hat{q}(t_1) \rangle \quad (10.195)$$

where $\langle \cdots \rangle$ means the vacuum expectation value. We have learned that in the case of free particles and harmonic oscillators, the path integral for the transition amplitude is proportional to the exponential of the classical action [Eqs. (10.89), (10.123)]. For this reason $W[J]$ is also called the effective action. The advantage of using $W(J)$ over $Z(J)$ is that it gives statistical deviations from the mean values. This can be shown by further differentiating Eq. (10.195):

$$\begin{aligned} & \left. \frac{\hbar}{i} \frac{\delta^2 W[J]}{\delta J(t_1) \delta J(t_2)} \right|_{J=0} \\ &= \left(\frac{\hbar}{i} \right)^2 \left[\frac{1}{Z[J]} \frac{\delta^2 Z[J]}{\delta J(t_1) \delta J(t_2)} - \frac{1}{Z^2[J]} \frac{\delta Z[J]}{\delta J(t_1)} \frac{\delta Z[J]}{\delta J(t_2)} \right]_{J=0} \\ &= \langle T[\hat{q}(t_1) \hat{q}(t_2)] \rangle - \langle \hat{q}(t_1) \rangle \langle \hat{q}(t_2) \rangle \\ &= \langle T[(\hat{q}(t_1) - \langle \hat{q}(t_1) \rangle)(\hat{q}(t_2) - \langle \hat{q}(t_2) \rangle)] \rangle \end{aligned} \quad (10.196)$$

We recognize this is the second-order deviation from the mean. The third-order mean deviation is given by

$$\begin{aligned}
 & \left(\frac{\hbar}{i} \right)^2 \frac{\delta^3 W[J]}{\delta J(t_1) \delta J(t_2) \delta J(t_3)} \Big|_{J=0} \\
 &= \left(\frac{\hbar}{i} \right)^3 \left[\frac{1}{Z[J]} \frac{\delta^3 Z[J]}{\delta J(t_1) \delta J(t_2) \delta J(t_3)} - \frac{1}{Z^2[J]} \frac{\delta^2 Z[J]}{\delta J(t_1) \delta J(t_2)} \frac{\delta Z[J]}{\delta J(t_3)} \right. \\
 &\quad - \frac{1}{Z^2[J]} \frac{\delta^2 Z[J]}{\delta J(t_2) \delta J(t_3)} \frac{\delta Z[J]}{\delta J(t_1)} - \frac{1}{Z^2[J]} \frac{\delta^2 Z[J]}{\delta J(t_3) \delta J(t_1)} \frac{\delta Z[J]}{\delta J(t_2)} \\
 &\quad \left. + \frac{1}{Z^3[J]} \frac{\delta Z[J]}{\delta J(t_1)} \frac{\delta Z[J]}{\delta J(t_2)} \frac{\delta Z[J]}{\delta J(t_3)} \right] \Big|_{J=0} \\
 &= \langle T [(\hat{q}(t_1) - \langle \hat{q}(t_1) \rangle)(\hat{q}(t_2) - \langle \hat{q}(t_2) \rangle)(\hat{q}(t_3) - \langle \hat{q}(t_3) \rangle)] \rangle \quad (10.197)
 \end{aligned}$$

We can go on to higher order deviations from the mean. In quantum field theory, $W[J]$ is known as the generating functional for the connected vacuum Green's function, which will be discussed later in Sect. 11.6.7.

10.7

Connection with Statistical Mechanics

A typical application of the path integral outside the field of particle physics is to calculate the partition functions in statistical mechanics. Once it is given as a function of the temperature and volume (or pressure), the internal energy, entropy, pressure and other standard thermodynamical quantities are obtained by simply differentiating the partition function.

In a statistical ensemble of many particles that are confined in a volume V and in thermal equilibrium with a heat bath at temperature T , the probability of each particle being in a state i is given by

$$p_i = \frac{1}{Z} e^{-\beta E_i}, \quad \beta = \frac{1}{k_B T}, \quad (10.198a)$$

where $k_B = 8.617 \times 10^{-5} \text{ eV K}^{-1}$ is the Boltzmann constant. The partition function Z and the Helmholtz free energy $F = U - TS$ are defined for a system in thermal equilibrium, hence the Hamiltonian can be considered as time independent. They are defined by

$$Z = \sum_i e^{-\beta E_i} = \sum_i \langle i | e^{-\beta \hat{H}} | i \rangle = \text{Tr} [\exp(-\beta \hat{H})] \quad (10.198b)$$

$$F = -\frac{1}{\beta} \ln Z = -k_B T \ln Z \quad (10.198c)$$

Thermodynamical variables such as entropy and pressure can be derived from the partition function:

$$S = - \left. \frac{\partial F}{\partial T} \right|_V = k_B \beta^2 \frac{\partial F}{\partial \beta}, \quad P = - \frac{\partial F}{\partial V} \quad (10.198d)$$

Just to refresh our memory of thermodynamics, we give a derivation of the Helmholtz free energy in the boxed paragraph at the end of this chapter.

On the other hand the functional integral is used to calculate

$$\begin{aligned} G(x, t : y) &\equiv \langle x | t_i = 0, y \rangle = \langle x | \exp \left(-\frac{it}{\hbar} \hat{H} \right) | y \rangle \\ &\stackrel{\text{Eq. (10.52)}}{=} \int \mathcal{D}q \mathcal{D} \left(\frac{p}{2\pi\hbar} \right) \exp \left[\frac{i}{\hbar} \int_{t_i}^{t_f} dt \{ p \dot{q} - H(p, q) \} \right] \end{aligned} \quad (10.199)$$

Namely, if one can calculate the propagator Eq. (10.199), all one has to do to obtain the partition function is “to replace $it/\hbar$ with β (or better to say analytically continue to $t \rightarrow -i\hbar\beta$) and to take the trace”. Note, the trace can be taken in any basis. In particular, if we choose the basis in the Schrödinger picture, we can write

$$Z = \int dq \langle q | e^{-\beta \hat{H}} | q \rangle \quad (10.200)$$

If we introduce an imaginary time $\tau = it/\hbar$, denote the τ derivative of q as $\dot{q}$ and set $\tau_i = 0$ for simplicity, the partition function is given by

$$\begin{aligned} Z &= \int dq \langle q | e^{-\beta \hat{H}} | q \rangle = \int dq \int_{q(0)=q(\beta)} \mathcal{D}q \mathcal{D} \left(\frac{p}{2\pi\hbar} \right) \\ &\quad \times \exp \left[- \int_0^\beta d\tau \{ H(p, q) - ip\dot{q} \} \right] \end{aligned} \quad (10.201a)$$

For the special case $H(p, q) = p^2/2m + V(q)$, this becomes

$$Z = \int dq \int_{q(0)=q(\beta)} \overline{\mathcal{D}}q \exp \left[- \int_0^\beta d\tau \left\{ \frac{1}{2} m \dot{q}^2 + V(q) \right\} \right] \quad (10.201b)$$

where

$$\overline{\mathcal{D}}q = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar\epsilon} \right)^{N/2} dq_1 dq_2 \cdots dq_{N-1} \quad (10.201c)$$

We have already encountered this type of functional integral in Sect. 10.6.2 as the Wick rotation.

We add a few simple exercises in statistical mechanics just to remind the reader of some basic uses of the partition function.

Example 1: Free Particle Using Eq. (10.201b) and substituting $t_f - t_i = -i\hbar\beta$, $q_f = q_i$ in Eq. (10.95)

$$Z = [(10.201b)]_{\text{free particle}} = \int dq [(10.95)]_{t_f-t_i=-i\hbar\beta, q_f=q_i} = \left[\frac{m}{2\pi\hbar^2\beta} \right]^{1/2} \int dq \quad (10.202)$$

Then the Helmholtz free energy is given by

$$F = -\frac{1}{\beta} \ln Z = -\frac{1}{\beta} \ln \left[\left(\frac{m}{2\pi\hbar^2\beta} \right)^{1/2} V \right] \quad (10.203)$$

Using Eqs. (10.203) and (10.198d), we obtain the equation of state for the ideal gas $PV = k_B T$.

Example 2: Harmonic Oscillator

Problem 10.7

(1) Calculate the propagator for the harmonic oscillator at imaginary time $t = -i\hbar\beta$ to obtain

$$K(x : y) = \left(\frac{m\omega}{2\pi\hbar \sinh(\hbar\omega\beta)} \right)^{1/2} \times \exp \left[-\frac{m\omega}{2\hbar \sinh \hbar\omega\beta} \{ (x^2 + y^2) \cosh(\hbar\omega\beta) - 2xy \} \right] \quad (10.204)$$

Hint: Use (10.124).

(2) Calculate the partition function and the Helmholtz free energy. Show they are given by

$$Z = \frac{1}{2 \sinh \hbar\omega\beta/2}, \quad F = kT \ln \left(2 \sinh \frac{\hbar\omega}{2kT} \right) \quad (10.205)$$

(3) Using the energy eigenvalues of the harmonic oscillator, derive the partition function directly by calculating Eq. (10.198b), and show it agrees with Eq. (10.205).

(4) Derive the Planck blackbody radiation formula from Eq. (10.205).

Derivation of the Helmholtz Free Energy Having defined the partition function for an ensemble of particles in a volume V and in thermal equilibrium with a heat bath at temperature T , we will prove that the Helmholtz free energy is given by

$$F \equiv U - TS = -k_B T \ln Z \quad (10.206)$$

Note, E_i is a function of V , hence Z is a function of T and V . The internal energy U of a system is given as the expectation (or average) value of the energy. Denoting the probability of a particle in the state i as p_i and using $\sum E_i e^{-\beta E_i} = -\partial Z / \partial \beta$, the total energy is given by

$$U \equiv \sum p_i E_i = \frac{1}{Z} \sum E_i e^{-\beta E_i} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \Big|_V = -\frac{\partial \ln Z}{\partial \beta} \Big|_V \quad (10.207)$$

Similarly, the pressure is given as an energy change ΔU divided by an infinitesimal volume change ΔV under equitemperature conditions:

$$\begin{aligned} P &= -\frac{1}{Z} \sum \frac{\partial E_i}{\partial V} e^{-\beta E_i} = \frac{1}{\beta Z} \frac{\partial}{\partial V} \sum e^{-\beta E_i} \Big|_T = \frac{1}{\beta} \left(\frac{\partial \ln Z}{\partial V} \right)_T \\ &= -\frac{\partial F}{\partial V} \Big|_T \end{aligned} \quad (10.208)$$

When the volume of the system changes, two things happen. First, each energy level shifts slightly, and second, heat energy dQ flows in from the heat bath to keep the temperature constant. The energy increase dU is given by

$$dU = -P dV + dQ \quad (10.209)$$

Noting $d(\beta U) = \beta dU + U d\beta$, we can calculate dQ/T from

$$\frac{dQ}{T} = k_B \beta (dU + P dV) = k_B [d(\beta U) - U d\beta + \beta P dV] \quad (10.210)$$

Using Eqs. (10.207) and (10.208), we obtain

$$\frac{dQ}{T} = k_B \left[d(\beta U) + \frac{\partial \ln Z}{\partial \beta} d\beta + \frac{\partial \ln Z}{\partial V} dV \right] \quad (10.211)$$

The rhs is a total derivative. Therefore writing $dQ/T = dS$, $k_B \beta = 1/T$ and integrating we have

$$S = \frac{U}{T} + k_B \ln Z(T, V) \equiv \frac{U}{T} - \frac{F}{T}, \quad \text{or} \quad F = U - TS \quad (10.212)$$

which proves Eq. (10.206). Using $F = -(\ln Z)/\beta$, $U = -\partial \ln Z / \partial \beta \Big|_V = \partial(\beta F) / \partial \beta \Big|_V$, we can rewrite the entropy as

$$S = \frac{U - F}{T} = k_B \beta \left(\frac{\partial(\beta F)}{\partial \beta} \Big|_V - F \right) = k_B \beta^2 \frac{\partial F}{\partial \beta} \Big|_V = -\frac{\partial F}{\partial T} \Big|_V \quad (10.213)$$

From Eqs. (10.212), (10.207) and (10.198a), it can also be shown that

$$S = -k_B \sum_i p_i \ln p_i \quad (10.214)$$

which is the well-known expression for the entropy.

11

Path Integral Approach to Field Theory

Applications of the path integral method to field theory are given in this chapter. Scattering amplitudes of QED processes treated in Chap. 7 and the self-energy of the photon and electron already considered in Chap. 8 are rederived.

11.1

From Particles to Fields

Up to now we have discussed the one-particle system. We can extend the path integral method to systems of particles with many degrees of freedom. Consider a system of N particles having coordinates q^r ($r = 1, \dots, N$). Then the transition amplitude is expressed as

$$\langle t_f q_f | t_i q_i \rangle \Rightarrow \langle t_f q_f^1, \dots, q_f^N | t_i q_i^1, \dots, q_i^N \rangle = N \int \prod_r \mathcal{D}q^r(t) e^{\frac{i}{\hbar} S[q]} \quad (11.1)$$

$$q_f^r = q^r(t_f), \quad q_i^r = q^r(t_i)$$

The path integral on the rhs is now N -fold. The action has the form

$$S[q] = \int_{t_i}^{t_f} dt \sum_r L(q^r, \dot{q}^r) \quad (11.2)$$

and the integration is over all paths starting at q_i^r at $t = t_i$ and ending at q_f^r at $t = t_f$. The field $\phi(x) = \phi(t, \mathbf{x})$ can be considered as an infinite set of q^r indexed by the position $\mathbf{x}$, namely

$$\begin{aligned} r &\rightarrow \mathbf{x} \\ q^r &\rightarrow \phi(\mathbf{x}) \Delta \mathbf{x} \\ \sum_r &\rightarrow \int \end{aligned} \quad (11.3)$$

To define the path integral, instead of dividing just time into segments, we divide space and time, and the Minkowski space is broken down into four-dimensional cubes of volume ε^4 . Inside each cube at coordinate (t_j, x_k, y_l, z_m) , ϕ is taken to be

a constant:

$$\phi(j, k, l, m) = \bar{\phi}(t_j, x_k, y_l, z_m) \quad (11.4)$$

Derivatives are approximated by, for example,

$$\left. \frac{\partial \phi}{\partial x} \right|_{j,k,l,m} \simeq \frac{\bar{\phi}(t_j, x_k + \varepsilon, y_l, z_m) - \bar{\phi}(t_j, x_k, y_l, z_m)}{\varepsilon} \quad (11.5)$$

We replace the four indices (j, k, l, m) formally by the single index a and write

$$\mathcal{L} \{ \phi(t_j, x_k, y_l, z_m), \partial_\mu \phi(t_j, x_k, y_l, z_m) \} = \mathcal{L}(\phi_a, \partial_\mu \phi_a) = \mathcal{L}_a \quad (11.6a)$$

$$S[\phi] = \int d^4x \mathcal{L} \simeq \sum_{a=1}^{N^4} \varepsilon^4 \mathcal{L}_a \quad (11.6b)$$

The path integral of the field is defined as

$$\begin{aligned} \int \mathcal{D}\phi e^{\frac{i}{\hbar} S[\phi]} &= \int \mathcal{D}\phi \exp \left[\frac{i}{\hbar} \int d^4x \mathcal{L}(x) \right] \\ &= \lim_{N \rightarrow \infty} \int \prod_{a=1}^{N^4} d\phi_a \exp \left[\frac{i}{\hbar} \sum_a^N \varepsilon^4 \mathcal{L}_a \right] \\ &= \lim_{\substack{N \rightarrow \infty \\ \varepsilon \rightarrow 0}} \int \prod_{a=1}^{N^4} d\phi_a \exp \left[\frac{i}{\hbar} \varepsilon^4 \mathcal{L}_a \right] \end{aligned} \quad (11.7)$$

11.2

Real Scalar Field

11.2.1

Generating Functional

Now we consider the real scalar field $\phi(x) = \phi(t, \mathbf{x})$, which obeys the Klein-Gordon equation. From now on, we adopt natural units and omit $\hbar$. We continue to consider that the field is interacting with an external field $J(x)$. Then the classical equation that ϕ_{cl} satisfies is

$$[\partial_\mu \partial^\mu + m^2 - i\varepsilon] \phi_{\text{cl}} = J(x) \quad (11.8)$$

We have introduced a small negative imaginary part to the mass squared to ensure that the path integral is well behaved, as we discussed in Sect. 10.6.2. The Feynman's Green function is defined as one that obeys

$$\begin{aligned} [\partial_\mu \partial^\mu + m^2 - i\varepsilon] \Delta_F(x - y) &= -\delta^4(x - y) \\ \Delta_F(x - y) &= -[\partial_\mu \partial^\mu + m^2 - i\varepsilon]^{-1} \delta^4(x - y) = \int \frac{d^4q}{(2\pi)^4} \frac{e^{-iq \cdot (x-y)}}{q^2 - m^2 + i\varepsilon} \end{aligned} \quad (11.9)$$

The second equation is written to demonstrate that Δ_F is formally defined as the inverse of the Klein–Gordon operator. The prescription of how to handle the singularity of the Green’s function is dictated by the small negative imaginary part of the mass, which reproduces the recipe discussed in Sect. 6.6.2 to choose positive energy for particles going forward in time and negative energy for backward-going particles. The vacuum-to-vacuum amplitude is given by

$$Z_0[J] = N \int \mathcal{D}\phi \exp \left[i \int d^4x \left\{ \frac{1}{2} \{ \partial_\mu \phi \partial^\mu \phi - (m^2 - i\varepsilon) \phi^2 \} + J\phi \right\} \right] \quad (11.10)$$

We have attached the subscript “0” to Z to denote that we have not included the interaction of the scalar field apart from the artificial external source and hence ϕ can be considered free except for the external source J . The notation $Z[J]$ is saved for future use to denote the generating functional of interacting fields. We remind the reader that ϕ in the path integral is not the field satisfying Eq. (11.8), but the integration variable, which can have any value. We modify the action slightly, using the identity

$$\int d^4x \partial_\mu \phi \partial^\mu \phi = \int d^4x \partial_\mu (\phi \partial^\mu \phi) - \int d^4x \phi \partial_\mu \partial^\mu \phi \quad (11.11)$$

The first term is a surface integral and vanishes if $\phi \rightarrow 0$ at infinity. Then $Z_0[J]$ is expressed as

$$Z_0[J] = N \int \mathcal{D}\phi \exp \left[-i \int d^4x \left\{ \frac{1}{2} \phi (\partial_\mu \partial^\mu + m^2 - i\varepsilon) \phi - J\phi \right\} \right] \quad (11.12)$$

To re-formulate the path integral, we use the fact that ϕ is an integration variable and we are free to shift the reference path. We choose the reference path as the solution to Eq. (11.8):

$$\phi(x) \rightarrow \phi'(x) = \phi(x) + \phi_{\text{cl}}(x) = \phi(x) - \int d^4y \Delta_F(x - y) J(y) \quad (11.13)$$

The action in Eq. (11.12) can be rewritten as

$$\begin{aligned} & -i \int d^4x \left[\frac{1}{2} \phi' (\partial_\mu \partial^\mu + m^2 - i\varepsilon) \phi' - J\phi' \right] \\ &= -i \int d^4x \left[\frac{1}{2} \phi (\partial_\mu \partial^\mu + m^2 - i\varepsilon) \phi \right. \\ & \quad + \frac{1}{2} \phi (\partial_\mu \partial^\mu + m^2 - i\varepsilon) \phi_{\text{cl}} + \frac{1}{2} \phi_{\text{cl}} (\partial_\mu \partial^\mu + m^2 - i\varepsilon) \phi \\ & \quad \left. + \frac{1}{2} \phi_{\text{cl}} (\partial_\mu \partial^\mu + m^2 - i\varepsilon) \phi_{\text{cl}} - J\phi_{\text{cl}} - J\phi \right] \end{aligned} \quad (11.14)$$

The second and third terms on the rhs give the same contribution when integrated over and are canceled by the last term because of Eq. (11.8). The fourth term is $\frac{1}{2} J\phi_{\text{cl}}$ again because of Eq. (11.8). Then using Eq. (11.13), $\int d^4x J\phi_{\text{cl}}$ can be rewritten as $-\iint d^4x d^4y J(x) \Delta_F(x - y) J(y)$.

Collecting all the terms, we can now write the action as

$$S[\phi] = -\frac{i}{2} \int d^4x \phi (\partial_\mu \partial^\mu + m^2 - i\varepsilon) \phi - \frac{i}{2} \int d^4x d^4\gamma J(x) \Delta_F(x - \gamma) J(\gamma) \quad (11.15)$$

The second term does not contain ϕ , hence it can be taken out of the functional integral. $Z_0[J]$ now takes the form

$$Z_0[J] = N \exp \left[-\frac{i}{2} \int d^4x d^4\gamma J(x) \Delta_F(x - \gamma) J(\gamma) \right] \times \int \mathcal{D}\phi \exp \left[-\frac{i}{2} \int d^4x \phi (\partial_\mu \partial^\mu + m^2 - i\varepsilon) \phi \right] \quad (11.16)$$

The second factor, formally giving $(\det[i(\partial_\mu \partial^\mu + m^2 - i\varepsilon)])^{-1/2}$, is a constant, because it is integrated over all functions ϕ . We will calculate $\det[i(\partial_\mu \partial^\mu + m^2 - i\varepsilon)]$ later (see Sect. 11.2.2). Here we simply denote it by N^{-1} . So the final expression for $Z_0[J]$ is

$$Z_0[J] = \exp \left[-\frac{i}{2} \int d^4x d^4\gamma J(x) \Delta_F(x - \gamma) J(\gamma) \right] \quad (11.17)$$

Alternatively, the above equation can be derived using the generalized Gaussian integral. Note that just as integration of quadratic polynomials in an exponent is given by

$$\int_{-\infty}^{\infty} dq e^{-aq^2 + bq + c} = \sqrt{\frac{\pi}{a}} \exp \left(\frac{b^2}{4a} + c \right) \quad (11.18)$$

the N -dimensional Gaussian integral in such a form is generally expressed as

$$\int d\gamma_1 \cdots d\gamma_N e^{-\frac{1}{2} \gamma^T A \gamma + B^T \gamma + C} = \frac{(2\pi)^{N/2}}{\sqrt{\det[A]}} \exp \left(\frac{1}{2} B^T A^{-1} B + C \right) \quad (11.19)$$

where $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_N)$ and A and B are $N \times N$ and $N \times 1$ matrices. The generalization to the path integral of the bilinear scalar field action can easily be done. It is to be understood that the integration can be carried out in Euclidian space and analytically continued back to give the real space-time result so that formal application of the Gaussian formula Eq. (11.19) is valid for the oscillating integrand. Setting

$$A = i(\partial_\mu \partial^\mu + m^2 - i\varepsilon), \quad B = iJ, \quad C = 0 \quad (11.20)$$

the straightforward application of Eq. (11.19) results in

$$\begin{aligned} Z_0[J] &= N \int \mathcal{D}\phi \exp \left[-i \int d^4x \left\{ \frac{1}{2} \phi \{ \partial_\mu \partial^\mu + m^2 - i\varepsilon \} \phi - J\phi \right\} \right] \\ &= \frac{N D}{\sqrt{\det A}} \exp \left(\frac{1}{2} B^T A^{-1} B \right) \\ &= \exp \left[-\frac{i}{2} \int d^4x d^4\gamma J(x) \Delta_F(x - \gamma) J(\gamma) \right] \end{aligned} \quad (11.21)$$

where D is a diverging constant, which can be absorbed in the normalization constant. Equation (11.21) reproduces the result Eq. (11.17).

Normalization of the Path Integral Let us settle the question of the normalization. In most field theoretical questions, one is primarily interested in calculating expectation values of time-ordered operator products between the vacuum states, namely vacuum to vacuum transition rate. Consequently, the value of the normalization constant N is irrelevant, because we are only interested in the normalized transition amplitude. As $Z_0[J]$ is the vacuum-to-vacuum transition amplitude in the presence of the source J , it is sensible to normalize it to $Z_0[J = 0] = 1$. Or equivalently we can eliminate the awkward extra constant by dividing the path integral by $Z_0[0]$, which produces exactly the same constant. This is exactly what we have done in defining the n -point time-ordered correlation functions in Eq. (10.183). Then Eq. (11.10) must be rewritten as

$$\begin{aligned} Z_0[J] &= \frac{\int \mathcal{D}\phi \exp \left[-i \int d^4x \left\{ \frac{1}{2} \phi (\partial_\mu \partial^\mu \phi + m^2 - i\varepsilon) \phi - J\phi \right\} \right]}{\int \mathcal{D}\phi \exp \left[-i \int d^4x \left\{ \frac{1}{2} \phi (\partial_\mu \partial^\mu \phi + m^2 - i\varepsilon) \phi \right\} \right]} \\ &= \exp \left[-\frac{i}{2} \int d^4x d^4\gamma J(x) \Delta_F(x - \gamma) J(\gamma) \right] \end{aligned} \quad (11.22)$$

We have two equivalent expressions for the generating functional. The first, in the form of the path integral, is convenient for defining time-ordered n -point vacuum correlation functions. The second does not contain the path integral, and is convenient in calculating their functional derivatives.

11.2.2

Calculation of $\det A$

As we discussed in the previous section, the actual value of $\det A$ is not important as long as it does not contain any dynamical variables. It can always be factored away by proper normalization. However, it is helpful to become familiar with the actual calculation of the determinant, and we do it for the case of the most familiar Klein-Gordon operator. Perhaps it may be of interest to mention that $\det A$ in Euclidian time is the partition function (see discussion in Sect. 10.7) and that it has a practical use in statistical mechanics:

$$\det A = \det [i (\partial_\mu \partial^\mu + m^2) \delta^4(x - \gamma)] \quad (11.23)$$

As we will show, this is a divergent integral, but it does no harm as stated above. To evaluate it we go to Euclidian space ($x^0 = -ix_E^4$) and write

$$\begin{aligned} iS &= -\frac{i}{2} \int d^4x \phi(x) (\partial_\mu \partial^\mu + m^2) \phi(x) \\ &= -\frac{1}{2} \int d^4x_E \phi(x_E) (-\partial_\mu \partial^\mu + m^2)_E \phi(x_E) \\ &\equiv -\frac{1}{2} \int d^4x_E d^4\gamma_E \phi(x_E) A(x_E, \gamma_E) \phi(\gamma_E) \end{aligned} \quad (11.24a)$$

$$x_E = (x_E^1, x_E^2, x_E^3, x_E^4) = (x^1, x^2, x^3, ix^0) \quad (11.24b)$$

where x_E^4 is treated as a real number. Now

$$A(x_E, y_E) = (-\partial_\mu \partial^\mu + m^2)_E \delta^4(x_E - y_E) \quad (11.25)$$

is a matrix whose row and column indices are continuous variables. The following reformulation is useful:

$$\det A = \exp[\ln \det A] = \exp \left[\ln \prod_j \lambda_j \right] = \exp\{\text{Tr}[\ln A]\} \quad (11.26)$$

where the λ 's are the eigenvalues of the matrix A . The validity of the equalities can be most easily seen in the diagonalized form, noting that the trace is invariant if the matrix is diagonalized or not. The eigenvalues of the matrix A are easily seen in its Fourier transform

$$\begin{aligned} A &= (-\partial_\mu \partial^\mu + m^2)_E \delta^4(x_E - y_E) = (-\partial_\mu \partial^\mu + m^2)_E \int \frac{d^4 k_E}{(2\pi)^4} e^{ik_E \cdot (x_E - y_E)} \\ &= \int \frac{d^4 k_E}{(2\pi)^4} e^{ik_E \cdot (x_E - y_E)} (k_E^2 + m^2) \end{aligned} \quad (11.27a)$$

$$k_E = (k_E^1, k_E^2, k_E^3, k_E^4) = (k^1, k^2, k^3, -ik^0) \quad (11.27b)$$

We see A^{-1} is not singular, namely A has an inverse, since the eigenvalues are positive definite. Of course, we knew the inverse exists. It was already given by Eq. (11.9) as the Feynman propagator.

Next we prove that a function of matrices $f[A(x, y)]$ is the Fourier transform of the same function of the Fourier-transformed matrices, namely

$$f[A(x, y)] = \int \frac{dk}{2\pi} \frac{dk'}{2\pi} e^{i(kx - k'y)} f[A(k, k')] \quad (11.28a)$$

$$A(k, k') = \int dx dy e^{-i(kx - k'y)} A(x, y) \quad (11.28b)$$

Proof: For simplicity we consider in one dimension. In the above statement, the function is to be understood as expandable in a Taylor series:

$$\begin{aligned} f(A) &= c_0 + c_1 A + c_2 A^2 + \cdots c_n A^n + \cdots \\ \text{or } f[A(x, y)] &= c_0 + c_1 A(x, y) + c_2 \int dz_1 A(x, z_1) A(z_1, y) + \cdots \\ &\quad + c_n \int dz_1 \cdots dz_{n-1} A(x, z_1) A(z_1, z_2) \cdots A(z_{n-1}, y) + \cdots \end{aligned} \quad (11.29)$$

For $n = 2$ it is easy to show that

$$\begin{aligned}
 & \int dz_1 A(x, z_1) A(z_1, y) \\
 &= \int dz_1 \frac{dk}{2\pi} \frac{dk'}{2\pi} \frac{dp}{2\pi} \frac{dp'}{2\pi} e^{i(kx - k'z_1)} e^{i(pz_1 - p'y)} A(k, k') A(p, p') \\
 &= \int \frac{dk}{2\pi} \frac{dp'}{2\pi} e^{i(kx - p'y)} \int \frac{dk'}{2\pi} A(k, k') A(k', p') \quad (11.30)
 \end{aligned}$$

The n th term has $n - 1$ integrations, yielding $n - 1$ δ functions, and the result is

$$\begin{aligned}
 & \int dz_1 \cdots dz_{n-1} A(x, z_1) A(z_1, z_2) \cdots A(z_{n-1}, y) \\
 &= \int \frac{dk}{2\pi} \frac{dk'}{2\pi} e^{i(kx - k'y)} \int \frac{dk_1}{2\pi} \cdots \frac{dk_{n-1}}{2\pi} A(k, k_1) A(k_1, k_2) \cdots A(k_{n-1}, k') \\
 &= \int \frac{dk}{2\pi} \frac{dk'}{2\pi} e^{i(kx - k'y)} A^n(k, k') \quad (11.31)
 \end{aligned}$$

□

which proves Eq. (11.28a). If $A(k, k')$ is diagonal, namely $A(k, k') = A(k)(2\pi)\delta(k - k')$, or, equivalently, if $A(x, y)$ is a function of $x - y$, Eq. (11.28a) can be expressed as

$$f[A(x, y)] = \int \frac{dk}{2\pi} e^{ik(x-y)} f[A(k)] \quad (11.32)$$

then

$$\begin{aligned}
 \text{Tr}[f[A(x, y)]] &= \int dx dy \delta(x - y) \int \frac{dk}{2\pi} e^{ik(x-y)} f[A(k)] \\
 &= \int dx \int \frac{dk}{2\pi} f[A(k)] \quad (11.33)
 \end{aligned}$$

Now we go back to four dimensions. Equation (11.27) tells us that

$$A(k_E, k'_E) = (k_E^2 + m^2) \delta^4(k_E - k'_E) \quad (11.34)$$

Then setting $f(A) = \ln A$, the determinant A of Eq. (11.27) is given as

$$\begin{aligned}
 \det A &= \exp [\text{Tr} \ln A(x_E, y_E)] \\
 &= \exp \left[\int d^4 x_E d^4 y_E \delta^4(x_E - y_E) \left\{ \int \frac{d^4 k_E}{(2\pi)^4} e^{ik \cdot (x_E - y_E)} \ln(k_E^2 + m^2) \right\} \right] \\
 &= \exp \left[\int d^4 x_E \int \frac{d^4 k_E}{(2\pi)^4} \ln(k_E^2 + m^2) \right] \quad (11.35)
 \end{aligned}$$

Analytically continuing back to Minkowski space by setting $x_E^4 = ix^0$, $k_E^4 = -ik^0$, $\det A$ is given by

$$\det A = \exp \left[\int d^4 x \int \frac{d^4 k}{(2\pi)^4} \ln(-k^2 + m^2) \right] \quad (11.36)$$

As announced, this is a divergent quantity. The factor $\int d^4x$ arose from the translational symmetry. The action S stays constant for a class of the field which can be obtained by symmetry operations. Given a constant integrand the trace gives a diverging contribution, so it should be removed from the path integral. Our definition of normalization Eq. (11.22) takes care of this.

11.2.3

n -Point Functions and the Feynman Propagator

In Sect. 10.6, we derived a relation between n -point functions and generating functionals. The field theoretical version of an n -point function is

$$\langle 0 | T[\hat{\phi}(x_1) \cdots \hat{\phi}(x_n)] | 0 \rangle = \left(\frac{1}{i} \right)^n \frac{\delta^n Z_0[J]}{\delta J(x_1) \cdots \delta J(x_n)} \Big|_{J=0} \quad (11.37)$$

which is easily obtained by taking functional derivatives of the first expression of the generating functional Eq. (11.22). Note, here we express the vacuum state by $|0\rangle$, which is the vacuum of the free field, because we are using $Z_0[J]$, the generating functional of the free field. Looking at the second expression of Eq. (11.22), we see that n -point functions can be expressed in terms of Feynman propagators.

First we show that for a quadratic action such as that of the scalar field, the one-point vacuum function vanishes:

$$\begin{aligned} \frac{1}{i} \frac{\delta Z_0[J]}{\delta J(x_1)} &= \frac{1}{i} \frac{\delta}{\delta J(x)} \exp \left[-\frac{i}{2} \int d^4x d^4y J(x) \Delta_F(x-y) J(y) \right] \\ &= - \int d^4z J(z) \Delta_F(x_1 - z) Z_0[J] \end{aligned} \quad (11.38)$$

Using Eq. (11.37), this clearly gives

$$\langle 0 | T[\hat{\phi}(x)] | 0 \rangle = \langle 0 | \hat{\phi}(x) | 0 \rangle = \frac{1}{i} \frac{\delta Z_0[J]}{\delta J(x_1)} \Big|_{J=0} = 0 \quad (11.39)$$

Next, as the two-point correlation function is given by

$$\langle 0 | T[\hat{\phi}(x_1) \hat{\phi}(x_2)] | 0 \rangle = \left(\frac{1}{i} \right)^2 \frac{\delta^2 Z_0[J]}{\delta J(x_1) \delta J(x_2)} \Big|_{J=0} \quad (11.40)$$

we obtain by differentiating Eq. (11.38)

$$\begin{aligned} \left(\frac{1}{i} \right)^2 \frac{\delta^2 Z_0[J]}{\delta J(x_1) \delta J(x_2)} &= \left[i \Delta_F(x_1 - x_2) + \int d^4z_1 \Delta_F(x_1 - z_1) J(z_1) \right. \\ &\quad \left. \times \int d^4z_2 \Delta_F(x_2 - z_2) J(z_2) \right] Z_0[J] \end{aligned} \quad (11.41)$$

After setting $J = 0$, we obtain

$$\langle 0 | T[\hat{\phi}(x_1) \hat{\phi}(x_2)] | 0 \rangle = \left(\frac{1}{i} \right)^2 \frac{\delta^2 Z_0[J]}{\delta J(x_1) \delta J(x_2)} \Big|_{J=0} = i \Delta_F(x_1 - x_2) \quad (11.42)$$

which shows that the two-point correlation function is nothing other than the Feynman propagator. This agrees with Eq. (6.124), of course, which was derived using the canonical quantization method.

11.2.4

Wick's Theorem

We continue with n -point functions. We note that the generating functional $Z_0[J]$ can be Taylor expanded around J :

$$\begin{aligned} Z_0[J] &= 1 + \int d^4x_1 \left. \frac{\delta Z_0[J]}{\delta J(x_1)} \right|_{J=0} J(x_1) \\ &\quad + \frac{1}{n!} \int d^4x_1 \cdots d^4x_n \left. \frac{\delta^n Z_0[J]}{\delta J(x_1) \cdots \delta J(x_n)} \right|_{J=0} J(x_1) \cdots J(x_n) + \cdots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int d^4x_1 \cdots d^4x_n i^n \langle 0 | T[\hat{\phi}(x_1) \cdots \hat{\phi}(x_n)] | 0 \rangle J(x_1) \cdots J(x_n) \end{aligned} \quad (11.43)$$

where we have used Eq. (11.37). Another way of expressing the generating functional is

$$\begin{aligned} Z_0[J] &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(-\frac{i}{2} \int d^4x d^4y J(x) \Delta_F(x-y) J(y) \right)^k \\ &= 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{i}{2} \right)^k \int d^4x_1 \cdots d^4x_{2k} \Delta_{12} \Delta_{34} \cdots \Delta_{2k-1,2k} J_1 J_2 \cdots J_{2k} \end{aligned} \quad (11.44a)$$

$$\Delta_{ij} = \Delta_F(x_i - x_j) \quad (11.44b)$$

Since, $Z_0[J]$ contains only even powers of J , it is obvious that

$$\langle 0 | T[\hat{\phi}(x_1) \cdots \hat{\phi}(x_{2k-1})] | 0 \rangle = \frac{1}{i^{2k-1}} \left. \frac{\delta^{2k-1} Z_0[J]}{\delta J(x_1) \cdots \delta J(x_{2k-1})} \right|_{J=0} = 0 \quad (11.45a)$$

If $n = 2k$ is even, the n -point function is a sum of products of two-point functions:

$$\langle 0 | T[\hat{\phi}(x_1) \cdots \hat{\phi}(x_{2k})] | 0 \rangle = \sum_{\text{perms}} i \Delta_F(x_{p_1} - x_{p_2}) \cdots i \Delta_F(x_{p_{2k-1}} - x_{p_{2k}}) \quad (11.45b)$$

where the sum is taken for all the permutations. This is one version of Wick's theorem.

Note the factor $1/2^k k!$ has disappeared from the expression for the $2k$ -point function. This can be seen as follows. n -Functional derivatives are completely symmetric and there are $n!$ permutations (different orders) of n variables $x_1 \cdots x_n$. When $n = 2k$, the coefficient of $J_1 \cdots J_{2k}$ in Eq. (11.44a) $\Delta_{12} \Delta_{34} \cdots \Delta_{2k-1, 2k}$ has 2^k pairwise symmetry because $\Delta_{ij} = \Delta_{ji}$ and another $k!$ -fold symmetry because the order of the product of $k \Delta_{ij}$'s can be arbitrary. That leaves $(2k)!/(2^k k!) = (2k-1)!! = (2k-1) \cdots 3 \cdot 1$ irreducible combinations in Eq. (11.45b). As an explicit example we consider the case $n = 4$. We then should have $4!/(2^2 2!) = 3$ terms for the 4-point function, which is given by

$$\begin{aligned} \langle 0 | T[\hat{\phi}(x_1) \hat{\phi}(x_2) \hat{\phi}(x_3) \hat{\phi}(x_4)] | 0 \rangle &= \frac{1}{2^2 2!} \sum_{4! \text{ permutations}} i \Delta_{i_1 i_2} i \Delta_{i_3 i_4} \\ &= -\Delta_F(x_1 - x_2) \Delta_F(x_3 - x_4) - \Delta_F(x_1 - x_3) \Delta_F(x_2 - x_4) \\ &\quad - \Delta_F(x_1 - x_4) \Delta_F(x_2 - x_3) \end{aligned} \quad (11.46)$$

Equations (11.43) and (11.45) are important because, as we will learn later (Sect. 11.5), the transition rate of any physical process can be expressed in terms of n -point vacuum correlation functions using the Lehmann–Symanzik–Zimmermann (LSZ) reduction formula.

So far we have treated only the generating functionals of the free Lagrangian. The relation between the n -point vacuum correlation functions and the functional derivatives is equally valid if the Lagrangian includes interaction terms.

11.2.5

Generating Functional of Interacting Fields

So far we have dealt with the action of up to quadratic terms or exactly soluble cases. In general, the action includes higher order polynomials or transcendental functions that are not analytically soluble. Consider adding a potential $V(\phi)$ to the free Lagrangian:

$$\mathcal{L} = \frac{1}{2} [\partial_\mu \phi \partial^\mu \phi - (m^2 - i\varepsilon) \phi^2] - V(\phi) \quad (11.47)$$

It is not possible to solve this problem exactly even when $V \ll 1$. But as long as V is small, we can solve it perturbatively in the following way. We retain the external source as before and write

$$S[J] = S_0[\phi, J] - \int d^4x V(\phi) \quad (11.48)$$

$$S_0[\phi, J] = \int d^4x \left(\frac{1}{2} [\partial_\mu \phi \partial^\mu \phi - (m^2 - i\varepsilon) \phi^2] + J(x) \phi(x) \right)$$

$$Z[J] = N \int \mathcal{D}\phi e^{-i \int d^4z V(\phi(z))} e^{i S_0[\phi, J]} \quad (11.49)$$

We now use the relation

$$\frac{\delta}{\delta J(x)} Z_0[J] = \frac{\delta}{\delta J(x)} \int \mathcal{D}\phi e^{i S_0[\phi, J]} = \int \mathcal{D}\phi i \phi(x) e^{i S_0[\phi, J]} \quad (11.50)$$

This relation is also true for any power of ϕ , or for a potential $V(\phi)$ that is Taylor expandable:

$$\therefore \int \mathcal{D}\phi V(\phi) e^{iS_0[\phi, J]} = V\left(\frac{1}{i} \frac{\delta}{\delta J(x)}\right) Z_0[J] \quad (11.51)$$

This relation allows us to take the V -dependent factor in Eq. (11.49) out of the path integral. Consequently, after exponentiation, we obtain

$$\begin{aligned} Z[J] &= N \int \mathcal{D}\phi e^{-i \int d^4z V(\phi(z))} e^{iS_0[\phi, J]} \\ &= \exp\left[-i \int d^4z V\left(\frac{1}{i} \frac{\delta}{\delta J(z)}\right)\right] Z_0[J] \end{aligned} \quad (11.52)$$

where $Z_0[J]$ is the vacuum functional without interaction. Using the expression (11.22) for $Z_0[J]$

$$\begin{aligned} Z[J] &= \exp\left[-i \int d^4z V\left(\frac{1}{i} \frac{\delta}{\delta J(z)}\right)\right] \\ &\quad \times \exp\left[-\frac{i}{2} \int d^4x d^4y J(x) \Delta_F(x-y) J(y)\right] \end{aligned} \quad (11.53)$$

If V is small we can expand the first exponential in powers of V , and the problem can be treated perturbatively order by order.

11.3

Electromagnetic Field

11.3.1

Gauge Fixing and the Photon Propagator

The electromagnetic field is gauge invariant and this produces some difficulties in the path integral. The action of the free electromagnetic field is given by

$$\begin{aligned} S &= \int d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right] = -\frac{1}{4} \int d^4x (\partial_\mu A_\nu - \partial_\nu A_\mu) (\partial^\mu A^\nu - \partial^\nu A^\mu) \\ &= -\frac{1}{4} \int d^4x [\partial_\mu A_\nu \partial^\mu A^\nu - \partial_\nu A_\mu \partial^\mu A^\nu - \partial_\mu A_\nu \partial^\nu A^\mu + \partial_\nu A_\mu \partial^\nu A^\mu] \end{aligned} \quad (11.54)$$

By partial integration

$$\begin{aligned} S &= -\frac{1}{4} \int d^4x (-2A_\rho \partial_\mu \partial^\mu A^\rho + 2A^\mu \partial_\mu \partial_\nu A^\nu) \\ &= \frac{1}{2} \int d^4x A^\mu(x) [g_{\mu\nu} \partial_\rho \partial^\rho - \partial_\mu \partial_\nu] A^\nu(x) \\ &= \frac{1}{2} \int \frac{d^4k}{(2\pi)^4} \tilde{A}_\mu(k) [-g^{\mu\nu} k^2 + k^\mu k^\nu] \tilde{A}_\nu(k) \end{aligned} \quad (11.55)$$

where $\tilde{A}(k)$ is the Fourier transform of $A(x)$. The action is gauge invariant, which means it does not change if the electromagnetic field undergoes gauge transformation:

$$\tilde{A}_\mu(k) \rightarrow \tilde{A}'_\mu(k) = \tilde{A}_\mu(k) + k_\mu \tilde{\lambda}(k) \quad (11.56)$$

or $\mathcal{L} = \text{constant}$ for a vast class of λ , which makes the path integral badly divergent:

$$\int \mathcal{D}A e^{iS[A]} \sim \prod_{a=1}^{N^4} \prod_{\mu=0}^3 dA_a^\mu \times \text{constant} \rightarrow \infty \quad (11.57)$$

where a stands for t_j, x_k, y_l, z_m . Namely, $A_\mu = \partial_\mu \lambda(x)$ is equivalent to $A_\mu(x) = 0$ and should not be included in the path integral. The divergence problem is linked to another difficulty: the operator $\mathcal{A}_{\mu\nu} = g_{\mu\nu} \partial_\rho \partial^\rho - \partial_\mu \partial_\nu$ does not have an inverse. This means that the equation of the Feynman propagator

$$\begin{aligned} [g_{\mu\nu} \partial_\rho \partial^\rho - \partial_\mu \partial_\nu] D_F^{\nu\sigma}(x - y) &= \delta_\mu^\sigma \delta^4(x - y) \\ \text{or } [-g_{\mu\nu} k^2 + k_\mu k_\nu] \tilde{D}_F^{\nu\sigma}(k) &= \delta_\mu^\sigma \end{aligned} \quad (11.58)$$

has no solution. To see this, put $D_F^{\nu\rho}(k) = A g^{\nu\rho} + B k^\nu k^\rho$, insert it in Eq. (11.58) and solve for A and B . The remedy we discussed in the Lorentz covariant treatment of the canonical quantization was to add a gauge-fixing term to the Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{EM}} + \mathcal{L}_{\text{gf}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\lambda} (\partial_\mu A^\mu)^2 \quad (11.59)$$

In particular, we saw that setting $\lambda = 1$ in the canonical quantization produces the Maxwell equation, which satisfies the Lorentz condition

$$\partial_\mu \partial^\mu A^\nu = 0 \quad (11.60a)$$

$$\langle \phi | \partial_\mu A^\mu | \psi \rangle = 0 \quad (11.60b)$$

where $\langle \phi |$ and $| \psi \rangle$ are physical state vectors. The same prescription works for the path integral. If we add the gauge-fixing term to the Lagrangian, it is no longer gauge invariant and there is no redundant integration which produces unwanted divergence. Following the same reformulating procedure as led to Eq. (11.55), we obtain a new action

$$\begin{aligned} S &= \int d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\lambda} (\partial_\mu A^\mu)^2 \right] \\ &= \frac{1}{2} \int d^4x A_\mu(x) \left[g^{\mu\nu} \partial_\rho \partial^\rho - \left(1 - \frac{1}{\lambda} \right) \partial^\mu \partial^\nu \right] A_\nu(x) \\ &= \frac{1}{2} \int d^4k \tilde{A}_\mu(k) \left[-g^{\mu\nu} k^2 + \left(1 - \frac{1}{\lambda} \right) k^\mu k^\nu \right] \tilde{A}_\nu(k) \end{aligned} \quad (11.61)$$

The operator

$$\mathcal{A}_{\mu\nu} = g_{\mu\nu} \partial_\rho \partial^\rho - \left(1 - \frac{1}{\lambda} \right) \partial_\mu \partial_\nu \quad (11.62)$$

now has an inverse and the Feynman propagator for the electromagnetic field (the photon) is a solution to the equation

$$\left[g_{\mu\nu} \partial_\rho \partial^\rho - \left(1 - \frac{1}{\lambda} \right) \partial_\mu \partial_\nu - i\varepsilon \right] D_F^{\nu\sigma}(x-y) = \delta_\mu^\sigma \delta^4(x-y) \quad (11.63)$$

where a small negative imaginary part is added to keep the path integral well behaved at large time. In momentum space it is given by

$$\tilde{D}_{F\mu\nu} = -\frac{1}{k^2 + i\varepsilon} \left[g_{\mu\nu} - (1-\lambda) \frac{k_\mu k_\nu}{k^2} \right] \quad (11.64)$$

The constant λ is arbitrary. Note, setting $\lambda \rightarrow \infty$ reproduces the gauge invariant Lagrangian, but no propagator exists in this case. Choosing $\lambda = 1$ reproduces the familiar Feynman propagator, which we will adopt from now on. In this case

$$\begin{aligned} D_{F\mu\nu}(x-y) &= \int \frac{d^4 k}{(2\pi)^4} e^{-ik \cdot (x-y)} \frac{-g_{\mu\nu}}{k^2 - i\varepsilon} = -g_{\mu\nu} \Delta_F(x-y)|_{m^2=0} \\ \partial_\rho \partial^\rho D_{F\mu\nu}(x-y) &= g_{\mu\nu} \delta^4(x-y) \end{aligned} \quad (11.65)$$

The addition of the gauge-fixing term has made the action and hence the path integral gauge dependent, consequently the propagator is also gauge dependent, as is reflected by the arbitrary choice of λ in Eq. (11.64). However, they all lead to the same physics because in the transition amplitude $D_{F\mu\nu}$ appears coupled to conserved currents, which fulfills $\partial_\mu j^\mu = 0 \rightarrow k_\mu \tilde{j}^\mu(k) = 0$. As a consequence, the λ -dependent term in the propagator always drops out and the physically observable transition amplitudes are gauge independent.

11.3.2

Generating Functional of the Electromagnetic Field

Once we have fixed the action for the electromagnetic field, the generating functional can be constructed in a similar manner as the scalar field. We add the external field to the Lagrangian and reformulate the action in terms of the external field and the Feynman propagator. The bilinear electromagnetic field term can be integrated out just like the scalar field in deriving Eq. (11.22) and again we have two alternative expressions for the generating functional, just as for the scalar field:

$$\begin{aligned} Z_0^{\text{EM}}[J] &= N \int \mathcal{D}A \exp \left[i \int d^4 x \left\{ -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2 + J_\mu A^\mu \right\} \right] \\ &= \frac{\int \mathcal{D}A \exp \left[i \int d^4 x \left\{ \frac{1}{2} A^\mu \mathcal{A}_{\mu\nu} A^\nu + J_\mu A^\mu \right\} \right]}{\int \mathcal{D}A \exp \left[i \int d^4 x \left\{ \frac{1}{2} A^\mu \mathcal{A}_{\mu\nu} A^\nu \right\} \right]} \\ &= \exp \left[-\frac{i}{2} \int d^4 x d^4 y J^\mu(x) D_{F\mu\nu}(x-y) J^\nu(y) \right] \end{aligned} \quad (11.66)$$

where $\mathcal{A}_{\mu\nu}$ is defined in Eq. (11.62). Note this is the generating functional of the free electromagnetic field. That of the interacting field will be discussed in Sect. 11.6.

The two-point vacuum expectation value can be derived from the above equation and again proves to be the Feynman propagator of the photon:

$$\begin{aligned}\langle 0|T[A_\mu(x)A_\nu(y)]|0\rangle &= \left(\frac{1}{i}\right)^2 \frac{\delta^2 Z_0[J]}{\delta J^\mu(x)\delta J^\nu(y)} \\ &= i D_{F\mu\nu}(x-y) = \int \frac{d^4 k}{(2\pi)^4} \frac{-i g_{\mu\nu}}{k^2 - i\epsilon} e^{-ik\cdot(x-y)}\end{aligned}\quad (11.67)$$

11.4

Dirac Field

The path integral deals only with classical numbers. However, we cannot write the Lagrangian for the anticommuting fermion fields using normal or classical c-numbers. We need the notion of anticommuting classical numbers, mathematically known as Grassmann variables.

11.4.1

Grassmann Variables

Polynomial We use G-variables or G-numbers to denote Grassmann variables or Grassmann generators. Two independent Grassmann variables η_i and η_j are anticommuting, i.e.

$$\eta_i \eta_j = -\eta_j \eta_i \quad (11.68)$$

In particular,

$$\eta_i^2 = 0 \quad (i \text{ not summed}) \quad (11.69)$$

This means the G-numbers cannot have a length or an absolute value. Because of Eq. (11.68), any function of Grassmann variables including a c-number has a simple Taylor expansion:

$$f(\eta) = a + b\eta \quad (11.70)$$

where a and b are functions of ordinary c-numbers. For instance, an exponential function of G-numbers is

$$e^{a\eta} = 1 + a\eta \quad (11.71)$$

Complex Conjugate A complex Grassmann number is defined by using two real G-numbers ξ and η

$$\zeta = \xi + i\eta, \quad \zeta^* = \xi - i\eta \quad (11.72)$$

If we require $\zeta^* \zeta$ to be real, i.e.

$$(\zeta^* \zeta)^* = \zeta^* \zeta = -\zeta \zeta^* \quad (11.73)$$

then

$$\begin{aligned} (\xi - i\eta)(\xi + i\eta) &= \xi^2 + \eta^2 + 2i\xi\eta = 2i\xi\eta = (2i\xi\eta)^* \\ \therefore (\xi\eta)^* &= -\xi\eta \rightarrow (\xi\eta)^* = \eta\xi \end{aligned} \quad (11.74)$$

Therefore

$$\begin{aligned} \zeta_1 \zeta_2 &= (\xi_1 + i\eta_1)(\xi_2 + i\eta_2) = (\xi_1 \xi_2 - \eta_1 \eta_2) + i(\xi_1 \eta_2 + \eta_1 \xi_2) \\ \zeta_2^* \zeta_1^* &= (\xi_2 - i\eta_2)(\xi_1 - i\eta_1) = \xi_2 \xi_1 - \eta_2 \eta_1 - i\xi_2 \eta_1 - i\eta_2 \xi_1 \\ &= [(\xi_1 \xi_2 - \eta_1 \eta_2) + i(\xi_1 \eta_2 + \eta_1 \xi_2)]^* \\ \therefore (\zeta_1 \zeta_2)^* &= \zeta_2^* \zeta_1^* \end{aligned} \quad (11.75)$$

This shows that the complex conjugate of a G-number product behaves like the hermitian conjugate of matrix products.

Derivative As Grassmann variables are anticommuting, two distinguishable derivatives are defined. A left derivative would give

$$\frac{\partial}{\partial \eta_i} (\eta_j \eta_k) = \left(\frac{\partial \eta_j}{\partial \eta_i} \right) \eta_k - \eta_j \left(\frac{\partial \eta_k}{\partial \eta_i} \right) = \delta_{ij} \eta_k - \delta_{ik} \eta_j \quad (11.76)$$

whereas a right derivative would give

$$\frac{\partial}{\partial \eta_i} (\eta_j \eta_k) = \eta_j \left(\frac{\partial \eta_k}{\partial \eta_i} \right) - \left(\frac{\partial \eta_j}{\partial \eta_i} \right) \eta_k = \delta_{ik} \eta_j - \delta_{ij} \eta_k \quad (11.77)$$

Thus we have to be careful as to which one is used. In our discussion, we assume we use the left derivatives unless otherwise noted. The left derivative can be generalized to multivariables:

$$\frac{\partial}{\partial \eta_b} (\eta_a \cdots \eta_b \cdots \eta_c) = (-1)^P (\eta_a \cdots \overset{\eta_b \text{ missing}}{\langle} \cdots \eta_c) \quad (11.78)$$

where P is the number of permutations to move η_b to the left. We note that two successive derivatives anticommute just like G-numbers. Namely,

$$\frac{\partial}{\partial \eta_i} \frac{\partial}{\partial \eta_j} = - \frac{\partial}{\partial \eta_j} \frac{\partial}{\partial \eta_i} \quad (11.79a)$$

which also means

$$\left(\frac{\partial}{\partial \eta} \right)^2 = 0 \quad (11.79b)$$

This relation implies that there is no inverse to the derivative. This can be seen by multiplying the defining equation for the inverse

$$\frac{\partial}{\partial \eta} \left(\frac{\partial}{\partial \eta} \right)^{-1} \phi(\eta) = \phi(\eta) \quad (11.80)$$

from the left by $\partial/\partial \eta$

$$\frac{\partial}{\partial \eta} \frac{\partial}{\partial \eta} \left(\frac{\partial}{\partial \eta} \right)^{-1} \phi(\eta) = 0 \cdot \left(\frac{\partial}{\partial \eta} \right)^{-1} \phi(\eta) = \frac{\partial}{\partial \eta} \phi(\eta) \quad (11.81)$$

which is a contradiction. Thus the inverse does not exist.

The conventional commutation relation between the derivative and the coordinate takes the form

$$\left\{ \frac{\partial}{\partial \eta_i}, \eta_j \right\} = \frac{\partial}{\partial \eta_i} \eta_j + \eta_j \frac{\partial}{\partial \eta_i} = \delta_{ij} \quad (11.82)$$

Integral Because there is no inverse to the derivative, the integral in the space of G-variables can only be defined formally, which reserves some general properties of ordinary integration. Since any function can be expressed as a polynomial in the form of Eq. (11.70), there are only two types of integral:

$$\int d\eta, \quad \int d\eta \eta \quad (11.83)$$

All other integrals are zero. To derive rules for the integrals, we require that Grassmann variables respect translational invariance. Let ξ be a constant G-number and change the integration variable from η to $\eta' = \eta - \xi$. If translational invariance holds, the two integrals should give the same result:

$$\int d\eta' = \int d\eta \quad (11.84a)$$

$$\int d\eta' \eta' = \int d\eta (\eta - \xi) = \int d\eta \eta + \xi \int d\eta = \int d\eta \eta \quad (11.84b)$$

$$\therefore \int d\eta = 0 \quad (11.85a)$$

To evaluate $\int d\eta \eta$, we use the fact that a G-number commutes with a pair of other G-numbers:

$$\int d\eta \eta \xi = \xi \int d\eta \eta$$

This means $\int d\eta \eta$ is a c-number and we can define it to take a value 1. Namely

$$\int d\eta \eta = 1 \quad (11.85b)$$

The definition can be justified if we consider the fact that the integral of a total derivative $\int \partial$ must vanish if we ignore surface terms, and, the definite integral being a constant, its derivative must also vanish, namely

$$\int \partial = \partial \int = 0 \quad (11.86)$$

Since the definitions of the integral Eqs. (11.85a) and (11.85b) mean that for any function we have

$$\int d\eta_i f(\eta_j) = \int d\eta_i (a + b\eta_j) = b\delta_{ij} = \frac{d}{d\eta_i} f(\eta_j) \quad (11.87)$$

Namely, the differential and integral operations are identical. Hence

$$\int \partial = \partial \int = \partial^2 = 0 \quad (11.88)$$

which confirms Eq. (11.86).

Because the integral is the same as the differential, the Jacobian of the variable transformation differs from that of ordinary integration. If we redefine the variables of the integration from η to $\eta' = a\eta$, then

$$\int d\eta f(\eta) = \frac{\partial f(\eta)}{\partial \eta} = a \frac{\partial f(\eta'/a)}{\partial \eta'} = a \int d\eta' f\left(\frac{\eta'}{a}\right) \quad (11.89)$$

The last equality is opposite to what happens to an ordinary integral. Namely, the Jacobian that appears in the redefinition of the Grassmann variables is the inverse of what one would expect for the ordinary integral.

Multiple Integral In multiple integrals the integration variable has to be brought first next to the integration measure:

$$\int d\eta_1 \int d\eta_2 \eta_1 \eta_2 = - \int d\eta_1 \int d\eta_2 \eta_2 \eta_1 = - \int d\eta_1 \eta_1 = -1 \quad (11.90)$$

Problem 11.1

Consider N -variable transformation $\xi_i = \sum_j U_{ij} \eta_j$. Write

$$\begin{aligned} \int \prod_{i=1}^N d\eta_i f(\eta_1, \dots, \eta_N) &= J_G \int \prod_{i=1}^N d\xi_i g(\xi_1, \dots, \xi_N) \\ g(\xi) &= f(U^{-1}\xi) \end{aligned} \quad (11.91)$$

Using Eq. (11.85) and anticommutativity of the G -variables, prove

$$J_G = \frac{\partial(\xi_1, \dots, \xi_N)}{\partial(\eta_1, \dots, \eta_N)} = \det[U_{ij}] \quad (11.92)$$

δ function A delta function in the space of Grassmann variables is written as

$$\delta(\eta - \eta') = \eta - \eta' \quad (11.93)$$

Proof: As $f(\eta) = a + b\eta$

$$\begin{aligned} \int d\eta(\eta - \eta')f(\eta) &= \int d\eta(\eta - \eta')(a + b\eta) \\ &= \frac{\partial}{\partial\eta}(\eta - \eta')(a + b\eta) = a + b\eta' = f(\eta') \end{aligned} \quad (11.94)$$

□

Note, as $\delta(\eta) = \eta$, $\delta(a\eta) = a\eta = \{\partial(a\eta)/\partial\eta\}\eta$. Generally

$$\delta(f(\eta)) = \frac{\partial f(\eta)}{\partial\eta} \delta(\eta) \quad (11.95)$$

which is consistent with the Jacobian defined by Eq. (11.92).

An integral representation of the δ -function is obtained by noting that if η is a G-variable

$$\frac{1}{i} \int d\eta e^{i\eta(\xi - \xi')} = \frac{1}{i} \int d\eta [1 + i\eta(\xi - \xi')] = \xi - \xi' = \delta(\xi - \xi') \quad (11.96)$$

Gaussian Integral Next we treat the Gaussian integral of the Grassmann variables. First we prove that for an integral of complex Grassmann variables

$$\int d\eta^* d\eta e^{-\eta^* \eta} = 1 \quad (11.97)$$

The above equality can easily be proved because

$$\int d\eta^* d\eta e^{-\eta^* \eta} = \int d\eta^* d\eta (1 - \eta^* \eta) = \frac{\partial}{\partial\eta^*} \frac{\partial}{\partial\eta} (1 - \eta^* \eta) = 1 \quad (11.98)$$

Next we generalize the Gaussian integral to n complex variables. Denoting a Grassmann vector $\boldsymbol{\eta}^T = (\eta_1, \eta_2, \dots, \eta_n)$, we consider an integral

$$I_G \equiv \int \prod_{i=1}^n d\eta_i^* d\eta_i \exp[-\boldsymbol{\eta}^\dagger \mathbf{A} \boldsymbol{\eta}] \quad (11.99)$$

where $\mathbf{A}$ is an $n \times n$ matrix and we assume $\mathbf{A}^\dagger = \mathbf{A}$, because in the path integral we are interested only in hermitian matrices. Then it can be diagonalized by using a unitary matrix U :

$$\mathbf{D} = \mathbf{U} \mathbf{A} \mathbf{U}^\dagger \quad \text{or} \quad \mathbf{A} = \mathbf{U}^\dagger \mathbf{D} \mathbf{U} \quad (11.100)$$

where $\mathbf{D}$ is a diagonal matrix and is expressed as

$$D_{jk} = \delta_{jk} \lambda^j \quad (11.101)$$

Changing the variables to $\xi = U\eta$

$$\begin{aligned}
 I_G &= \int \prod_{i=1}^n d\eta_i^* d\eta_i \exp[-\xi^\dagger U A U^\dagger \xi] \\
 &= \det U^\dagger \det U \int \prod_{i=1}^n d\xi_i^* d\xi_i \exp[-\xi^\dagger D \xi] \\
 &= \int \left(\prod_{i=1}^n d\xi_i^* d\xi_i \right) \exp \left[-\sum_j \lambda^j \xi_j^* \xi_j \right] \\
 &= \prod_{j=1}^n \left(\int d\xi_j^* d\xi_j \exp \left[-\lambda^j \xi_j^* \xi_j \right] \right) \\
 &= \prod_{j=1}^n \left(\int d\xi_j^* d\xi_j \left[1 - \lambda^j \xi_j^* \xi_j \right] \right) = \prod_{i=j}^n \lambda^j = \det A \quad (11.102)
 \end{aligned}$$

We have used the fact that the Jacobian is unity:

$$J = (\det U^\dagger)(\det U) = \det(U^\dagger U) = 1 \quad (11.103)$$

We also used the fact that $\int d\xi^* d\xi$ is a c-number and can be moved freely in the product. Note, the integral is $2n$ dimensional but the matrix is n dimensional. This formula is to be contrasted with the n -fold Gaussian integral over ordinary c-number complex variables [Eq. (10.148)]:

$$I_c = \int dz_1^* dz_1 \cdots dz_n^* dz_n e^{-z^\dagger A z} = \frac{(2\pi i)^n}{\det A} \quad (11.104)$$

We note that the determinant of the Grassmann Gaussian integral appears in the numerator whereas that of the ordinary Gaussian appears in the denominator.

The generalization to the functional integral is straightforward. Change

$$\begin{aligned}
 \eta_i &\rightarrow \eta(x) \\
 \prod_{i=1}^n d\eta_i^* d\eta_i &\rightarrow \mathcal{D}\bar{\eta}(x) \mathcal{D}\eta(x) \\
 \eta^\dagger A \eta &\rightarrow \int d^4x d^4\gamma \eta^*(x) A(x, \gamma) \eta(\gamma)
 \end{aligned} \quad (11.105)$$

and we obtain

$$\int \mathcal{D}\bar{\eta}(x) \mathcal{D}\eta(x) \exp \left[-i \int d^4x d^4\gamma \eta^*(x) A(x, \gamma) \eta(\gamma) \right] = \det iA \quad (11.106)$$

Next, we evaluate an integral of the form

$$I_F = \int \prod_{i=1}^n d\eta_i^* d\eta_i \eta_k \eta_l^* \exp[-\eta^\dagger A \eta] \quad (11.107)$$

It is clear that an integral containing one or an odd number of G -variables in front of the exponential function vanishes. After unitary transformation the integral becomes [see Eq. (11.102)]

$$I_f = \prod_{j=1}^n \int d\xi_j^* d\xi_j (U_{ks}^{-1} \xi_s \xi_t^* U_{tl}) \left[1 - \lambda_j \xi_j^* \xi_j \right] \quad (11.108)$$

From this equation we see only terms of the form

$$\int d\xi_j^* d\xi_j \xi_j \xi_j^* (1 - \lambda_j \xi_j^* \xi_j) = \int d\xi_j^* d\xi_j \xi_j \xi_j^* \quad (11.109)$$

gives non-zero contribution. Namely, $\xi_s \xi_t^*$ with $s = t = j$ survive and λ_j is lost. Therefore, the effect of inserting an extra $\eta_k \eta_l^*$ in the integral I_G is to make the change

$$\lambda_j \rightarrow U_{kj}^{-1} U_{jl} \quad (11.110)$$

or equivalently to multiply I_G [Eq. (11.102)] by

$$U_{kj}^{-1} \frac{1}{\lambda_j} U_{jl} = (A^{-1})_{kl} \quad (11.111)$$

Thus the end result becomes

$$I_F = \int \prod_{i=1}^n d\eta_i^* d\eta_i \eta_k \eta_l^* \exp[-\boldsymbol{\eta}^\dagger \mathbf{A} \boldsymbol{\eta}] = \det \mathbf{A} (A^{-1})_{kl} \quad (11.112)$$

Inserting another pair $\eta_m \eta_n^*$ in the integrand would yield a second factor $(A^{-1})_{mn}$.

Gaussian Integral of Real Grassmann Variables We can obtain similar expressions for real-variable Gaussian integrals with the orthogonal matrix $\mathbf{A}$:

$$\int d\eta_1 \cdots d\eta_n e^{-\frac{1}{2} \boldsymbol{\eta}^\top \mathbf{A} \boldsymbol{\eta}} = \begin{cases} \sqrt{\det \mathbf{A}} & n \text{ even} \\ 0 & n \text{ odd} \end{cases} \quad (11.113)$$

The proof is more complicated than that for complex variables because the matrix is antisymmetric in this case. As real-variable integrals are used less than complex-variable integrals we leave the proof as an exercise.

Problem 11.2

Let $\boldsymbol{\eta}$ be an n -dimensional real Grassmann vector. Expand the Gaussian integral in a Taylor series:

$$e^{-\frac{1}{2} \boldsymbol{\eta}^\top \mathbf{A} \boldsymbol{\eta}} = 1 - \frac{1}{2} \boldsymbol{\eta}^\top \mathbf{A} \boldsymbol{\eta} + \frac{\left(\frac{1}{2}\right)^2}{2!} (-\boldsymbol{\eta}^\top \mathbf{A} \boldsymbol{\eta})^2 + \cdots + \frac{\left(\frac{1}{2}\right)^j}{j!} (-\boldsymbol{\eta}^\top \mathbf{A} \boldsymbol{\eta})^j + \cdots \quad (11.114)$$

Since the j th term contains $2j$ G -variables, the series does not go beyond $j > n/2$.

Show that

(1) For an n -dimensional integral I_n

$$I_2 = \int d\eta_1 d\eta_2 e^{-\frac{1}{2}\eta^T A \eta} = \frac{1}{2}(A_{12} - A_{21}) = A_{12} = \sqrt{\det A} \quad (11.115a)$$

$$I_3 = 0 \quad (11.115b)$$

$$\begin{aligned} I_4 &= \int d\eta_1 \cdots d\eta_4 e^{-\frac{1}{2}\eta^T A \eta} \\ &= \frac{\left(\frac{1}{2}\right)^2}{2} \sum_{ijkl=1}^4 A_{ij} A_{kl} \int d\eta_1 \cdots d\eta_4 \eta_i \eta_j \eta_k \eta_l \\ &= (A_{12} A_{34} - A_{13} A_{24} + A_{14} A_{23}) = \sqrt{\det A} \end{aligned} \quad (11.115c)$$

(2) Prove Eq. (11.113).*

Problem 11.3

Prove that a similar equation to Eq. (11.112) holds for ordinary numbers, namely for an N -dimensional real vector $\mathbf{x} = (x_1, \dots, x_N)$ and real symmetric and positive definite $D \times D$ matrix A

$$\begin{aligned} &\int \prod_{i=1}^N dx^i x_{k_1} \cdots x_{k_n} e^{-\frac{1}{2}\mathbf{x}^T A \mathbf{x}} \\ &= \begin{cases} \frac{(2\pi)^{N/2}}{\sqrt{\det A}} (A_{k_1 k_2}^{-1} \cdots A_{k_{n-1} k_n}^{-1} + \text{permutations}) & n = \text{even} \\ 0 & n = \text{odd} \end{cases} \end{aligned} \quad (11.116)$$

11.4.2

Dirac Propagator

We are now ready to express the anticommuting Dirac field in terms of Grassmann variables. We may define a Grassmann field $\psi(x)$ in terms of any set of orthonormal basis functions:

$$\psi(x) = \sum_j \psi_j \phi_j(x) \quad (11.117)$$

The basis function $\phi_j(x)$ are ordinary c -number functions, but the coefficients ψ_j are Grassmann numbers. Then the Dirac field can be expressed by taking $\phi(x)$ to be the four-component Dirac spinors $u_r(p)e^{-ip \cdot x}$ and $v_r(p)e^{-ip \cdot x}$. Using the

Grassmann–Dirac field, the 2-point correlation function for the Dirac field is expressed as

$$\langle 0 | T[\hat{\psi}(x_1) \hat{\bar{\psi}}(x_2)] | 0 \rangle = \frac{\int \mathcal{D}\bar{\psi} \mathcal{D}\psi \psi(x_1) \bar{\psi}(x_2) \exp \left[i \int d^4x \bar{\psi}(i\partial - m)\psi \right]}{\int \mathcal{D}\bar{\psi} \mathcal{D}\psi \exp \left[i \int d^4x \bar{\psi}(i\partial - m)\psi \right]} \quad (11.118)$$

It is also to be understood that the integral can be analytically continued in the time variable complex plane to ensure the good behavior of the integral or equivalently a small imaginary part $i\varepsilon$ is implicit to produce the right propagator. Comparing Eq. (11.118) with Eq. (11.112), we obtain

$$\begin{aligned} \langle 0 | T[\hat{\psi}(x) \hat{\bar{\psi}}(y)] | 0 \rangle &= i S_F(x - y) = [-i(i\partial - m)]^{-1} \delta^4(x - y) \\ S_F(x - y) &= \int \frac{d^4p}{(2\pi)^4} \frac{e^{-ip \cdot (x-y)}}{\not{p} - m + i\varepsilon} \\ &= \int \frac{d^4p}{(2\pi)^4} e^{-ip \cdot (x-y)} \frac{\not{p} + m}{p^2 - m^2 + i\varepsilon} \end{aligned} \quad (11.119)$$

11.4.3

Generating Functional of the Dirac Field

Using the Feynman propagator, the generating functional of the Dirac field is written as

$$\begin{aligned} Z_0^{\text{Dirac}}[\bar{\eta}, \eta] &= \frac{\int \mathcal{D}\bar{\psi} \mathcal{D}\psi \exp \left[i \int d^4x \left\{ \bar{\psi}(i\partial - m)\psi + \bar{\eta}\psi + \bar{\psi}\eta \right\} \right]}{\int \mathcal{D}\bar{\psi} \mathcal{D}\psi \exp \left[i \int d^4x \left\{ \bar{\psi}(i\partial - m)\psi \right\} \right]} \\ &= \exp \left[-i \int d^4x d^4y \bar{\eta}(x) S_F(x - y) \eta(y) \right] \end{aligned} \quad (11.120)$$

The derivation of the second line from the first is completely in parallel with that of the scalar field [see Eq. (11.22)]. Then the n -point correlation function can be derived by successive differentiation of the generating functional. We have to be careful only about the sign change depending on the position of the Grassmann variables. The operation $-i\delta/\delta\bar{\eta}$ pulls down ψ but to pull down $\bar{\psi}$, $i\delta/\delta\eta$ is necessary because it has to be anticommutated through $\bar{\psi}$ before it can differentiate η . If we apply two derivatives to Z_0 to pull down $\psi(x)\bar{\psi}(y)$:

$$\begin{aligned} &\left(-\frac{1}{i} \frac{\delta}{\delta\eta(y)} \right) \left(\frac{1}{i} \frac{\delta}{\delta\bar{\eta}(x)} \right) \int \mathcal{D}\bar{\psi} \mathcal{D}\psi \exp \left[i \int d^4z \left\{ \bar{\psi}(i\partial - m)\psi + \bar{\eta}\psi + \bar{\psi}\eta \right\} \right] \\ &= \left(-\frac{1}{i} \frac{\delta}{\delta\eta(y)} \right) \int \mathcal{D}\bar{\psi} \mathcal{D}\psi \psi(x) \exp \left[i \int d^4z \left\{ \bar{\psi}(i\partial - m)\psi + \bar{\eta}\psi + \bar{\psi}\eta \right\} \right] \\ &= - \int \mathcal{D}\bar{\psi} \mathcal{D}\psi \psi(x) \bar{\psi}(y) \exp \left[i \int d^4z \left\{ \bar{\psi}(i\partial - m)\psi + \bar{\eta}\psi + \bar{\psi}\eta \right\} \right] \end{aligned} \quad (11.121)$$

Note the minus sign was added because $\delta/\delta\eta$ has to go through ψ , which was pulled down by the first derivative. If the ordering of the operators and of the derivatives is the same, no extra sign appears, namely

$$\left(\frac{1}{i}\right)^2 \frac{\delta^2 Z_0[\bar{\eta}, \eta]}{\delta\eta(\gamma)\delta\bar{\eta}(x)} \Big|_{\bar{\eta}=\eta=0} = \langle 0 | T[\hat{\psi}(x)\hat{\bar{\psi}}(\gamma)] | 0 \rangle \quad (11.122)$$

This formula can be extended to a general $2n$ -point function:

$$\begin{aligned} & \left(\frac{1}{i}\right)^{2n} \frac{\delta^{2n} Z_0[\bar{\eta}, \eta]}{\delta\eta(\gamma_1) \cdots \delta\eta(\gamma_n) \delta\bar{\eta}(x_1) \cdots \delta\bar{\eta}(x_n)} \Big|_{\bar{\eta}=\eta=0} \\ &= \langle 0 | T[\hat{\psi}(x_1) \cdots \hat{\psi}(x_n) \hat{\bar{\psi}}(\gamma_1) \cdots \hat{\bar{\psi}}(\gamma_n)] | 0 \rangle \end{aligned} \quad (11.123)$$

To prove that we have the right phase in the above equation, we first show that if the derivatives are ordered in exactly the same way as the operator products, the relation between the two is expressed as

$$\begin{aligned} & \left(-\frac{1}{i}\right)^n \left(\frac{1}{i}\right)^n \frac{\delta^{n+n} Z_0[\bar{\eta}, \eta]}{\delta\eta(\gamma_n) \cdots \delta\eta(\gamma_1) \delta\bar{\eta}(x_n) \cdots \delta\bar{\eta}(x_1)} \Big|_{\bar{\eta}=\eta=0} \\ &= \langle 0 | T[\hat{\bar{\psi}}(\gamma_n) \cdots \hat{\bar{\psi}}(\gamma_1) \hat{\psi}(x_n) \cdots \hat{\psi}(x_1)] | 0 \rangle \end{aligned} \quad (11.124)$$

The proof goes as follows. The first derivative $\delta/\delta\bar{\eta}(x_1)$ pulls $\psi(x_1)$ down from Z_0 and produces no sign change. The second derivative has to be anticommutated with $\psi(x_1)$ before pulling down $\psi(x_2)$ from Z_0 , producing one extra -1 . But it is canceled by moving the pulled-down $\psi(x_2)$ to the left of $\psi(x_1)$. Similarly, the i th derivative has to be anticommutated i times before it pulls $\psi(x_i)$ down from Z_0 , but the generated sign is canceled exactly by moving the pulled down ψ back to the i th position. The argument is similar for $\delta/\delta\eta$ except for an extra -1 to go to the right of $\bar{\psi}$ in the exponent before the derivative is applied to η .

Rearranging the ordering of the operators to move all $\bar{\psi}(\gamma_i)$ to the right of $\bar{\psi}(x_j)$ produces the minus sign of $(-1)^{n^2}$. Hence the total number of -1 's becomes $n^2 + n$ and $(-1)^{n^2+n} = 1$. Renaming x_i and γ_i in ascending order recovers Eq. (11.123).

11.5

Reduction Formula

11.5.1

Scalar Fields

Here we introduce the LSZ (Lehmann–Symanzik–Zimmermann) reduction formula [260], which relates the scattering matrix to n -point vacuum correlation functions. We first consider the case of scalar particles. It is straightforward to generalize to other types of particles. We assume the weak asymptotic condition of

Eq. (6.26) with Z set to be 1, namely

$$\lim_{t \rightarrow \mp\infty} \langle f | \hat{\phi}(x) | i \rangle = \langle f | \hat{\phi}_{\text{out}}^{\text{in}}(x) | i \rangle \quad (11.125)$$

The “in” and “out” states are free states, i.e. on the mass shell, and can be described by the free field. For scalar fields, the free-field operator can be expressed as

$$\hat{\phi}_{\text{out}}^{\text{in}}(x) = \int \frac{d^3 q}{(2\pi)^3 \sqrt{2\omega_q}} \left[a_{\text{out}}^{\text{in}}(q) e^{-iq \cdot x} + a_{\text{out}}^{\text{in}\dagger}(q) e^{iq \cdot x} \right] \quad (11.126)$$

The scattering matrix is given as

$$S_{fi} = \langle \text{out} : f | i : \text{in} \rangle = \langle \text{in} : f | S | i : \text{in} \rangle = \langle \text{out} : f | S | i : \text{out} \rangle \quad (11.127)$$

Then we can define the in (out) state vector of definite momentum as

$$\begin{aligned} |q_{\text{out}}^{\text{in}}\rangle &= \sqrt{2\omega} a_{\text{out}}^{\text{in}\dagger}(q) |\Omega\rangle = -i\sqrt{2\omega} \int d^3 x e_q(x) \overleftrightarrow{\partial}_t \hat{\phi}_{\text{out}}^{\text{in}}(x) |\Omega\rangle \\ &= \lim_{\substack{t \rightarrow -\infty \\ t \rightarrow +\infty}} \left[-i \int d^3 x f_q(x) \overleftrightarrow{\partial}_t \hat{\phi}(x) |\Omega\rangle \right] \end{aligned} \quad (11.128)$$

where

$$e_q = \frac{e^{-iq \cdot x}}{\sqrt{2\omega}}, \quad f_q = e^{-iq \cdot x}, \quad A \overleftrightarrow{\partial}_t B = A \partial_t B - (\partial_t A) B \quad (11.129)$$

Here, although not explicitly written, we assume that the operators appear sandwiched between the free states, as they should be consistent with the condition Eq. (11.125). Next, we wish to convert the three-dimensional integral into a Lorentz-invariant four-dimensional integral. To do that we use the identity

$$\left(\lim_{t \rightarrow \infty} - \lim_{t \rightarrow -\infty} \right) \int d^3 x A(x) = \int_{-\infty}^{\infty} dt \partial_t \int d^3 x A(x) \quad (11.130)$$

Consider, for example, an S-matrix element for the scattering of m particles with momenta $(q_1, q_2, \dots, q_m)$ to n particles of momenta $(p_1, \dots, p_n)$. We want to extract the free field $\hat{\phi}_{\text{out}}^{\text{in}}(x)$ from it. First note

$$\langle \text{out} : \beta | \alpha : \text{in} \rangle = \langle \text{out} : \beta | \sqrt{2\omega} a_{\text{in}}^{\dagger}(q) | \alpha - q : \text{in} \rangle \quad (11.131)$$

where we have used $\langle \beta |$, $|\alpha\rangle$ and $|\alpha - q\rangle$ to denote states containing n , m and $m - 1$ particles, respectively. In $|\alpha - q\rangle$ a particle having momentum q is taken out. Then using Eqs. (11.128) and (11.130), Eq. (11.131) becomes

$$\begin{aligned} \langle \text{out} : \beta | \alpha : \text{in} \rangle &= \lim_{t \rightarrow -\infty} \left[-i \int d^3 x f_q(x) \overleftrightarrow{\partial}_t \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle \right] \\ &= i \int d^4 x \partial_t \left[f_q(x) \overleftrightarrow{\partial}_t \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle \right] \\ &\quad - \lim_{t \rightarrow \infty} \left[i \int d^3 x f_q(x) \overleftrightarrow{\partial}_t \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle \right] \end{aligned} \quad (11.132)$$

The second term is just a creation operator of the “out” state, or an annihilation operator if it acts on the left, therefore

$$\begin{aligned}
 \langle \text{out} : \beta | a_{\text{out}}^\dagger(q) \sqrt{2\omega_q} &= \langle \text{out} : \beta - p_j, p_j | a_{\text{out}}^\dagger(q) \sqrt{2\omega_q} \\
 &= \langle \text{out} : \beta - p_j | a_{\text{out}}(p_j) a_{\text{out}}^\dagger(q) \sqrt{2\omega_j} \sqrt{2\omega_q} \\
 &= (\sqrt{2\omega_j 2\omega_q}) \langle \text{out} : \beta - p_j | \{ [a_{\text{out}}(p_j), a_{\text{out}}^\dagger(q)] + a_{\text{out}}^\dagger(q) a_{\text{out}}(p_j) \} \\
 &= 2\omega_j (2\pi)^3 \delta^3(\mathbf{p}_j - \mathbf{q}) \langle \text{out} : \beta - \mathbf{q} | \\
 &\quad + \sqrt{2\omega_j} \sqrt{2\omega_q} \langle \text{out} : \beta - p_j | a_{\text{out}}^\dagger(q) a_{\text{out}}(p_j) \rangle
 \end{aligned} \tag{11.133}$$

The process can be repeated until $a_{\text{out}}^\dagger(q)$ goes to the left end and destroys the vacuum. The term vanishes if $\langle \beta |$ contains no particles with momentum q . This term is called a “disconnected diagram”, because one particle emerges unaffected by the scattering process and is hence disconnected from the rest of the particles.

The first term on the rhs of Eq. (11.132) gives

$$\begin{aligned}
 &= i \int d^4x \partial_t \left[f_q(x) \overleftrightarrow{\partial}_t \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle \right] \\
 &= -i \int d^4x \{ \partial_t^2 f_q(x) - f_q(x) \partial_t^2 \} \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle \\
 &= i \int d^4x \{ (-\nabla^2 + m^2) f_q(x) + f_q(x) \partial_t^2 \} \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle \\
 &= i \int d^4x \{ f_q(x) (\partial_t^2 - \nabla^2 + m^2) \} \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle
 \end{aligned} \tag{11.134}$$

where in deriving the third line we have used the fact that the plane wave function $f_q(x)$ satisfies the Klein–Gordon equation and the last line was obtained by partial integration. We now have a Lorentz-invariant expression. Collecting terms, we find

$$\begin{aligned}
 \langle \text{out} : \beta | \alpha : \text{in} \rangle &= (\text{disconnected diagram}) \\
 &\quad + i \int d^4x f_q(x) (\partial_\mu \partial^\mu + m^2) \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle
 \end{aligned} \tag{11.135}$$

We have extracted one particle from the abstract S-matrix and converted it into a matrix element of the field operator $\hat{\phi}$. Next, we extract the second particle with momentum p from the “out” state. Disregarding the disconnected diagram, we get

$$\begin{aligned}
 \langle \text{out} : \beta | \hat{\phi}(x) | \alpha - q : \text{in} \rangle &= \langle \text{out} : \beta - p | a_{\text{out}}(p) \sqrt{2\omega_p} \hat{\phi}(x) | \alpha - q : \text{in} \rangle \\
 &= \langle \text{out} : \beta - p | \hat{\phi}(x) \sqrt{2\omega_p} a_{\text{in}}(p) | \alpha - q : \text{in} \rangle \\
 &\quad + \langle \text{out} : \beta - p | \{ a_{\text{out}}(p) \sqrt{2\omega_p} \hat{\phi}(x) - \hat{\phi}(x) \sqrt{2\omega_p} a_{\text{in}}(p) \} | \alpha - q : \text{in} \rangle \\
 &= \langle \text{out} : \beta - p | \hat{\phi}(x) | \alpha - q - p : \text{in} \rangle \\
 &\quad + i \int d^3x' \left[\lim_{t' \rightarrow \infty} \left\{ f_p^*(x') \overleftrightarrow{\partial}_{t'} \langle \text{out} : \beta - p | \hat{\phi}(x') \hat{\phi}(x) | \alpha - q : \text{in} \rangle \right\} \right] \\
 &\quad - i \int d^3x' \left[\lim_{t' \rightarrow -\infty} \left\{ f_p^*(x') \overleftrightarrow{\partial}_{t'} \langle \text{out} : \beta - p | \hat{\phi}(x) \hat{\phi}(x') | \alpha - q : \text{in} \rangle \right\} \right]
 \end{aligned}$$

(11.136a)

Disregarding again the disconnected diagram (the first term), the second and the third terms can be combined to give

$$\begin{aligned}
 &= i \int d^4 x' \partial_t' \left[f_p^*(x') \overleftrightarrow{\partial_t'} \langle \text{out} : \beta - p | T[\hat{\phi}(x') \hat{\phi}(x)] | \alpha - q : \text{in} \rangle \right] \\
 &= i \int d^4 x' f_p^*(x') (\partial_\mu' \partial^{\mu'} + m^2) \langle \text{out} : \beta - p | T[\hat{\phi}(x') \hat{\phi}(x)] | \alpha - q : \text{in} \rangle
 \end{aligned}
 \tag{11.136b}$$

where $T[\dots]$ is a time-ordered product. The process can be repeated on both sides until we extract all the particles. We can always choose the “in” and “out” momenta so that there are no disconnected diagrams. Then the final expression for the scattering matrix for m incoming particles and n outgoing particles is given by

$$\begin{aligned}
 \langle \beta | S | \alpha \rangle &= i^{m+n} \int \prod_{j=1}^n d^4 x'_j \prod_{k=1}^m d^4 x_k f_{p_j}^*(x'_j) (\overrightarrow{\partial_\mu' \partial^{\mu'}} + m^2)_j \\
 &\quad \times \langle \Omega | T \left[\hat{\phi}(x'_1) \cdots \hat{\phi}(x'_n) \hat{\phi}(x_1) \cdots \hat{\phi}(x_m) \right] | \Omega \rangle (\overleftarrow{\partial_\mu \partial^\mu} + m^2)_k f_{q_k}(x_k)
 \end{aligned}
 \tag{11.137}$$

where we have arranged the incoming part of the operation to the right of the correlation function, although how it is ordered is irrelevant for the case of boson fields. This is called the LSZ reduction formula to convert a S-matrix to a vacuum expectation value of an $m+n$ -point correlation function. Note, we have derived the above formula in the Heisenberg representation. The field $\hat{\phi}$ is that of interacting particles, which contains all the information on the interaction and the vacuum state is also that of the full, interacting theory.

As the n -point time-ordered correlation function, in turn, can be expressed as a functional derivative [see Eq. (11.37)], the S-matrix expressed in terms of generating functionals is given by

$$\begin{aligned}
 \langle \beta | S | \alpha \rangle &= i^{m+n} \int \prod_{j=1}^n d^4 x'_j \prod_{k=1}^m d^4 x_k f_{p_j}^*(x'_j) (\overrightarrow{\partial_\mu' \partial^{\mu'}} + m^2)_j \\
 &\quad \times \left(\frac{1}{i} \right)^{m+n} \frac{\delta^{m+n} Z_0[J]}{\delta J(x_1) \cdots \delta J(x_{m+n})} \bigg|_{J=0} (\overleftarrow{\partial_\mu \partial^\mu} + m^2)_k f_{q_k}(x_k)
 \end{aligned}
 \tag{11.138}$$

11.5.2

Electromagnetic Field

Since the free electromagnetic field (in the Lorentz gauge) obeys the Klein–Gordon equation with mass = 0 and is expressed as

$$A_{\text{out}}^\mu(x) = \int \frac{d^3k}{(2\pi)^3 \sqrt{2\omega_k}} \sum_\lambda \left[\varepsilon^\mu(k, \lambda) a_{\text{out}}^{\text{in}}(k, \lambda) e^{-ik \cdot x} + \varepsilon^{\mu*}(k, \lambda) a_{\text{out}}^{\dagger}(k, \lambda) e^{ik \cdot x} \right] \quad (11.139)$$

the reduction formula of the electromagnetic field is almost the same as that of the scalar field, except for the addition of the polarization vector. Namely,

$$\begin{aligned} \langle \beta | S | \alpha \rangle &= i^{m+n} \int \prod_{j=1}^n d^4x'_j \prod_{k=1}^m d^4x_k f_{p_j}^*(x'_j) \varepsilon_{\rho_j}^*(p_j, \lambda_j) (\overrightarrow{\partial_\mu' \partial^{\mu'}})_j \\ &\times \langle \Omega | T \left[\hat{A}^{\rho_1}(x'_1) \cdots \hat{A}^{\rho_n}(x'_n) \hat{A}^{\sigma_1}(x_1) \cdots \hat{A}^{\sigma_m}(x_m) \right] | \Omega \rangle \\ &\times (\overleftarrow{\partial_\mu \partial^\mu})_k \varepsilon_{\sigma_k}(q_k, \lambda_k) f_{q_k}(x_k) \end{aligned} \quad (11.140)$$

We can express the vacuum correlation function in terms of functional derivatives. Referring to Eq. (11.37),

$$\begin{aligned} \langle \beta | S | \alpha \rangle &= \int \prod_{j=1}^n d^4x'_j \prod_{k=1}^m d^4x_k f_{p_j}^*(x'_j) \varepsilon_{\rho_j}^*(p_j, \lambda_j) (\overrightarrow{\partial_\mu' \partial^{\mu'}})_j \\ &\times \frac{\delta^{m+n} Z[J]}{\delta J^{\rho_1}(x'_1) \cdots \delta J^{\rho_n}(x'_n) \delta J^{\sigma_1}(x_1) \cdots \delta J^{\sigma_m}(x_m)} \\ &\times (\overleftarrow{\partial_\mu \partial^\mu})_k \varepsilon_{\sigma_k}(q_k, \lambda_k) f_{q_k}(x_k) \end{aligned} \quad (11.141)$$

11.5.3

Dirac Field

It is straightforward to apply the reduction formula to the Dirac field. The difference between the scalar field is that we have to be careful about the ordering of the operators since the Dirac field is a four-component spinor and anticommuting. As before, we first postulate the asymptotic condition as

$$\lim_{\substack{t \rightarrow \infty \\ t \rightarrow -\infty}} \langle f | \hat{\psi}(x) | i \rangle = \langle f | \hat{\psi}_{\text{out}}(x) | i \rangle \quad (11.142)$$

where the free Dirac field can be expressed as

$$\hat{\psi}_{\text{in}}(x) = \int \frac{d^3p}{(2\pi)^3 \sqrt{2E_p}} \left[U_p^r(x) a_{\text{in}}(p, r) + V_p^r(x) b_{\text{in}}^\dagger(p, r) \right] \quad (11.143a)$$

$$U_p^r(x) = u_r(p) e^{-ip \cdot x}, \quad V_p^r(x) = v_r(p) e^{ip \cdot x} \quad (11.143b)$$

$$(i\vec{\partial} - m)\hat{\psi}_{\text{in}}(x) = 0, \quad \hat{\bar{\psi}}_{\text{in}}(x)(-i\overleftarrow{\partial} - m) = 0 \quad (11.143c)$$

$$a_{\text{in}}(p, r) = \int d^3x \frac{\overline{U}_p^r(x)}{\sqrt{2E_p}} \gamma^0 \hat{\psi}_{\text{in}}(x), \quad a_{\text{in}}^\dagger(p, r) = \int d^3x \hat{\bar{\psi}}_{\text{in}}(x) \gamma^0 \frac{U_p^r(x)}{\sqrt{2E_p}} \quad (11.143d)$$

$$b_{\text{in}}(p, r) = \int d^3x \hat{\bar{\psi}}_{\text{in}}(x) \gamma^0 \frac{V_p^r(x)}{\sqrt{2E_p}}, \quad b_{\text{in}}^\dagger(p, r) = \int d^3x \frac{\overline{V}_p^r(x)}{\sqrt{2E_p}} \gamma^0 \hat{\psi}_{\text{in}}(x) \quad (11.143e)$$

The reduction formula is obtained following the same logical flow as the scalar field:

$$\begin{aligned} \langle \text{out} : p | q : \text{in} \rangle &= \langle \text{out} : p | \sqrt{2E_q} a_{\text{in}}^\dagger(q, r) | \Omega : \text{in} \rangle \\ &= -\langle \text{out} : p | \{ a_{\text{out}}^\dagger(q, r) \sqrt{2E_q} - \sqrt{2E_q} a_{\text{in}}^\dagger(q, r) \} | \Omega : \text{in} \rangle \\ &\quad + \langle \text{out} : p | \sqrt{2E_q} a_{\text{out}}^\dagger(q, r) | \Omega : \text{in} \rangle \\ &= -(\lim_{t \rightarrow \infty} - \lim_{t \rightarrow -\infty}) \int d^3x \langle \text{out} : p | \hat{\bar{\psi}}(x) | \Omega : \text{in} \rangle \gamma^0 U_q^r(x) \\ &\quad + \text{disconnected part} \\ &= i \int d^4x \langle \text{out} : p | i \partial_0 \{ \hat{\bar{\psi}}(x) \gamma^0 U_q^r(x) \} | \Omega : \text{in} \rangle + (\dots) \\ &= i \int d^4x \langle \text{out} : p | \{ \hat{\bar{\psi}}(x) i \overleftarrow{\partial}_0 \gamma^0 U_q^r(x) + \hat{\bar{\psi}}(x) i \overrightarrow{\partial}_0 \gamma^0 U_q^r(x) \} \\ &\quad \times | \Omega : \text{in} \rangle + (\dots) \end{aligned} \quad (11.144)$$

We convert the time derivative of $U_q^r(x)$ to a space derivative using the equation of motion

$$\begin{aligned} \int d^4x \hat{\bar{\psi}}(x) i \overrightarrow{\partial}_0 \gamma^0 U_q^r(x) &= \int d^4x \hat{\bar{\psi}}(x) \{ -i \gamma \cdot \vec{\nabla} + m \} U_q^r(x) \\ &= \int d^4x \hat{\bar{\psi}}(x) \{ i \gamma \cdot \vec{\nabla} + m \} U_q^r(x) \end{aligned} \quad (11.145)$$

where the last equation is obtained by partial integration. Then the first term in the last equality of Eq. (11.144) is combined to become

$$\langle \text{out} : p | q : \text{in} \rangle = i \int d^4x \langle p | \hat{\bar{\psi}}(x) | \Omega : \text{in} \rangle (i \overleftarrow{\partial} + m) U_q^r(x) \quad (11.146)$$

Similar calculations can be carried out and give

$$\begin{aligned}
 \sqrt{2\omega} \langle \text{out} : f | a_{\text{in}}^\dagger(q) | i : \text{in} \rangle &= (-i) \int d^4x \langle f | \hat{\bar{\psi}}(x) | i \rangle (-i \overleftarrow{\not{\partial}} - m)_x U_q^r(x) \\
 \sqrt{2E_p} \langle \text{out} : f | a_{\text{out}}(p) | i : \text{in} \rangle &= (-i) \int d^4y \bar{U}_p^s(y) (i \overrightarrow{\not{\partial}} - m)_y \langle f | \hat{\psi}(y) | i \rangle \\
 \sqrt{2\omega'} \langle \text{out} : f | b_{\text{in}}^\dagger(q') | i : \text{in} \rangle &= i \int d^4x' \bar{V}_{q'}^m(x') (i \overrightarrow{\not{\partial}} - m)_{x'} \langle f | \hat{\psi}(x') | i \rangle \\
 \sqrt{2E_{p'}} \langle \text{out} : f | b_{\text{out}}(p') | i : \text{in} \rangle &= i \int d^4y' \langle f | \hat{\bar{\psi}}(y') | i \rangle (-i \overleftarrow{\not{\partial}} - m)_{y'} V_{p'}^n(y')
 \end{aligned} \tag{11.147}$$

For the one-particle transition $q \rightarrow p$, we have

$$\begin{aligned}
 &\langle \text{out} : p | q : \text{in} \rangle \\
 &= (-i)^2 \int d^4x d^4y \bar{U}_p(y) (i \overrightarrow{\not{\partial}} - m)_y \langle \Omega | T[\psi(y) \bar{\psi}(x)] \Omega \rangle \\
 &\quad \times (-i \overleftarrow{\not{\partial}} - m)_x U_q(x) \\
 &= (-i)^2 \int d^4x d^4y \bar{U}_p(y) (i \overrightarrow{\not{\partial}} - m)_y \\
 &\quad \times \left(-\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(y)} \right) \left(\frac{1}{i} \frac{\delta}{\delta \eta(x)} \right) Z_0[\bar{\eta}, \eta] (-i \overleftarrow{\not{\partial}} - m)_x U_q(x)
 \end{aligned} \tag{11.148}$$

We now have to extend the reduction formula to many-particle systems. As the ordering of the operators is important, we begin with the transition amplitude. In the transition amplitude, particles are arranged in such a way that the subscripts increase from left to right and antiparticles are placed to the right of particles. Namely, denoting the momenta by

$$\underbrace{q_1 + \cdots + q_m}_{\text{particles}} + \underbrace{\bar{q}_1 + \cdots + \bar{q}_{m'}}_{\text{antiparticles}} \rightarrow \underbrace{p_1 + \cdots + p_n}_{\text{particles}} + \underbrace{\bar{p}_1 + \cdots + \bar{p}_{n'}}_{\text{antiparticles}}$$

the transition amplitude can be written as

$$\begin{aligned}
 &\langle \text{out} : \beta | \alpha : \text{in} \rangle = \langle \text{out} : p_1, \dots, p_n; \bar{p}_1, \dots, \bar{p}_{n'} | q_1, \dots, q_m; \\
 &\quad \bar{q}_1, \dots, \bar{q}_{m'} : \text{in} \rangle \\
 &= (i)^{m'+n'} (-i)^{m+n} \int \prod_{i=1}^m \prod_{i'=1}^{n'} d^4x_i d^4y_{i'} \prod_{j=1}^n \prod_{j'=1}^{m'} d^4x'_j d^4y_j \\
 &\quad \times \bar{U}_p(y_j) (i \overrightarrow{\not{\partial}} - m)_{y_j} \bar{V}_{q'}(x'_{j'}) (i \overrightarrow{\not{\partial}} - m)_{x'_{j'}} \\
 &\quad \times \langle \Omega | T[\psi(y_1) \cdots \psi(y_n) \bar{\psi}(y'_1) \cdots \bar{\psi}(y'_{n'}) \bar{\psi}(x_1) \cdots \bar{\psi}(x_m) \\
 &\quad \psi(x'_1) \cdots \psi(x'_{m'})] \Omega \rangle (-i \overleftarrow{\not{\partial}} - m)_{x_i} U_q(x_i) (-i \overleftarrow{\not{\partial}} - m)_{y_{i'}} V_{p'}(y'_{i'})
 \end{aligned} \tag{11.149}$$

Here, the ordering of the operators matches that in $\langle \text{out} : \beta | \alpha : \text{in} \rangle$ exactly. Rearrangement of the operators produces additional sign changes, of course. To express the above formula in terms of functional derivatives, the vacuum expectation value of the operator T-product can be replaced by the functional derivatives

$$\hat{\psi}(\gamma_i) \leftrightarrow \frac{1}{i} \frac{\delta}{\delta \bar{\eta}(\gamma_i)}, \quad \hat{\bar{\psi}}(x_i) \leftrightarrow -\frac{1}{i} \frac{\delta}{\delta \eta(x_i)} \quad (11.150)$$

provided we keep the ordering of the functional derivatives exactly as that of the operators in the T-product [see Eq. (11.124)]. In the following, we only give the case for $n = m, n' = m' = 0$:

$$\begin{aligned} \langle \text{out} : p_1 \cdots p_n | q_1 \cdots q_n : \text{in} \rangle &= \int \prod_{i=1}^n d^4 x_i d^4 \gamma_i \bar{U}_{p_i}(\gamma_i) (i \overrightarrow{\not{D}} - m)_{\gamma_i} \\ &\times \frac{\delta^{2n} Z_0[\bar{\eta}, \eta]}{\delta \eta(x_1) \cdots \eta(x_n) \delta \bar{\eta}(\gamma_1) \cdots \delta \bar{\eta}(\gamma_n)} (-\overleftarrow{\not{D}} - m)_{x_i} U_{q_i}(x_i) \end{aligned} \quad (11.151)$$

where the phases are exactly canceled first by that of the functional derivatives and second by moving $\delta/\delta \bar{\eta}(\gamma_i)$'s to the right of $\delta/\delta \eta(x)$'s.

11.6

QED

We are now ready to apply the method of the path integral to actual calculation of QED processes. We shall derive the scattering amplitude for reactions such as the Compton and $e\text{-}\mu$ scatterings and show that they produce the same results as calculated using the canonical quantization method.

11.6.1

Formalism

We now want to write down the scattering matrix describing the interactions of electrons and photons. In the previous section, we learned to express the scattering matrix in the form of the LSZ reduction formula. As it is expressed in terms of the functional derivatives of the generating functional of the interacting fields, the remaining task is to derive the generating functional of interacting fields involving the Dirac and electromagnetic fields.

In QED, which describes phenomena involving the electron (a Dirac field) interacting with the electromagnetic field, the whole Lagrangian $\mathcal{L}_{\text{QED}}$ can be obtained from the free Lagrangian of the electron and the electromagnetic field by replacing the derivative ∂_μ by the covariant derivative $D_\mu = \partial_\mu + iqA_\mu$, a procedure called “gauging”.

First we write the gauge-invariant full Lagrangian for the electron interacting with the electromagnetic field (photon).

$$\mathcal{L}_{\text{free}} \xrightarrow{\partial_\mu \rightarrow D_\mu} \mathcal{L}_{\text{QED}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i \not{D} - m)\psi \quad (11.152a)$$

$$D = \gamma^\mu (\partial_\mu + i q A_\mu), \quad (q = -e)$$

$$\mathcal{L}_{\text{QED}} = \mathcal{L}_0 + \mathcal{L}_{\text{int}} \quad (11.152b)$$

$$\mathcal{L}_0 = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i \not{\partial} - m)\psi$$

$$\mathcal{L}_{\text{int}} = -q \bar{\psi} \gamma^\mu \psi A_\mu$$

The action of the generating functional of the interacting electrons and photons should include, in addition to $\mathcal{L}$ of Eq. (11.152b), a gauge-fixing term

$$-\frac{1}{2\lambda} (\partial_\mu A^\mu)^2 \quad (11.153)$$

and source terms, and hence is not gauge invariant. We adopt the Feynman gauge and set $\lambda = 1$. Then we can write down the generating functional:

$$\begin{aligned} Z[\bar{\eta}, \eta, J] = N \int \mathcal{D}\bar{\psi} \mathcal{D}\psi \mathcal{D}A \exp \left[i \int d^4x \right. \\ \left. \times \left\{ \mathcal{L}_0 + \mathcal{L}_{\text{int}} - \frac{1}{2} (\partial_\mu A^\mu)^2 + J_\mu A^\mu + \bar{\eta} \psi + \bar{\psi} \eta \right\} \right] \end{aligned} \quad (11.154)$$

where N is a normalization constant to make $Z[0, 0, 0] = 1$. After the standard manipulation to integrate the bilinear terms [refer to Eqs. (11.66) and (11.120)], the generating functional $Z_0[\bar{\eta}, \eta, J^\mu]$ for the free fields, namely those with $\mathcal{L}_{\text{int}} = 0$, can be expressed as

$$\begin{aligned} Z_0[\bar{\eta}, \eta, J] = \exp \left[-i \int d^4x d^4y \right. \\ \left. \times \left\{ \frac{1}{2} J^\mu(x) D_{F\mu\nu}(x-y) J^\nu(y) + \bar{\eta}(x) S_F(x-y) \eta(y) \right\} \right] \end{aligned} \quad (11.155)$$

We learned in Sect. 11.2.5, that the generating functional of the interacting fields can be obtained by replacing the field in the interaction Lagrangian by its functional derivatives [see Eq. (11.53)]. In QED, we replace the fields in the interaction Lagrangian by

$$\bar{\psi} \rightarrow -\frac{1}{i} \frac{\delta}{\delta \eta}, \quad \psi \rightarrow \frac{1}{i} \frac{\delta}{\delta \bar{\eta}}, \quad A_\mu \rightarrow \frac{1}{i} \frac{\delta}{\delta J^\mu} \quad (11.156)$$

Therefore, the generating functional for QED is given by

$$\begin{aligned}
 Z[\bar{\eta}, \eta, J] &= N \exp \left[-i \int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta} \right) (q \gamma^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\mu} \right) \right] \\
 &\quad \times Z_0[\bar{\eta}, \eta, J] \\
 &= N \exp \left[\int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta} \right) (-i q \gamma^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\mu} \right) \right] \\
 &\quad \times Z_0[\bar{\eta}, \eta, J]
 \end{aligned} \tag{11.157}$$

We retain the minus sign in front of the coupling constant because the action is $iS = i \int d^4x \mathcal{L}_{\text{int}} = \int d^4x \{-iV(x)\}$ and the perturbation expansion is carried out in powers of minus the coupling constant, i.e. in $(-iq)^n$.

11.6.2

Perturbative Expansion

The exponential of the interaction part in Eq. (11.157) can be Taylor-expanded in the coupling strength $|q|$

$$\begin{aligned}
 Z[\bar{\eta}, \eta, J] &= N \sum_{n=0}^{\infty} \frac{1}{n!} \\
 &\quad \times \left\{ \int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta} \right) (-i q \gamma^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\mu} \right) \right\}^n Z_0[\bar{\eta}, \eta, J]
 \end{aligned} \tag{11.158}$$

Note, η and $\bar{\eta}$ are actually four-component spinors and the above expression is an abbreviation of

$$\left(-\frac{\delta}{\delta \eta} \right) \gamma^\mu \left(\frac{\delta}{\delta \bar{\eta}} \right) = \sum_{a,b=1}^4 \left(-\frac{\delta}{\delta \eta_a} \right) (\gamma^\mu)_{ab} \left(\frac{\delta}{\delta \bar{\eta}_b} \right) \tag{11.159}$$

In the following we calculate the first- and second-order terms explicitly and express these results graphically. The Feynman rules that establish the connection between the algebraic and graphical representations are very useful in understanding the so far very formal treatment of the functional derivatives and are of great help in performing actual calculations.

11.6.3

First-Order Interaction

The first-order interaction term is given by

$$Z_1 \equiv N \left\{ \int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta(x)} \right) (-iq\gamma^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(x)} \right) \right. \\ \left. \times \left(\frac{1}{i} \frac{\delta}{\delta J^\mu(x)} \right) \right\} Z_0[\bar{\eta}, \eta, J] \quad (11.160a)$$

$$= N \left\{ \int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta_a(x)} \right) (-iq\gamma_{ab}^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}_b(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\mu(x)} \right) \right\} \\ \times \exp \left[-i \int d^4z_2 d^4z_3 \left\{ \frac{1}{2} J^\rho(z_2) D_{F\rho\sigma}(z_2 - z_3) J^\sigma(z_3) \right. \right. \\ \left. \left. + \bar{\eta}_l(z_2) \{S_F(z_2 - z_3)\}_{lm} \eta_m(z_3) \right\} \right] \quad (11.160b)$$

$$= N \left\{ \int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta_a(x)} \right) (-iq\gamma_{ab}^\mu) \right\} \\ \times \left[\int d^4z_1 d^4z_3 D_{F\mu\sigma}(x - z_1) J^\sigma(z_1) \{S_F(x - z_3)\}_{bm} \eta_m(z_3) \right] \\ \times \exp \left[-i \int d^4w_2 d^4w_3 \left\{ \frac{1}{2} J^\alpha(w_2) D_{F\alpha\beta}(w_2 - w_3) J^\beta(w_3) \right. \right. \\ \left. \left. + \bar{\eta}_s(w_2) \{S_F(w_2 - w_3)\}_{st} \eta_t(w_3) \right\} \right] \\ = \left[+iN \int d^4x d^4z_1 D_{F\mu\sigma}(x - z_1) J^\sigma(z_1) \{(-iq\gamma_{ab}^\mu)\}_{ba} \{S_F(x - x)\}_{ba} \right. \\ \left. + N \int d^4x d^4z_1 d^4z_2 d^4z_3 D_{F\mu\sigma}(x - z_1) J^\sigma(z_1) \right. \\ \left. \times (-iq\gamma_{ab}^\mu) \{S_F(x - z_3)\}_{bm} \eta_m(z_3) \bar{\eta}_l(z_2) \{S_F(z_2 - x)\}_{la} \right] Z_0[\bar{\eta}, \eta, J] \\ = \left[iN \int d^4x d^4z_1 \text{Tr} [(-iq\gamma^\mu) S_F(x - x)] D_{F\mu\nu}(x - z_1) J^\nu(z_1) \right. \\ \left. + (-1)N \int d^4x d^4z_1 d^4z_2 d^4z_3 D_{F\mu\nu}(x - z_1) J^\nu(z_1) \right. \\ \left. \times \bar{\eta}(z_2) S_F(z_2 - x) (-iq\gamma^\mu) S_F(x - z_3) \eta(z_3) \right] Z_0[\bar{\eta}, \eta, J] \quad (11.160c)$$

The first term vanishes because

$$\text{Tr}[\gamma^\mu S_F(0)] = \int \frac{d^4p}{(2\pi)^4} \frac{\text{Tr}[\gamma^\mu (\not{p} - m)]}{p^2 - m^2 - i\varepsilon} = \int \frac{d^4p}{(2\pi)^4} \frac{4p^\mu}{p^2 - m^2 - i\varepsilon} = 0 \quad (11.161)$$

Note the extra (-1) phase in the second term of the last line, which was generated by moving $\bar{\eta}$ to the left of η . This does not happen for boson fields. The final expression for the first-order generating functional is

$$Z_1[\bar{\eta}, \eta, J] = -N \int d^4x d^4z_1 d^4z_2 d^4z_3 D_{F\mu\nu}(x - z_1) J^\nu(z_1) \times \bar{\eta}(z_2) S_F(z_2 - x) (-i q \gamma^\mu) S_F(x - z_3) \eta(z_3) Z_0[\bar{\eta}, \eta, J] \quad (11.162)$$

The expression Z_1/Z_0 is pictorially illustrated in Figure 11.1.

It represents several different processes. The reduction formula decides which one to pick up. In Figure 11.1a, an electron is created by a source denoted by a cross at z_3 , goes upward in time, interacts at x with the electromagnetic field and is destroyed by a source (or sink) at z_2 . Here, a photon is created at z_1 and is absorbed by the electron. But in (b), the photon is emitted by the electron at x and destroyed by the source (sink). The electromagnetic field is real and hence the photon is its own antiparticle. This is why the same source can act both as a source to create a photon and as a sink (negative source) to destroy it. In Figure 11.1c, the sense of time direction is reversed to denote a positron (electron's antiparticle) created at z_3 and destroyed at z_2 ; otherwise it is the same as (a).

The pictorial description of the expression is as follows:

1. Crosses represent sources or sinks ($\eta, \bar{\eta}, J^\nu$).
2. Dots represent interactions. " $-i q \gamma^\mu$ " denotes an interaction in which a Dirac spinor comes in and goes out, at the same time emitting or absorbing a photon.
3. Lines with an arrow represent the electron propagator, $S_F(x - z_3)$ going from z_3 to x and $S_F(z_2 - x)$ from x to z_2 . The arrow is directed downward if the line represents an antiparticle (positron). The spin state of the electron was suppressed but can be attached if necessary.
4. Wavy lines represent the photon propagator $D_{F\mu\nu}(x - z_1)$.

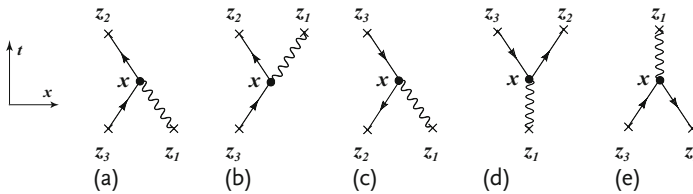


Figure 11.1 First-order generating functional of the interaction with three external sources. The same expression can represent various processes. (a,b) An electron is created at lower marked by a cross (z_3), interacts at the vertex (x) and is destroyed at the upper cross (z_2). A photon is created by a source at z_1

and absorbed by the electron at the vertex (a) or emitted by the electron and destroyed by the source (b). (c) The same as (a) except a positron is created at z_3 and destroyed at z_2 . (d) Electron positron pair creation by a photon. (e) Pair annihilation to a photon.

11.6.4

Mott Scattering

As a preparatory step to quantum field theory of the scattering matrix, we first describe a semiclassical treatment of electron scattered by a fixed Coulomb potential. This applies to the scattering of electrons by heavy nuclei. The nucleus is very heavy and its recoil can be neglected in low-energy scattering. Then the electric field created by the nucleus can be treated classically and only the electrons are treated using the quantized field. The electromagnetic field is not replaced by a functional derivative but is a classical field; it is represented by a Coulomb potential

$$A_\mu = (\phi, 0, 0, 0), \quad \phi = \frac{Zq'}{4\pi} \frac{1}{r} \quad (11.163)$$

where the electric charge of the nucleus is taken to be Zq' and that of the electron $q = -e$. The generating functional becomes

$$Z_1 \equiv N \left\{ \int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta(x)} \right) (-iq\gamma^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(x)} \right) A_\mu(x) \right\} \\ \times \exp \left[-i \int d^4z_2 d^4z_3 \bar{\eta}(z_2) S_F(z_2 - z_3) \eta(z_3) \right] \quad (11.164a)$$

$$= N \left[-i \int d^4x \text{Tr} [(-iq\gamma^\mu) S_F(x - x)] A_\mu(x) \right. \\ \left. + \int d^4x d^4z_2 d^4z_3 \bar{\eta}(z_2) S_F(z_2 - x) (-iq\gamma^\mu) S_F(x - z_3) \eta(z_3) A_\mu(x) \right] \\ \times \exp \left[-i \int d^4w_2 d^4w_3 \bar{\eta}(w_2) S_F(w_2 - w_3) \eta(w_3) \right] \quad (11.164b)$$

The first term does not contribute. The reduction formula Eq. (11.148) for an electron of momentum p coming in and one of p' going out becomes

$$S_{fi} = \langle \text{out} : p' | p : \text{in} \rangle \\ = (-i)^2 \int d^4w_2 d^4w_3 \bar{U}_{p'}(w_2) (i\vec{\not{\partial}} - m)_{w_2} \\ \times \left(\frac{1}{i^2} \right) \frac{\delta^2 Z_1[\bar{\eta}, \eta]}{\delta \eta(w_3) \delta \bar{\eta}(w_2)} \Big|_{\bar{\eta}=\eta=0} (-i\overleftarrow{\not{\partial}} - m)_{w_3} U_p(w_3) \\ = \int d^4x d^4w_2 d^4w_3 \bar{U}_{p'}(w_2) (i\vec{\not{\partial}} - m)_{w_2} \\ \times S_F(w_2 - x) (-iq\gamma^\mu) S_F(x - w_3) (-i\overleftarrow{\not{\partial}} - m)_{w_3} U_p(w_3) A_\mu(x) \quad (11.165)$$

From the properties of the Feynman propagators and the Dirac wave functions

$$(i\vec{\not{\partial}} - m)_{w_2} S_F(w_2 - x) = \delta^4(w_2 - x) \\ S_F(x - w_3) (-i\overleftarrow{\not{\partial}} - m)_{w_3} = \delta^4(x - w_3) \\ U_p(x) = u_r(p) e^{-ip \cdot x}, \quad \bar{U}_{p'}(x) = u_s(p') e^{ip' \cdot x} \quad (11.166)$$

the scattering amplitude becomes

$$\begin{aligned}
 S_{fi} &= \int d^4x e^{-i(p-p') \cdot x} \bar{u}_s(p') (-i q \gamma^\mu) u_r(p) A_\mu(x) \\
 &= -2\pi i \delta(E' - E) \int d^3x e^{i(p-p') \cdot x} \left(\frac{Z q q'}{4\pi r} \right) \bar{u}_s(p') \gamma^0 u_r(p) \\
 &= -2\pi i \delta(E' - E) \left(\frac{q q'}{|q q'|} \right) 4\pi Z \alpha \frac{\bar{u}_s(p') \gamma^0 u_r(p)}{q^2} \\
 q^2 &= (p - p')^2 = 4|p|^2 \sin^2 \frac{\theta}{2}
 \end{aligned} \tag{11.167}$$

where $\alpha = e^2/4\pi = 1/137$ is the fine structure constant and θ is the scattering angle. Equation (11.167) has reproduced the Mott scattering cross section formula Eq. (6.136).

11.6.5

Second-Order Interaction

Referring to Eq. (11.162), the second-order interaction is given by

$$\begin{aligned}
 Z_2 &\equiv \frac{N}{2!} \left\{ \int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta(x)} \right) (-i q \gamma^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\mu(x)} \right) \right\}^2 \\
 &\quad \times Z_0[\bar{\eta}, \eta, J] \\
 &= \frac{N}{2} \left\{ \int d^4\gamma \left(-\frac{1}{i} \frac{\delta}{\delta \eta(\gamma)} \right) (-i q \gamma^\rho) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(\gamma)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\rho(\gamma)} \right) \right\} \\
 &\quad \times \left[- \int d^4x d^4z_1 d^4z_2 d^4z_3 \bar{\eta}(z_2) S_F(z_2 - x) (-i q \gamma^\mu) \right. \\
 &\quad \times S_F(x - z_3) \eta(z_3) D_{F\mu\nu}(x - z_1) J^\nu(z_1) Z_0[\bar{\eta}, \eta, J] \left. \right] \\
 &\equiv \frac{N}{2} \left\{ \int d^4\gamma \left(-\frac{1}{i} \frac{\delta}{\delta \eta(\gamma)} \right) (-i q \gamma^\rho) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(\gamma)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\rho(\gamma)} \right) \right\} \\
 &\quad \times \left[\int d^4x d^4z_1 d^4z_2 d^4z_3 C[\bar{\eta}, \eta, J] Z_0[\bar{\eta}, \eta, J] \right]
 \end{aligned} \tag{11.168}$$

where we have separated the part of the fields already pulled out of $Z_0[\bar{\eta}, \eta, J]$ by the first interaction derivatives and abbreviated it as $C[\bar{\eta}, \eta, J]$. Then the effect of the second interaction derivatives on Z_0 is to pull an electron or photon line with external source out of Z_0 and connect it to the second vertex at γ . Namely, it adds an incoming or outgoing particle to the vertex at γ . On the other hand if the derivatives act on $C[\bar{\eta}, \eta, J]$, they connect either an electron line or photon line from the vertex x to the vertex γ . Namely it draws an internal line between the two vertices. Depending on which of the terms ($C[\bar{\eta}, \eta, J]$ or $Z_0[\bar{\eta}, \eta, J]$) are acted on by the three functional derivatives, there are six distinct processes, which will be explained one by one.

Disconnected Diagrams When all three derivatives act on $Z_0[\bar{\eta}, \eta, J]$, they pull down another trio of external lines and there are no internal lines to connect x and y . The contribution is

$$\begin{aligned}
 Z_{21} &= \frac{1}{2} \left[- \int d^4x d^4z_1 d^4z_2 d^4z_3 \bar{\eta}(z_2) S_F(z_2 - x) (-i q \gamma^\mu) \right. \\
 &\quad \times S_F(x - z_3) \eta(z_3) D_{F\mu\nu}(x - z_1) J^\nu(z_1) \left. \right] \\
 &\quad \times \left[- \int d^4y d^4w_1 d^4w_2 d^4w_3 \bar{\eta}(w_2) S_F(w_2 - y) (-i q \gamma^\rho) \right. \\
 &\quad \times S_F(y - w_3) \eta(w_3) D_{F\rho\sigma}(y - w_1) J^\sigma(w_1) \left. \right] Z_0[\bar{\eta}, \eta, J] \\
 &= \frac{1}{2} \left(\frac{Z_1}{Z_0} \right)^2 Z_0
 \end{aligned} \tag{11.169}$$

where Z_1 is defined by Eq. (11.162). This is two separate processes generated by two independent “trios of external sources”, each one identical to the first-order interaction, and is depicted in Fig. 11.2.

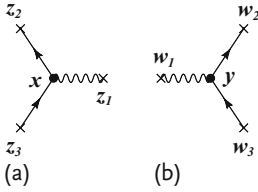


Figure 11.2 Among the second-order processes, these disconnected diagrams describe two independent processes, each one identical to the first-order interaction in Fig. 11.1.

This is the so-called disconnected diagram and describes two independent processes, each of which has already appeared in the lower order perturbation. We can eliminate the disconnected diagrams by using $W[\bar{\eta}, \eta, J] = -i \ln Z[\bar{\eta}, \eta, J]$ instead of $Z[\bar{\eta}, \eta, J]$, which will be explained later in Sect. 11.6.7.

Four External Electron Lines This is the case when both $\delta/\delta\bar{\eta}$ and $\delta/\delta\eta$ act on Z_0 and pull down two electron lines while $\delta/\delta J$ acts on C and connects the two vertices by a photon line. This contributes to electron–electron or Bhabha (e^-e^+)

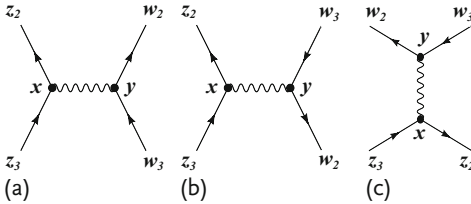


Figure 11.3 Four-point functions: Diagrams that contribute to (a) electron–electron or (b,c) Bhabha (e^-e^+) scattering.

scattering (see Fig. 11.3):

$$\begin{aligned}
 Z_{22} &= \frac{N}{2} \left\{ \int d^4 \gamma \left(-\frac{1}{i} \frac{\delta}{\delta \eta(\gamma)} \right) (-iq\gamma^\rho) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(\gamma)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\rho(\gamma)} \right) \right\} \\
 &\quad \times \left[- \int d^4 x d^4 z_1 d^4 z_2 d^4 z_3 \{ \bar{\eta}(z_2) S_F(z_2 - x) (-iq\gamma^\mu) \right. \\
 &\quad \times S_F(x - z_3) \eta(z_3) D_{F\mu\nu}(x - z_1) J^\nu(z_1) \} Z_0[\bar{\eta}, \eta, J] \Big] \\
 &= \frac{i}{2} N \int d^4 x d^4 \gamma d^4 z_2 d^4 z_3 d^4 w_2 d^4 w_3 \\
 &\quad \times \{ \bar{\eta}(z_2) S_F(z_2 - x) (-iq\gamma^\mu) S_F(x - z_3) \eta(z_3) D_{F\mu\rho}(x - \gamma) \\
 &\quad \times \bar{\eta}(w_2) S_F(w_2 - \gamma) (-iq\gamma^\rho) S_F(\gamma - w_3) \eta(w_3) \} Z_0 \quad (11.170)
 \end{aligned}$$

The S-matrix of the process will be discussed later.

Compton-Type Process Here $\delta Z_0/\delta J$ and $\delta Z_0/\delta \bar{\eta}$ pull down one more photon line and an electron line while the remaining $\delta C/\delta \eta$ connects the two vertices by an electron line. This gives the first term in the next equation and can represent Compton scattering, pair annihilation into two gammas or pair creation by two photons. Another one where the role of η and $\bar{\eta}$ is exchanged gives the second term:

$$\begin{aligned}
 Z_{23} &= \frac{N}{2} \left\{ \int d^4 \gamma \left(-\frac{1}{i} \frac{\delta}{\delta \eta(\gamma)} \right) (-iq\gamma^\rho) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(\gamma)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\rho(\gamma)} \right) \right\} \\
 &\quad \times \left[- \int d^4 x d^4 z_1 d^4 z_2 d^4 z_3 \bar{\eta}(z_2) S_F(z_2 - x) (-iq\gamma^\mu) \right. \\
 &\quad \times S_F(x - z_3) \eta(z_3) D_{F\mu\nu}(x - z_1) J^\nu(z_1) Z_0[\bar{\eta}, \eta, J] \Big] \\
 &= \frac{i}{2} N \int d^4 x d^4 \gamma d^4 z_1 d^4 z_3 d^4 w_1 d^4 w_3 \{ J^\sigma(w_1) D_{F\sigma\rho}(w_1 - \gamma) \\
 &\quad \times D_{F\mu\nu}(x - z_1) J^\nu(z_1) \} \left\{ \bar{\eta}(z_3) S_F(z_3 - x) (-iq\gamma^\mu) \right. \\
 &\quad \times S_F(x - \gamma) (-iq\gamma^\rho) S_F(\gamma - w_3) \eta(w_3) \\
 &\quad + \bar{\eta}(w_3) S_F(w_3 - \gamma) \\
 &\quad \times (-iq\gamma^\rho) S_F(\gamma - x) (-iq\gamma^\mu) S_F(x - z_3) \eta(z_3) \Big\} Z_0 \quad (11.171)
 \end{aligned}$$

The second term is identical to the first and compensates the factor two. This is seen if we exchange the variables $x \leftrightarrow \gamma$, $z_3 \leftrightarrow w_3$, $z_1 \leftrightarrow w_1$ and $\mu \leftrightarrow \rho$. Note, the diagrams (a-d) in Fig. 11.4 are described using the notation of the first term, but every diagram is included in both terms as every variable is integrated and can be renamed. The S-matrix of the process will be treated later.

Vacuum Energy When all three derivatives act on $C[\bar{\eta}, \eta, J]$, they connect all the external lines with each other and no external lines remain (Fig. 11.5a). This is a

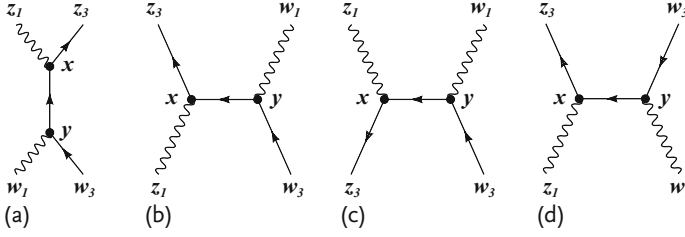


Figure 11.4 Four-point functions: Diagrams that contribute to Compton-like processes. (a,b) Compton scattering, (c) pair annihilation and (d) pair creation of photons.

vacuum-to-vacuum transition. The expression has the form

$$Z_{24} = \frac{i}{2} N \int d^4x d^4y \text{Tr} [(-iq\gamma^\rho) S_F(y-x)(-iq\gamma^\mu) \times S_F(x-y) D_{F\rho\mu}(x-y)] Z_0[\bar{\eta}, \eta, J] \quad (11.172)$$

It is absorbed by the normalization constant N to give $Z_2[0, 0, 0] = 1$.

Photon Self-Energy When only the photon derivative acts on Z_0 , and the others on C , the two electron lines already pulled down into C are connected to each other to make a loop. This is called a photon self-energy diagram (Fig. 11.5b). The expression becomes

$$Z_{25} = -\frac{N}{2} \int d^4x d^4y d^4z d^4w J^\sigma(w) D_{F\sigma\rho}(w-y) \times \text{Tr} [(-iq\gamma^\rho) S_F(y-x)(-iq\gamma^\mu) S_F(x-y)] \times D_{F\mu\nu}(x-z) J^\nu(z) Z_0 \quad (11.173)$$

This is a correction to the photon propagator. Note the extra minus sign for the loop. Whenever there are fermion loops, an extra sign (-1) is added, which comes from reordering the fermion fields. This appears in the higher order correction in the vacuum expectation value of the two-point time-ordered correlation function. The whole propagator consists of functional derivative of Z_0 and Z_{25} and it is given

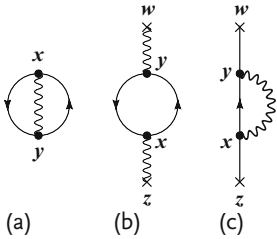


Figure 11.5 (a) Vacuum-to-vacuum transition amplitude, (b) photon self-energy and (c) electron self-energy.

by

$$\begin{aligned}
\langle \Omega | T[\hat{A}_\mu(x) \hat{A}_\nu(y)] | \Omega \rangle &= \left(\frac{1}{i} \right)^2 \frac{\delta^2}{\delta J^\mu(x) \delta J^\nu(y)} Z[\bar{\eta}, \eta, J] \Big|_{\bar{\eta}=\eta=J=0} \\
&= \left(\frac{1}{i} \right)^2 \frac{\delta^2}{\delta J^\mu(x) \delta J^\nu(y)} \{ Z_0 + Z_1 + (\cdots + Z_{25} + \cdots) + \cdots \} \Big|_{\bar{\eta}=\eta=J=0} \\
&= i D_{F\mu\nu}(x-y) + \int d^4w d^4z D_{F\mu\rho}(x-w) \\
&\quad \times \text{Tr} \left[(-iq\gamma^\rho) S_F(w-z) (-iq\gamma^\sigma) S_F(z-w) \right] D_{F\sigma\nu}(z-y) \quad (11.174)
\end{aligned}$$

The factor 1/2 was canceled by taking second derivatives of the quadratic $J^\sigma J^\nu$ term. Using the momentum representation of the propagators Eqs. (11.65) and (11.119) and integrating over w and z , the second term becomes

$$\begin{aligned}
&\int \frac{d^4k}{(2\pi)^4} e^{-ik \cdot (x-y)} \frac{-g_{\mu\rho}}{k^2 + i\varepsilon} \int \frac{d^4p}{(2\pi)^4} \\
&\quad \times \text{Tr} \left[(-iq\gamma^\rho) \frac{1}{\not{p} - m + i\varepsilon} (-iq\gamma^\sigma) \frac{1}{\not{p} - \not{k} - m + i\varepsilon} \right] \frac{-g_{\sigma\nu}}{k^2 + i\varepsilon} \quad (11.175)
\end{aligned}$$

Then the momentum representation of the photon propagator including the second-order correction is expressed as

$$\begin{aligned}
i \tilde{D}_{F\mu\nu}(k) &= \frac{-ig_{\mu\nu}}{k^2 + i\varepsilon} + \frac{-ig_{\mu\rho}}{k^2 + i\varepsilon} (-i\Pi^{\rho\sigma}) \frac{-ig_{\sigma\nu}}{k^2 + i\varepsilon} \\
-i\Pi^{\rho\sigma} &= - \int \frac{d^4p}{(2\pi)^4} \\
&\quad \times \text{Tr} \left[(-iq\gamma^\rho) \frac{1}{\not{p} - m + i\varepsilon} (-iq\gamma^\sigma) \frac{1}{\not{p} - \not{k} - m + i\varepsilon} \right] \quad (11.176)
\end{aligned}$$

Thus we have reproduced the radiative correction to the photon propagator Eq. (8.49b).

Electron Self-Energy When only one of the electron derivatives acts on Z_0 and the others on C , this pulls down an electron line from the source and connects it to the second vertex through both a photon and another electron line, i.e. this is the self-energy correction to the electron vertex (See Fig. 11.5c).

Problem 11.4

Derive the electron self-energy corresponding to Fig. 11.5c and show that its vacuum expectation value becomes a higher order correction term to $iS_F(x-y)$

$$\begin{aligned}
& i S_F(x - y)_{2\text{nd order}} \\
&= \int d^4 w d^4 z S_F(x - w) (-i q \gamma^\mu) S_F(w - z) (-i q \gamma^\nu) S_F(z - y) \\
&\quad \times D_{F\mu\nu}(w - z) \\
&= \int \frac{d^4 p}{(2\pi)^4} e^{-i p \cdot (x - y)} \frac{i}{\not{p} - m + i\epsilon} i \Sigma_e(p) \frac{i}{\not{p} - m + i\epsilon} \\
& i \Sigma_e(p) = \int \frac{d^4 k}{(2\pi)^4} \frac{-i g_{\mu\nu}}{k^2} (-i q \gamma^\mu) \frac{i}{\not{p} - \not{k} - m + i\epsilon} (-i q \gamma^\nu) \quad (11.177)
\end{aligned}$$

which reproduces Eq. (8.66).

11.6.6

Scattering Matrix

Electron–Muon Scattering To treat the interaction of the muon and the electron, we only need to add the muon Lagrangian to the QED Lagrangian. The former is identical to that of the electron except the particle species must be distinguished. The interaction Lagrangian becomes

$$\mathcal{L}_{\text{int}} = \mathcal{L}_{\text{int},e} + \mathcal{L}_{\text{int},\mu} = -q \bar{\psi}_e \gamma^\mu \psi_e - q \bar{\psi}_\mu \gamma^\mu \psi_\mu \quad (11.178)$$

where the electric charge of the electron and muon is $q = -e$. The second-order generating functional $Z_2[\bar{\eta}, \eta, J]$ contains

$$\begin{aligned}
& \frac{1}{2!} \left[i \int d^4 x \mathcal{L}_{\text{int}} \right]^2 = \frac{1}{2} \left[i \int d^4 x (\mathcal{L}_{\text{int},e} + \mathcal{L}_{\text{int},\mu}) \right]^2 \\
&= \frac{1}{2} \left(i \int d^4 x \mathcal{L}_{\text{int},e} \right)^2 + \frac{1}{2} \left(i \int d^4 x \mathcal{L}_{\text{int},\mu} \right)^2 + i \int d^4 x \mathcal{L}_{\text{int},e} i \int d^4 y \mathcal{L}_{\text{int},\mu} \quad (11.179)
\end{aligned}$$

Therefore, the generating functional of e – μ scattering is readily obtained from Eq. (11.170):

$$\begin{aligned}
Z_{e\mu} &= N \left\{ \int d^4 y \left(-\frac{1}{i} \frac{\delta}{\delta \eta(y)} \right) (-i q \gamma^\rho) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(y)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\rho(y)} \right) \right\} \\
&\quad \times \left[- \int d^4 x d^4 z_1 d^4 z_2 d^4 z_3 \{ \bar{\eta}(z_2) S_F(z_2 - x) (-i q \gamma^\mu) \right. \\
&\quad \times S_F(x - z_3) \eta(z_3) D_{F\mu\nu}(x - z_1) J^\nu(z_1) \} Z_0[\bar{\eta}, \eta, J] \Big] \\
&= i \int d^4 x d^4 y d^4 z_2 d^4 z_3 d^4 w_2 d^4 w_3 \\
&\quad \times \{ \bar{\eta}(z_2) S_F(z_2 - x) (-i q \gamma^\mu) S_F(x - z_3) \eta(z_3) \}_e D_{F\mu\rho}(x - y) \\
&\quad \times \{ \bar{\eta}(w_2) S_F(w_2 - y) (-i q \gamma^\rho) S_F(y - w_3) \eta(w_3) \}_\mu Z_0 \quad (11.180)
\end{aligned}$$

where we have distinguished the electron and the muon. To calculate the transition amplitude of the electron–muon scattering, we assign the following kinematic variables:

$$e(p_1) + \mu(p_2) \rightarrow e(p_3) + \mu(p_4) \quad (11.181)$$

Application of the LSZ reduction formula Eq. (11.151) to e – μ scattering gives

$$\begin{aligned} \langle \text{out} : p_3 p_4 | p_1 p_2 : \text{in} \rangle &= (-i)^4 \int d^4 x_1 d^4 x_2 d^4 x_3 d^4 x_4 \\ &\times \overline{U}_{p_3}(x_3)(i\vec{\partial} - m_e)_{x_3} \overline{U}_{p_4}(x_4)(i\vec{\partial} - m_\mu)_{x_4} \\ &\times \left(\frac{1}{i}\right)^4 \frac{\delta^4 Z_{e\mu}[\overline{\eta}, \eta, J]}{\delta \eta^\mu(x_2) \delta \eta^e(x_1) \delta \overline{\eta}^e(x_3) \delta \overline{\eta}^\mu(x_4)} \Big|_{\overline{\eta}=\eta=J=0} \\ &\times (-i\overleftarrow{\partial} - m_e)_{x_1} U_{p_1}(x_1) (-i\overleftarrow{\partial} - m_\mu)_{x_2} U_{p_2}(x_2) \end{aligned} \quad (11.182)$$

The effect of the 4-derivatives in the reduction formula on the generating functionals $Z_{e\mu}$ in Eq. (11.180) is given by

$$\begin{aligned} &\frac{\delta^4 Z_{e\mu}[\overline{\eta}, \eta, J]}{\delta \eta^\mu(x_2) \delta \eta^e(x_1) \delta \overline{\eta}^e(x_3) \delta \overline{\eta}^\mu(x_4)} \Big|_{\overline{\eta}=\eta=J=0} \\ &= i \int d^4 x d^4 y \left[\{ S_F(x_3 - x)(-i q \gamma^\mu) S_F(x - x_1) \}_e \right. \\ &\quad \left. \times D_{F\mu\rho}(x - y) \{ S_F(x_4 - y)(-i q \gamma^\rho) S_F(y - x_2) \}_\mu \right] \end{aligned} \quad (11.183)$$

Functional derivatives acting on Z_0 do not contribute since they contain the source functions. As the effect of the Dirac operators in the reduction formula is only to convert the outermost propagators to the corresponding wave functions, the final expression for the scattering amplitude becomes

$$\begin{aligned} \langle \text{out} : p_3, p_4 | p_1, p_2 : \text{in} \rangle &= i \int d^4 x d^4 y \{ \overline{U}_{p_3}(x)(-i q \gamma^\mu) U_{p_1}(x) \}_e \\ &\times D_{F\mu\rho}(x - y) \{ \overline{U}_{p_4}(y)(-i q \gamma^\rho) U_{p_2}(y) \}_\mu \end{aligned} \quad (11.184)$$

The formula is depicted in Fig. 11.6a.

When we insert the expressions

$$U_p(x) = u_r(p) e^{-ip \cdot x}, \quad \overline{U}_{p'} = \overline{u}_s(p')(x) e^{ip' \cdot x} \quad (11.185a)$$

$$D_{F\mu\nu}(x - y) = \int \frac{d^4 k}{(2\pi)^4} e^{-ik \cdot (x-y)} \frac{-g_{\mu\nu}}{k^2 + i\epsilon} \quad (11.185b)$$

(where we have recovered the subscript for the spin degree of freedom) in this formula, the integration of x and y produces two δ -functions, and after k integration

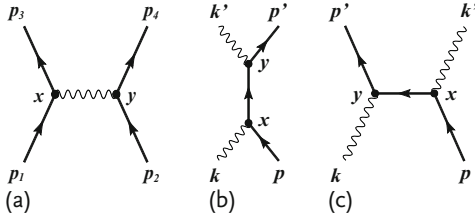


Figure 11.6 (a) Electron–muon scattering and (b,c) Compton scattering.

the scattering amplitude becomes

$$\begin{aligned}
 \langle \text{out} : p_3, p_4 | p_1, p_2 : \text{in} \rangle &= (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \\
 &\times [\bar{u}_s(p_3)(-iq\gamma^\mu)u_r(p_1)]_e \frac{-ig_{\mu\nu}}{(p_1 - p_3)^2 + i\varepsilon} [\bar{u}_n(p_4)(-iq\gamma^\nu)u_m(p_2)]_\mu
 \end{aligned}
 \quad (11.186)$$

This reproduces the scattering amplitude Eq. (7.11) obtained using the standard canonical quantization method.

Problem 11.5

Derive the Compton scattering amplitude and confirm it agrees with Eq. (7.51).

Hint: As the photon source J^μ has no directionality, it could serve either as a source or a sink. There are two ways to make the photon derivatives, producing two distinct scattering processes, as depicted in Fig. 11.6b,c.

11.6.7

Connected Diagrams

In the expansion of the generating functional, we saw that it contained the disconnected diagrams that describe two (or more) independent processes. We will show that the generating functional W produces only connected diagrams. It is related to Z by

$$Z[J] = e^{iW[J]} \quad \text{or} \quad W[J] = -i \ln Z[J] \quad (11.187)$$

We consider the case of the second-order interaction. The generating functional of QED was given by Eq. (11.157), which we reproduce here:

$$\begin{aligned}
 Z[\bar{\eta}, \eta, J] &= N \exp \left[\int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta} \right) (-iq\gamma^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\mu} \right) \right] \\
 &\times Z_0[\bar{\eta}, \eta, J]
 \end{aligned}
 \quad (11.188)$$

For ease of calculation we abbreviate the exponent by iL_I and express $Z[J]$ as follows:

$$Z[J] \equiv Z[\bar{\eta}, \eta, J] = N Z_0 Z_0^{-1} e^{iL_I} Z_0[J] \quad (11.189)$$

Taking the logarithm of $Z[J]$, we can Taylor expand to the second order of the interaction Lagrangian iL_I , assuming it is small:

$$\begin{aligned}
 \ln Z[J] &= \ln N Z_0 + \ln (Z_0^{-1} e^{iL_I} Z_0) \\
 &= \ln N Z_0 + \ln \{1 + Z_0^{-1} (e^{iL_I} - 1) Z_0\} \\
 &= \ln N Z_0 + \ln \left\{1 + Z_0^{-1} \left(iL_I + \frac{1}{2!}(iL_I)^2 + \dots\right) Z_0\right\} \\
 &= \ln N Z_0 + Z_0^{-1} \left\{iL_I + \frac{1}{2!}(iL_I)^2\right\} Z_0 - \frac{1}{2} \{Z_0^{-1} (iL_I) Z_0\}^2 + \dots \\
 &= \ln N + S_0 + S_1 + \left(S_2 - \frac{1}{2} S_1^2\right) + \dots
 \end{aligned} \tag{11.190}$$

Recovering the original notation

$$\begin{aligned}
 S_0 &= \ln Z_0 \\
 &= -\frac{i}{2} \int d^4x d^4y [J^\mu(x) D_{F\mu\nu}(x-y) J^\nu(y) + \bar{\eta}(x) S_F(x-y) \eta(y)] \\
 S_1 &= Z_0^{-1} (iL_I) Z_0 \\
 &= \frac{i}{Z_0} \int d^4x \mathcal{L}_{\text{int}} \left(-\frac{1}{i} \frac{\delta}{\delta \eta}, \frac{1}{i} \frac{\delta}{\delta \bar{\eta}}, \frac{1}{i} \frac{\delta}{\delta J^\mu}\right) Z_0[\bar{\eta}, \eta, J] = \frac{Z_1}{Z_0} \\
 S_2 &= Z_0^{-1} \frac{1}{2!} (iL_I)^2 Z_0 \\
 &= \frac{1}{Z_0} \frac{1}{2!} \left\{i \int d^4x \mathcal{L}_{\text{int}} \left(-\frac{1}{i} \frac{\delta}{\delta \eta}, \frac{1}{i} \frac{\delta}{\delta \bar{\eta}}, \frac{1}{i} \frac{\delta}{\delta J^\mu}\right)\right\}^2 Z_0[\bar{\eta}, \eta, J] \\
 &= \frac{Z_2}{Z_0} \\
 \frac{1}{2} S_1^2 &= \frac{1}{2} \left(\frac{Z_1}{Z_0}\right)^2
 \end{aligned} \tag{11.191}$$

where Z_1 and Z_2 are the first- and second-order terms of $Z[J]$ in the power expansion of the interaction Lagrangian defined in Eqs. (11.162) and (11.168). The disconnected diagram that appeared in the second-order interaction was given by $Z_{21} = \frac{1}{2} \left(\frac{Z_1}{Z_0}\right)^2 Z_0$ [see Eq. (11.169)], which is exactly canceled in $\ln Z[J]$, as is seen in the parentheses in the last equation of Eq. (11.190). Higher order disconnected diagrams can also be shown to be canceled by the corresponding products of lower order contributions.

11.7

Faddeev–Popov’s Ansatz*

This section is provided to explain why a ghost field is necessary in the covariant formalism of non-Abelian gauge theory. The discussion is formal and mathematical. The reader is advised to read Chap. 18 first and then come back here.

11.7.1

A Simple Example*¹⁾

The concept of Faddeev and Popov’s gauge-fixing method is rather hard to grasp. However, the method constitutes an important concept for handling symmetry in the path integral and becomes an indispensable tool when we extend our discussion to non-Abelian gauge theory. Therefore we shall make a small detour, spending a few extra pages rederiving the gauge-fixing condition that could be phrased in one sentence for Abelian gauge theory. Perhaps it helps if we do a simple exercise that follows their strategy closely to isolate the redundant degrees of freedom. Consider an integral

$$W = \iint dx dy G(x, y) = \iint d\mathbf{r} G(r), \quad r = \sqrt{x^2 + y^2} \quad (11.192)$$

Here, we assume $G(x, y)$ is a function of r and has rotational symmetry. The nature of the symmetry is apparent in polar coordinates:

$$W = \int d\theta \int dr r G(r) \quad (11.193)$$

The integral is invariant under coordinate rotations

$$\mathbf{r} = (r, \theta) \rightarrow \mathbf{r}_\phi = (r, \theta + \phi) \quad (11.194)$$

The essential part of the integrand that does not change by rotation is isolated as a function of r . The symmetry manifests itself in that it does not depend on θ and that θ integration merely adds an integration volume over the symmetry space, which is the cause of an infinity in the path integral. Our aim is to embed this symmetry condition in the original coordinates. To do that we insert

$$1 = \int d\phi \delta(\phi - \theta) \quad (11.195)$$

Then Eq. (11.193) becomes

$$W = \int d\phi \int dr d\theta \delta(\phi - \theta) r G(r) \quad (11.196)$$

By this insertion, the second integration has been reduced from a two-dimensional to a one-dimensional one by the delta function, which constrains the two-dimensional integration to be carried out at a fixed angle $\theta = \phi$.

More generally, a condition $f(\mathbf{r}) = 0$ that fixes the angle at some value of ϕ may be used. ϕ can vary as a function of r as long as $f(\mathbf{r}) = 0$ gives a unique solution of ϕ for a given value of r . As $\int d\phi \delta[f(\mathbf{r})] \neq 1$, we define a quantity $\Delta_f(\mathbf{r})$ by

$$\Delta_f(\mathbf{r}) \int d\phi \delta[f(\mathbf{r}_\phi)] = 1 \quad (11.197)$$

1) Here we closely follow the arguments of Cheng and Li [97].

Then

$$[\Delta_f(\mathbf{r})]^{-1} \equiv \int d\phi \delta[f(\mathbf{r}_\phi)] = \int d\phi \left(\frac{df}{d\phi} \right)^{-1} \delta[\phi - \theta(\mathbf{r})] \quad (11.198)$$

$$\therefore \Delta_f(\mathbf{r}) = \left. \frac{\partial f(\mathbf{r})}{\partial \theta} \right|_{f=0}$$

$\Delta_f(\mathbf{r})$ itself is invariant under the rotation because

$$[\Delta_f(\mathbf{r}_{\phi'})^{-1}] = \int d\phi \delta[f(\mathbf{r}_{\phi+\phi'})] = \int d\phi'' \delta[f(\mathbf{r}_{\phi''})] = [\Delta_f(\mathbf{r})]^{-1} \quad (11.199)$$

Insertion of Eq. (11.197) in Eq. (11.192) gives

$$W = \int d\phi W_\phi = \int d\phi \left[\int d\mathbf{r} G(\mathbf{r}) \Delta_f(\mathbf{r}) \delta[f(\mathbf{r})] \right] \quad (11.200)$$

W_ϕ is rotationally invariant because

$$\begin{aligned} W_{\phi'} &= \int d\mathbf{r} G(\mathbf{r}) \Delta_f(\mathbf{r}) \delta[f(\mathbf{r}_{\phi'})] = \int d\mathbf{r}' G(\mathbf{r}') \Delta_f(\mathbf{r}') \delta[f(\mathbf{r}')] \\ &= W_\phi \end{aligned} \quad (11.201)$$

where $\mathbf{r}' = (r, \phi')$ and we have used the fact that both $G(\mathbf{r})$ and $\Delta_f(\mathbf{r})$ are invariant under rotation. We have succeeded in isolating the symmetry-independent volume factor from the symmetry-invariant integral.

11.7.2

Gauge Fixing Revisited*

In gauge theories, we consider a local gauge transformation²⁾

$$A_\mu \rightarrow A_\mu^\omega \quad (11.202a)$$

$$A_\mu^\omega = U(\omega) A_\mu U^{-1}(\omega) - \frac{i}{g} U(\omega) \partial_\mu U^{-1}(\omega) \quad (11.202b)$$

$$U(\omega) = \exp \left[-i \sum_{a=1}^N \omega_a(x) T_a \right] \quad (11.202c)$$

In the following discussion, we consider only an infinitesimal transformation, but the conclusion is general. Then the transformation is written as

$$A_a^\mu{}^\omega = A_a^\mu + f_{abc} \omega_b A_c^\mu + \frac{1}{g} \partial^\mu \omega_a \quad (11.203)$$

2) Gauge transformation properties are discussed in detail in Chap. 18. Here, it suffices to accept that general infinitesimal gauge transformations are given by Eq. (11.203).

where a, b, c ($a, b, c = 1 \cdots N_c$) are color indices for non-Abelian gauge transformation and f_{abc} ’s are structure constants of the unitary group representation. For Abelian groups, $f_{abc} = 0$. For a given set of fields $A_a^\mu(x)$ in the gauge-invariant Lagrangian, the action is constant on any path of $A_a^{\mu\omega}(x)$ that can be obtained by the gauge transformation $U(\omega)$. This will naturally give a infinite volume because the path integral is over the whole range of A , including A^ω ’s that are gauge transformed from A . We want to separate this part of the path integral. The procedure will be exactly the same as in the simple example we have just discussed. First we have to fix the gauge at some value. The gauge condition is expressed generically as

$$g_a(A_\mu) = 0, \quad \text{i.e.} \quad g_a(A_\mu^\omega)|_{\omega=0} = 0 \quad (11.204)$$

Some commonly used gauge-fixing conditions may be represented in the form

$$\text{Lorentz gauge:} \quad \partial_\mu A_a^\mu = 0 \quad (11.205a)$$

$$\text{Coulomb gauge:} \quad \nabla \cdot \mathbf{A}_a = 0 \quad (11.205b)$$

$$\text{Axial gauge:} \quad A_a^3 = 0 \quad (11.205c)$$

$$\text{Temporal gauge:} \quad A_a^0 = 0 \quad (11.205d)$$

The generating functional of the gauge fields is expressed as

$$Z = \int \mathcal{D}A e^{iS[A]} \quad (11.206)$$

Then we insert

$$\Delta[A^\omega] \int \mathcal{D}\omega \delta[g(A^\omega)] = 1 \quad (11.207)$$

which is analogous to Eq. (11.197) and defines $\Delta[A^\omega]$. The generating functional is transformed to

$$\begin{aligned} Z &= \int \mathcal{D}A e^{iS[A]} \int \mathcal{D}\omega \Delta[A^\omega] \delta[g(A^\omega)] \\ \mathcal{D}A \mathcal{D}\omega &= \prod_{v=0}^3 \prod_{a=1}^{N_c} \prod_{x^\rho} dA_a^v(x^\rho) \prod_{b=1}^{N_c} \prod_{y^\sigma} d\omega_b(y^\sigma) \end{aligned} \quad (11.208)$$

Notice, however, that Eq. (11.207) is a functional integral and “the delta functional” should be considered as an infinite product of δ -functions, one for each space-time point x :

$$\delta[g(A^\omega)] = \prod_{a=1}^{N_c} \prod_{x^\mu} \delta(g_a(A^\omega(x^\mu))) \quad (11.209)$$

Namely Eq. (11.207) should read

$$\begin{aligned}
 1 &= \left(\prod_a^{N_c} \int d\omega_a \delta[g_a(A^\omega)] \right) \det \left(\frac{\partial g_a}{\partial \omega_b} \right) \\
 &= \lim_{N \rightarrow \infty} \prod_n^N \int d\omega_n \delta[g_n(A^\omega)] \frac{\partial(g_1, \dots, g_N)}{\partial(\omega_1, \dots, \omega_N)}
 \end{aligned} \tag{11.210}$$

where n includes both color and the discrete space-time point. Then

$$\begin{aligned}
 \Delta[A]^{-1} &= \prod_{a=1}^{N_c} \prod_{x^\rho} \int d\omega_a(x^\rho) \delta(g_a(A^\omega(x^\rho))) \\
 &= \underbrace{\prod_{a=1}^{N_c} \prod_x \int d g_a(x) \delta(g_a(A^\omega(x)))}_{=1} \frac{\partial[\omega_1(x), \dots, \omega_{N_c}(x)]}{\partial[g_1(x), \dots, g_{N_c}(x)]} \\
 &\equiv \det \left[\frac{\partial \omega}{\partial g} \right]_{\substack{y=x \\ g=0}}
 \end{aligned} \tag{11.211}$$

The last step defines the functional determinant of the continuous matrix $\partial \omega_a(x)/\partial g_b(y)$, with rows labeled by (a, x) and columns by (b, y) . The gauge condition $g_a(A(x)) = 0$ can be rewritten as $g_a(A^\omega(x))|_{\omega=0} = 0$. $\Delta[A]$ is then expressed as

$$\Delta[A] = \det \left[\frac{\delta g(A^\omega)}{\delta \omega} \right]_{\omega=0} \quad \text{for } A \text{ satisfying } g(A) = 0 \tag{11.212}$$

The functional determinant in Eq. (11.212) can be calculated using Eq. (11.203). Clearly, $\Delta(A^\omega)$ is gauge invariant

$$\Delta[A] = \Delta[A^\omega] \tag{11.213}$$

because, in Eq. (11.207), if we perform another gauge transformation $g_a(A^\omega) \rightarrow g_a(A^{\omega+\omega'})$, the group nature of the measure means $\mathcal{D}\omega = \mathcal{D}(\omega + \omega')$ and hence

$$\begin{aligned}
 \Delta[A^{\omega+\omega'}] &= \int \mathcal{D}\omega \delta[g(A^{\omega+\omega'})] = \int \mathcal{D}(\omega + \omega') \delta[g(A^{\omega+\omega'})] \\
 &= \int \mathcal{D}\omega'' \delta[g(A^{\omega''})] = \Delta[A^\omega]
 \end{aligned} \tag{11.214}$$

At this stage, we can look at the generating functional Z in Eq. (11.208) and change the order of $\mathcal{D}A$ and $\mathcal{D}\omega$. Since $e^{iS[A]}$, $\Delta[A^\omega]$ are all gauge invariant, they can be replaced by $e^{iS[A^\omega]}$, $\Delta[A^\omega]$. Now we change the integration variable to $\mathcal{D}A^\omega$, but the new integration variable can be renamed back to A , leading to

$$\begin{aligned}
 Z &= \int \mathcal{D}A e^{iS[A]} \int \mathcal{D}\omega \Delta[A^\omega] \delta[g(A^\omega)] \\
 &= \int \mathcal{D}\omega \left[\int \mathcal{D}A e^{iS[A]} \Delta[A] \delta[g(A)] \right]
 \end{aligned} \tag{11.215}$$

where we have been able to take out the factor $\mathcal{D}\omega$ since the integral in large square brackets is independent of ω . This is the Faddeev–Popov ansatz [136]. The integration over ω simply gives an overall multiplicative constant to Z and can be absorbed in the normalization constant. Thus we have successfully isolated the integration over redundant degrees of freedom originating from the gauge symmetry. In order to obtain a more tractable expression, however, we go one step further to manipulate Eq. (11.207) before we put it back into the generating functional.

The identity Eq. (11.207) can be generalized to

$$\Delta(A) \int \mathcal{D}\omega \prod_{a=1}^N \delta[g_a(A^\omega) - h_a(x)] = 1 \quad (11.216)$$

where $h_a(x)$ is a set of arbitrary functions. This merely redefines g_a to make the gauge condition $g_a(A^\omega(x)) = h_a(x)$. As $h_a(x)$ is independent of A^μ , $\Delta[A]$ is unaffected. Now multiply both sides by an arbitrary functional $H[h] = H[h_1(x), \dots, h_N(x)]$ and integrate over all $h_a(x)$. The result is a more general identity:

$$\Delta[A] \frac{\int \mathcal{D}\omega \mathcal{D}h \prod_{a=1}^N \delta[g_a(A^\omega) - h_a(x)] H[h]}{\int \mathcal{D}h H[h]} = \Delta[A] \frac{\int \mathcal{D}\omega H[g(A^\omega)]}{\int \mathcal{D}h H[h]} = 1 \quad (11.217)$$

Now insert this identity instead of Eq. (11.207) in the generating functional Z , and as before change the order of $\mathcal{D}A$ and $\mathcal{D}\omega$. The same arguments we used in obtaining Eq. (11.215) can isolate $\int \mathcal{D}\omega$ and we obtain

$$Z = \Delta[A] \frac{\int \mathcal{D}\omega}{\int \mathcal{D}h H[h]} \int \mathcal{D}A e^{iS[A]} H[g(A)] = N \int \mathcal{D}A e^{iS[A]} H[g(A)] \Delta[A] \quad (11.218)$$

It remains to evaluate $\Delta[A]$ and fix $H[g(A)]$.

11.7.3

Faddeev–Popov Ghost*

Lorentz Gauge

To proceed further, we need to fix the gauge condition. Let us choose the Lorentz gauge, namely

$$g_a(A) = \partial_\mu A_a^\mu \quad (11.219)$$

Then for A satisfying $\partial_\mu A^\mu = 0$

$$\begin{aligned} g_a(A^\omega) &= \partial_\mu \left(A_a^\mu + \frac{1}{g} \partial^\mu \omega_a + f_{abc} \omega_b A_c^\mu \right) \\ &= \frac{1}{g} \partial_\mu \partial^\mu \omega_a + f_{abc} A_c^\mu \partial_\mu \omega_b \end{aligned} \quad (11.220)$$

Hence

$$\begin{aligned} \left. \frac{\delta g(x)}{\delta \omega(y)} \right|_{\omega=0} &= \left[\frac{1}{g} \delta_{ab} \partial_\mu \partial^\mu + f_{abc} A_c^\mu \partial_\mu \right] \delta(x-y) \\ \therefore \Delta[A] &= \det \left[\frac{\delta_{ab}}{g} \partial_\mu \partial^\mu + f_{abc} A_c^\mu \partial_\mu \right] \end{aligned} \quad (11.221)$$

For the Abelian field, $f_{abc} = 0$ and $\Delta[A]$ does not depend on A . It can be absorbed in the normalization constant. However, for the non-Abelian field, $\Delta[A]$ depends on A and is therefore a dynamical object. It cannot be absorbed by the normalization constant. A convenient way is to rewrite it as a path integral in terms of anticommuting c-numbers. We cannot use the ordinary commuting c-numbers because the determinant has to be inverse in that case. We refer to Eq. (11.106), which is reproduced as the next equation:

$$\int \mathcal{D}\bar{\eta}(x) \mathcal{D}\eta(x) \exp \left[-i \int d^4x d^4y \bar{\eta}^*(x) A(x, y) \eta(y) \right] = \det iA \quad (11.222)$$

Using this formula, we have

$$\begin{aligned} \Delta[A] &= \int \mathcal{D}\bar{\eta}(x) \mathcal{D}\eta(x) \exp \left[i \int d^4x \mathcal{L}_{\text{ghost}} \right] \\ \mathcal{L}_{\text{ghost}} &= -\eta_a^*(x) [\delta_{ab} \partial_\mu \partial^\mu + g f_{abc} A_c^\mu(x) \partial_\mu] \eta_b(x) \end{aligned} \quad (11.223)$$

The factor $1/g$ is absorbed into the normalization of the η and $\bar{\eta}$ fields. By expressing the determinant as a path integral, we have introduced a new field. It has peculiar characteristics: it is a scalar field obeying the Klein–Gordon equation but is anticommuting. It is called the Faddeev–Popov ghost. As it has a color index, there are as many (i.e. N_c) ghost fields as gauge fields. It interacts only with the gauge field through derivative couplings. Since there is no reason to consider it as a physical particle we do not introduce an external source and hence it only appears in the intermediate states, namely, in closed loops in Feynman diagrams.

We still have the freedom to adopt an arbitrary functional $H[h]$ and we choose it to be a Gaussian integral

$$H[h] = \exp \left[-\frac{i}{2\lambda} \int d^4x g^2 \right] = \exp \left[-\frac{i}{2\lambda} \int d^4x (\partial_\mu A_a^\mu)^2 \right] \quad (11.224)$$

It amounts to changing the gauge field Lagrangian

$$\mathcal{L} \rightarrow \mathcal{L} - \frac{1}{2\lambda} (\partial_\mu A_a^\mu)^2 \quad (11.225)$$

which is exactly the gauge-fixing term we introduced before to circumvent the difficulty (see Eq. (11.59)).

In summary, the final form for Z , taking note of Eqs. (11.223) and (11.224), takes the form

$$Z = N \int \mathcal{D}A \mathcal{D}\bar{\eta} \mathcal{D}\eta \exp \left[i \int \mathcal{L}_{\text{nAbel}} \right] \quad (11.226a)$$

$$\mathcal{L}_{\text{nAbel}} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{FPG}} \quad (11.226b)$$

$$= \mathcal{L}_{\text{free}} - \frac{1}{2\lambda} (\partial_\mu A_a^\mu)^2 - \bar{\eta}^a M_{ab} \eta^b \quad (11.226c)$$

$$M_{ab} = \delta_{ab} \partial_\mu \partial^\mu + f_{abc} A_c^\mu \partial_\mu \quad (11.226d)$$

Axial Gauge

The ghosts have the wrong spin-statistics and appear only in intermediate states. Their existence, however, is essential to retain the unitarity of the theory. This is analogous to the scalar photon with indefinite norm, i.e. negative probability, that appeared in the covariant Lorentz gauge of the Abelian theory. In the non-Abelian gauge, for instance in quantum chromodynamics (QCD), similar situations occur for scalar and longitudinal gluons. In addition, because of the nonlinear nature of the non-Abelian theory, they couple to themselves. The creation of unphysical gluon pairs cannot be canceled by themselves. The ghost’s role is precisely to compensate these unwanted contributions and to maintain the unitarity of the theory order by order. In the Abelian case, when the Coulomb gauge was adopted only physical photons appeared and we did not need unphysical photons. Similarly, we may ask if there is a gauge in non-Abelian gauge theory in which the ghosts do not appear. Indeed there is: the axial gauge,³⁾ where $A_a^3(x) = 0$. More generally, we define the axial gauge with a unit vector t^μ that satisfies the condition

$$t_\mu A_a^\mu = 0, \quad t_\mu t^\mu = \pm 1 \quad (11.227)$$

Then the gauge-fixing term is

$$g_a(A) = t_\mu A_a^\mu \quad (11.228)$$

and under the gauge transformation Eq. (11.203)

$$\delta g_a[A^\omega] = f_{abc} A_b^\mu t_\mu \omega_c + t^\mu \partial_\mu \omega_a = t^\mu \partial_\mu \omega_a \quad (11.229)$$

Hence

$$\frac{\delta g_a}{\delta \omega_b} = \delta_{ab} t^\mu \partial_\mu \quad (11.230)$$

This does not contain A_a^μ , and its contribution can be integrated out. We do not need ghosts in the axial (or temporal) gauge. However, just like the Coulomb gauge, the Feynman propagator is complicated in these gauges and calculation is more difficult.

3) By choosing t_μ to be timelike, the temporal gauge is included also.

12

Accelerator and Detector Technology

12.1

Accelerators

Particle physics explores the deepest inner structure of matter. The smallest size we can explore is directly proportional to the energy we can transfer to the particles. For this reason, accelerators of the highest possible energy are the major apparatus one needs. With them we can boost particles, make them collide and decompose and investigate their inner structure by looking at their debris. The role of detectors is to identify secondary particles, reconstruct their space-time structure by measuring energy-momentum and elucidate their nature. Particle physics is a science based on observation and experimentation. Its progress is inextricably related to the advent of technologies. The ability to devise an experiment to test theoretical hypotheses and to understand connections between natural phenomena and predictions induced from first principles are attributes that researchers in the field have to acquire. To interpret experimental data correctly, it is essential to grasp the essence of experimental techniques and their underlying principles and to recognize their advantages and disadvantages.

There are numerous examples from the past in which people, confused by false data, have taken a roundabout route. Acquiring correct knowledge of experimental techniques and cultivating one's own critical power help to assess the observed data correctly and enhance one's ability to devise new methods that can lead to new ideas. However, the general knowledge that is necessary for the experimental techniques is mostly based on electromagnetism, condensed matter physics, and the physics of molecules, atoms and nuclei. These are, to say the least, detached from elementary particle physics. To avoid writing a manual-like textbook on experimental methods by going into too much detail, I restrict myself here to describing only basic ideas and a few examples of representative accelerators and detectors. In this chapter we discuss only experimental tools. Important experiments are introduced whenever possible as we pick up each subject throughout the book.

12.2

Basic Parameters of Accelerators¹⁾

12.2.1

Particle Species

The basic idea of particle acceleration is to apply an electric field to charged particles. Let the charge of the particle be q , the applied voltage V (in volts) and assume the field is uniform. Then the particle is accelerated to energy

$$E = qV = q\varepsilon d \quad (12.1)$$

where $\varepsilon = dV/dt$ is the gradient of the electric potential and d the length of the field. When a charged particle has unit charge ($= 1.6 \times 10^{-19}$ Coulomb) and is accelerated by 1 Volt of electric potential, we say it has acquired the energy of 1 electron volt (eV for short). The electron volt is a small unit, and we list below larger units and their applied physics in that energy region:

keV	=	10^3 eV	: X-rays, structure of atoms and molecules
MeV	=	10^6 eV	: nuclear reactions
			: m_π (pion mass) = $139.6 \text{ MeV}/c^2$
			: hadron size $\sim \hbar/m_\pi c \simeq 1.4 \times 10^{-13} \text{ cm}$
GeV	=	10^9 eV	: m_p (proton mass) = $0.938 \text{ GeV}/c^2$
			: m_W (W boson mass) = $80.26 \text{ GeV}/c^2$
			: $(\sqrt{2}G_F)^{-1}$ (energy scale of electroweak interaction)
			: = 250 GeV
TeV	=	10^{12} eV	: typical energy of particle physics frontier
			: Note: total energy of LHC ²⁾ = 14 TeV

Particles that can be accelerated are usually limited to stable charged particles, which are electrons (e) and protons (p), including their antiparticles. When one wants to use a beam of unstable particles such as pions (π), kaons (K) and hyperons (Y), one irradiates protons on matter (beryllium, iron, etc.) and uses the secondary particles that are produced. In the case of neutrinos (ν) and muons (μ), tertiary particles, which are the decay products of secondary particles, are used. The energy of the secondary or tertiary beams is usually less than half that of the primary and the intensity is of the order of 10^{-6} – 10^{-7} of the primary. Until the unified theories were established, strong, electromagnetic and weak interactions were considered as independent phenomena and experiments with different particle species were considered as belonging to different research fields. Historically, hadron beams (p , π , K) on fixed targets were very powerful in discovering new particles and resonances. Charting and classifying them (hadron spectroscopy) was an important

1) [266].

2) The Large Hadron Collider (LHC) is a proton–proton collider at CERN that started operation in 2009.

discipline of the field. The hyperon beam [247], a secondary hadron beam that contains a large fraction of hyperons, contributed to measuring the magnetic moment of hyperons and elucidating their decay properties. Neutrino beam experiments in the 1970s [145] were used to discover the neutral current and, together with intense electron beam experiments, clarified the quark-parton structure of hadrons. Both approaches, spectroscopy and dynamics, were instrumental in establishing the unified theories. Muon beam experiments [124] extended the knowledge obtained from the electron beam to the deeper inner structure of the hadron and enhanced the accuracy obtained by the neutrino experiments, thereby clarifying the quark-gluon structure, and contributed to establishing QCD (quantum chromodynamics).

With the advent of the unified theories, however, it became possible to interpret various phenomena produced by different particles in a unified way. From this viewpoint, the energy is the most important parameter and the particle species plays a secondary role.³⁾ The discovery of new particles at the energy frontier has been accorded the foremost importance in testing the correctness of proposed models by a potential breakthrough. In terms of energy, colliding accelerators are far more advantageous than those utilizing fixed targets. From the accelerator technological point of view, the proton is much easier to accelerate to higher energies than the electron. Consequently, hadron colliders (usually p on p or $\bar{p}$) have an advantage compared to electron colliders as the energy frontier machine. However, the hadron is a quark composite while the electron is elementary. Therefore hadron-associated phenomena have an intrinsic difficulty in their analysis and theoretical interpretation. It is also not clean experimentally, because the hadron is a strongly interacting particle and produces a great number of soft gluons, which appear as QCD background to a reaction one wants to investigate. The search for a new particle in hadron reactions is literally looking for “a needle in a haystack”. Thus electron machines are usually more advantageous in doing precise and decisive experiments to test theoretical models, except for those unreachable in the electron machine.

Today, fixed target experiments are carried out mostly in cases where the intensity of the beam and the species of the particle are the most important assets, namely when high precision in flavor physics is at stake. Examples are CP violation tests, rare decay modes in kaons and B -mesons, and neutrino and muon experiments to probe lepton flavor mixing. Generally speaking, for collider experiments at the energy frontier, a general purpose detector that is a combination of a variety of devices covering the largest possible solid angle is required. On the other hand, a clear and definite purpose is ascribed to fixed target experiments. They are most often accompanied by unique and well thought out measuring devices specific to the purpose.

- 3) To be fair, since the Standard Model has been established, the origin of particle species is an unanswered question. The universe can live with just the first generation (u , d) quarks and (ν_e , e) leptons and the existence of the second and third families remains an unsolved “flavor problem”.

12.2.2

Energy

The highest energy is the most fundamental parameter of an accelerator because it determines the minimum size of the object being probed. Equally important is its luminosity, which determines the number of events one can sample. It may be compared with the clearness of an image if the energy is the magnification of a microscope. The minimum scale (λ) one can probe with an accelerator having energy E is related to the de Broglie wavelength of the incident particle:

$$\lambda = \frac{h}{p} \simeq \frac{\hbar c}{E} = \frac{2 \times 10^{-16}}{\sqrt{s}/2[\text{GeV}]} \text{ m} \quad (12.2)$$

E is the energy of colliding particles in the center of mass system. Throughout in this book, we use “ $\sqrt{s} = 2E$ ” to denote the total center of mass energy. s is a Mandelstam variable and is defined by the following formula using the energy-momentum of two colliding particles (see Sect. 6.4):

$$s = c^2(p_1 + p_2)^2 = (E_1 + E_2)^2 - (\mathbf{p}_1 c + \mathbf{p}_2 c)^2 \quad (12.3)$$

In the collider accelerator, $\mathbf{p}_1 = -\mathbf{p}_2$ and $\sqrt{s} = E_1 + E_2$, the total energy of colliding particles.⁴⁾ For fixed target experiments $\mathbf{p}_2 = 0$ and when $E_1 \gg m_1 c^2, m_2 c^2$

$$\sqrt{s} = \sqrt{(E_1 + m_2 c^2)^2 - (\mathbf{p}_1 c)^2} \simeq \sqrt{2E_1 m_2 c^2} \quad (12.4)$$

This means the equivalent center of mass energy of the fixed target experiments increases only proportionally to the square root of the accelerator energy and is inefficient compared to the collider accelerators. This is why, with a few exceptions, most of the important results in the last 30–40 years have been obtained in collider experiments.

Problem 12.1

The energy of the Tevatron at Fermilab in the USA has $\sqrt{s} = 1.96 \text{ TeV} = (980 + 980) \text{ GeV}$. Derive the center of mass energy when one uses this beam for a fixed target experiment and prove it is not possible to produce a W boson ($m_W = 80.4 \text{ GeV}$). Assume a hydrogen target, $m_p = 0.938 \text{ GeV}$.

Note the Tevatron [343] started as an accelerator with a beam energy of 400 GeV for fixed targets in 1973 and W production was one of its major aims.

Effective Energy of Proton Colliders The proton is a composite of quarks, and the energy that participates in the reactions is that of the constituent quarks. Assuming

4) For asymmetric colliders ($\mathbf{p}_1 \neq -\mathbf{p}_2$) such as B-factories (KEKB and PEP II), $\sqrt{s}$ has to be calculated using Eq. (12.3). The B -factory is a high luminosity accelerator that operates at $\sqrt{s} = 2m_B \sim 10 \text{ GeV}$ designed to produce copious B -particles, which contain b -quarks.

the quark in the proton has a fraction x of the parent proton momentum, the total energy $\sqrt{\hat{s}}$ of the colliding quarks is given by

$$\sqrt{\hat{s}} = \sqrt{(x_1 p_1 c + x_2 p_2 c)^2} \simeq \sqrt{x_1 x_2} \sqrt{s} \quad (12.5)$$

The average fractional momentum $\langle x \rangle$ has been measured in deep elastic scatterings by electrons and muons (see Chap. 17) to be 0.1–0.3, which means $0.01\sqrt{s} \lesssim \sqrt{\hat{s}} \lesssim \sqrt{s}$. However, the probability of quarks having $x_1 x_2 \sim 1$ is not zero and reactions with $\sqrt{\hat{s}} \simeq \sqrt{s}$ are not impossible. Whether the reaction is useful or not depends on the luminosity and the skill of experimenters in sorting out signals from the background.

12.2.3

Luminosity

Events that a detector picks up are mixtures of signals and background. The luminosity L plays a critical role in extracting clean signals out of unwanted backgrounds. It is defined as the quantity that gives an event rate when multiplied with the cross section.

$$N(\text{reaction rate/s}) = L\sigma \quad (12.6)$$

The unit of L is “per square centimeter per second” ($\text{cm}^{-2} \text{s}^{-1}$). The cross section σ is predetermined by nature and is not controllable. Therefore the number of events, a decisive factor in reducing statistical errors, is directly proportional to the luminosity. When large background signals exist, one has to filter them out usually by applying cuts in various kinematical variables, such as momenta, angles and spatial as well as time locations. However it is not possible to make an ideal filter that is capable of cutting out all the unwanted events while passing all the signals. In most cases, the signals are reduced, too. When the background signals are abundant, the events have to go through many filters and the number of surviving signals critically depends on the luminosity. The luminosity of a circular collider is given by [240, 316]

$$L = \frac{N_1 N_2 f}{n_b A}$$

N_1, N_2 : total number of particles in the beam

f : revolving frequency of the particles

n_b : number of bunches in a circle
(the beam is not continuous but bunched)

A : the beam's cross-sectional area at the colliding point

$$A = \sqrt{\varepsilon_x \beta_x} \sqrt{\varepsilon_y \beta_y}$$

β_x, β_y : beam amplitude at the interaction point

$\varepsilon_x, \varepsilon_y$: beam emittance

The beam emittance is defined as a phase space area in the $(\beta_i, \beta'_i = d\beta_i/dz)$ -plane of the particles in the beam, where z is the coordinate in the beam direction. Namely

$$\varepsilon_i = \int \beta'_i d\beta_i \quad (12.7)$$

Luminosities that can be achieved by electron and proton beams are slightly different. $L_e = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ has been achieved at B-factories, which is an approximately 10^2 increase compared to their predecessors, and $L_p \simeq 3 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ has been achieved at the Tevatron. At LHC the designed luminosity is $L_p = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and up to $10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ is envisaged in the future. For fixed-target experiments, the reaction rate $N \text{ (s}^{-1}\text{)}$ can be calculated using the number of incoming particles (N_{in}) and target particles (N_{tgt}) as

$$\begin{aligned} N &= N_{\text{in}} \sigma N_{\text{tgt}} = N_{\text{in}} \sigma \rho \ell N_A \\ \rho &: \text{density of the target in g/cm}^3 \\ \ell &: \text{the length of the target in cm} \\ N_A &: \text{Avogadro number} = 6.022 \times 10^{23} \text{ mol}^{-1} \end{aligned} \quad (12.8)$$

which gives

$$L = N_{\text{in}} \rho \ell N_A \quad (12.9)$$

As a typical example, taking $\rho = 10 \text{ g/cm}^3$, $\ell = 1 \text{ cm}$, $N_{\text{in}} = 10^7 \text{ s}^{-1}$ (secondary beam) to 10^{12} s^{-1} (primary beam) gives $L = 10^{32}$ to 10^{37} s^{-1} . Until the end of the 20th century, the luminosity at a fixed target was generally larger but the gap is closing fast. In electron collider experiments, a basic cross section is given by the process $e^- e^+ \rightarrow \mu^+ \mu^-$. Note, in the Standard Model, for colliding $e^- e^+$ all the production cross sections by the electroweak interaction are of the same order and the following number is a good number to keep in mind:

$$\begin{aligned} \sigma_{\mu\mu} &\equiv \sigma(e^- e^+ \rightarrow \mu^+ \mu^-) \\ &= \frac{8\pi}{3s} = \frac{87[\text{nb}]}{s[\text{GeV}^2]} \sim \frac{10^{-31}}{s[\text{GeV}^2]} [\text{cm}^2] \end{aligned} \quad (12.10)$$

where nb means nanobarn. At $\sqrt{s} = 10 \text{ GeV}$, $L = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, $N \sim 100 \text{ s}^{-1}$.

Signal-to-Noise Ratio (S/N) As the proton is a quark composite, quarks that have not participated in the reaction (called spectators) escape from the parent proton, radiating soft gluons, which hadronize, producing jets and constituting the so-called QCD backgrounds. For this reason, clean signals (with large S/N ratio) are hard to obtain in hadron colliders. Suppose one wants to detect a Higgs meson. Its production cross section is $\sim 10^{-35} \text{ cm}^2$. Furthermore, the decay branching ratio to lepton pairs ($H^0 \rightarrow e^- e^+, \mu^- \mu^+$), which are suitable for identifying the Higgs,

is $\sim 10^{-4}$. In contrast, a typical hadronic cross section that appears as QCD backgrounds in hadron colliders is $\sigma \sim 10^{-25} \text{ cm}^2$ and with $L = 10^{32}$ to $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ produces events at a rate of $N = 10^7$ to 10^{10} s^{-1} , a formidable number to cope with. The S/N ratio at its most naive level is estimated to be 10^{-14} , a challenge for experimentalists.

At electron colliders, production cross sections of interest relative to that of $e e \rightarrow \mu \mu$ are usually of the order of 10^{-1} and that of hadronic production is $\sim 3 - 4$, therefore S/N ratios are much better. Generally speaking, electron colliders are advantageous in that because the incoming particles are elementary theoretical interpretations are straightforward. Furthermore, the whole energy of the colliding electrons can be used to produce the wanted particles and the backgrounds are few. Therefore, signals are cleaner and easier to interpret theoretically. However, the electron, being much lighter than the proton ($m_e/m_p \simeq 2000$), emits photons if bent by magnetic fields (synchrotron radiation). Synchrotron light source facilities take advantage of this and are extensively used in the fields of materials science as well as atomic, molecular and biophysics. From the high energy physics point of view, it is a pure loss. Since the loss increases as $\sim E^4$, it is increasingly difficult to construct a high energy, high luminosity circular accelerator for electrons. It is believed that the LEP⁵⁾ at CERN with its energy $\sqrt{s} = 200 \text{ GeV}$ is about the limit of a circular machine. If higher energy is desired, one has to rely on linear colliders. So far only one linear collider has been operational, the SLC at SLAC with $\sqrt{s} \sim 90 \text{ GeV}$, and its technology is not as mature as that of circular accelerators.

In summary, with hadron colliders, innovative approaches to detector designs are required to cope with copious backgrounds, whereas with electron colliders developing advanced accelerator technology is the critical factor. With present technology, when physics of the TeV region is to be investigated, hadron machines are easier to construct, but linear colliders of the TeV range are a must for future experiments and extensive R&D is being carried out.

12.3

Various Types of Accelerators

12.3.1

Low-Energy Accelerators

One accelerator on its own, whatever its mechanism, cannot accelerate particles above an energy of about 1 GeV, and a variety of accelerators are utilized in cascade to achieve this. The final stage of a high energy accelerator complex, so far, with the exception of the SLAC linear accelerator (SLC), has always been a synchrotron, a kind of circular accelerator. It confines particles inside a circular pipe with fixed radius, places an accelerator tube at one point or two and accelerates

5) An electron collider that operated from 1989 to 2000 with $\sqrt{s} = 2m_z(180 \text{ GeV}) \sim 200 \text{ GeV}$ which provided precision data to confirm the validity of the Standard Model.

the particles every time they pass the tube. As a single acceleration tube can be used many times, it is very effective and has a high performance-to-cost ratio. Electrons or protons emanating from a source are first accelerated by a static electric field accelerator (usually a Cockroft–Walton accelerator) to a few hundred thousand volts, then further accelerated by a linac to hundreds of million electron volts, and finally injected into a synchrotron. As the particles can be accelerated more effectively if the ratio of the final energy to the injected energy is small, two or more synchrotrons are used in cascade. The first of them is often called a booster or an accumulator and that in the final stage the main ring.

Cockroft–Walton This type of accelerator uses a static electric field to accelerate particles and has the structure shown in Fig. 12.1. A combination condenser–diode circuit is arranged to work as a parallel connection in charging the capacitors and a series connection in discharging them. Figure 12.1 shows an eightfold multiplication of the applied voltage. The Cockroft–Walton accelerator used for accelerating the 12 GeV proton synchrotron (hereafter referred to as PS) at KEK⁶⁾ produced 750 kV.

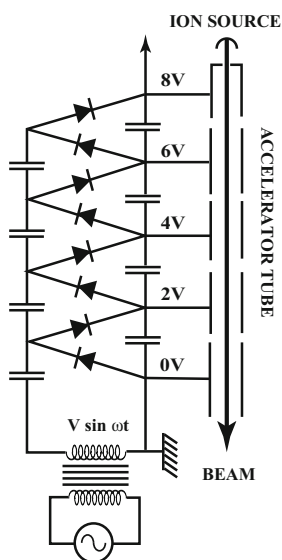


Figure 12.1 Cockroft–Walton accelerator: The circuit is connected in such way that applied voltages are arranged in parallel when charging and in series when discharging, giving the n -fold ($n = 8$ in the figure) voltage.

Linac Protons accelerated by the Cockroft–Walton accelerator are injected into a linear accelerator (linac). It is composed of a series of electrodes in a cavity (a hollow tube made of copper), to which an alternating electric voltage is applied by a radio frequency (RF) supplier (see Fig. 12.2).

The polarity of the applied voltage is arranged to decelerate the particles when they pass the cavities and to accelerate them when they cross the gaps between

6) High Energy Accelerator Organization in Japan.

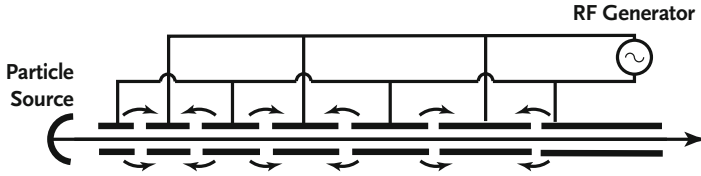


Figure 12.2 Proton Linear Accelerator: The accelerating RF voltage is applied to particles when they cross gaps and a decelerating voltage is applied when they are in the tube. As the particles gain energy, their velocity increases and the corresponding tube is longer.

them. They feel no voltage inside the cavity because of its electric shielding. The flow of injected protons is continuous, but only those synchronized with the applied voltage are accelerated and consequently they emerge bunched as a series of discontinuous bundles. The length of the cavity needs to be longer as the particles proceed because they fly faster each time they are accelerated. The accelerating voltage is usually limited to a few MeV (million electron volts) per meter, and an energy of 100–200 MeV is reached at the final stage. Electrons acquire near-light velocity at a few MeV, so the length becomes constant after a meter or two. The longest electron linac is SLC at SLAC (Stanford Linear Accelerator Center), which has $\sqrt{s} = 90 (45 + 45)$ GeV, and a length of 3 km. It uses 230 klystrons (large vacuum-tube RF amplifiers) and makes two beams of electrons and positrons collide 60 times per second. The positrons are made by bombarding high-energy electrons (34 GeV at SLC) on matter, returning them to the starting point and re-accelerating them using the same accelerator tube a half-period of the RF later.

Cyclotron A charged particle with its velocity perpendicular to a uniform magnetic field makes a circular motion. When the particle with mass m , charge q and momentum p moves on a circular orbit of radius R in a magnetic field of strength B , the equation of motion with centrifugal force supplied by the magnetic field is given by

$$\frac{mv^2}{R} = qvB \quad (12.11)$$

$$p[\text{GeV}/c] = mv = qBR = 0.3BR[\text{T m}]$$

where T stands for tesla. The equation is valid at relativistic energy if the relativistic mass ($m \rightarrow m\gamma = m/\sqrt{1 - (v/c)^2}$) is used. The radius of curvature is proportional to the momentum and the inverse of the field strength. The angular frequency is given by

$$\omega = 2\pi \frac{v}{2\pi R} = \frac{qB}{m\gamma} \quad (12.12)$$

and is referred to as the cyclotron frequency. When the particle motion is nonrelativistic ($v \ll c$), ω is a constant if the magnetic field strength is constant. In the cyclotron a pair of half-cut hollow discs called D-electrodes or “dees” are placed in

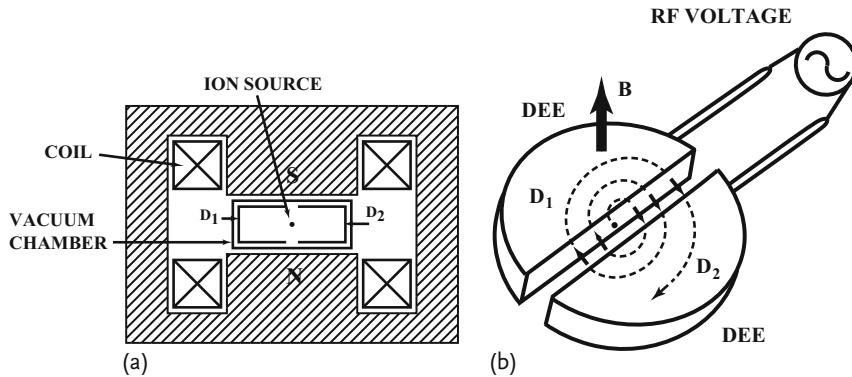


Figure 12.3 Layout of a cyclotron: (a) The cross section. Two electrodes, the dees, are in the magnetic field. (b) RF voltage is applied between the two dees to accelerate particles

when they pass the gap. The orbital radius grows as the particles are accelerated, but the revolution frequency stays the same.

a uniform magnetic field (Fig. 12.3a) and a RF voltage is applied across the dees. The particles, injected near the center of the magnetic field, are accelerated only when they pass through the gap between the electrodes. The combined effect of the magnetic field and the acceleration forces the particles to travel in a spiral path and eventually leave the dees (Fig. 12.3b).

Synchrocyclotron As particles are accelerated, their mass increases, the cyclotron frequency is no longer constant and their circular motion is not synchronous with the accelerating field. The cyclotron's acceleration is effectively limited to $v/c \lesssim 0.1$ or in the case of protons about 15 MeV. To overcome this limit, either ω or B has to be changed. The synchrocyclotron increases the maximum energy by decreasing the frequency as the mass increases. The disadvantage of this method is that only clusters of particles can be accelerated. Once the RF is synchronized to an injected cluster, the next cluster has to wait until the preceding cluster has reached its minimum frequency and is extracted. The intensity is reduced to a fraction ($\sim 1/100$) of that of the cyclotron.

Isochronous Cyclotron In the isochronous cyclotron, in contrast to the synchrocyclotron, the RF is kept constant and the magnetic field is synchronized with the mass increase. This is achieved by dividing the poles of the magnet into concentric sectors, and this type of cyclotron is referred to as a sector or ring cyclotron. The particles fly freely between the sectors and their trajectories are bent only when they pass through the sector magnets. The shape of the sectors is designed to keep the combined total flight time in free space and in the sectors constant as the particle becomes heavier. Thus the sector cyclotron can achieve high energy (several hundred MeV for protons) without loss of intensity. Figure 12.4 shows an example, the 590 MeV Ring Cyclotron of PSI (Paul Sherrer Institute in Switzerland), from

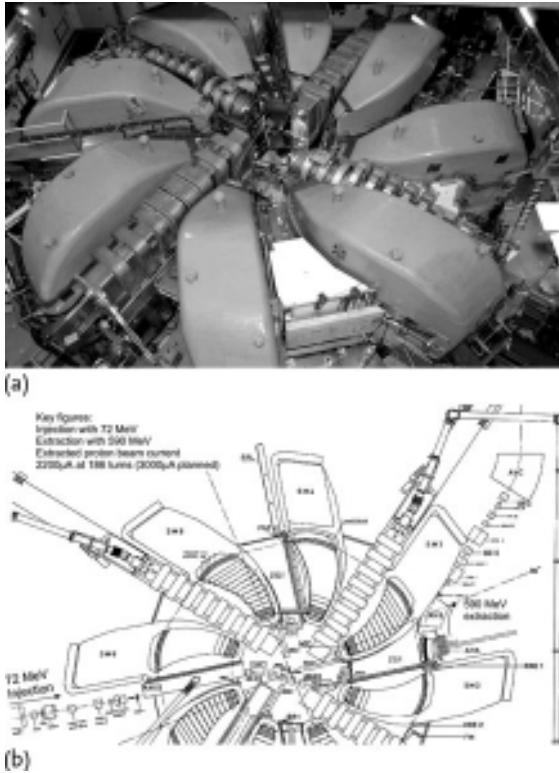


Figure 12.4 (a) PSI 590 MeV Ring Cyclotron. The beam is accelerated first to 870 kV by a Cockroft–Walton accelerator, then pre-accelerated to 72 MeV by a small ring cyclotron and injected into the 590 MeV Ring Cyclotron. (b) Floor layout of the Ring cyclotron. Four accelerating tubes (KAV1–4) and eight sec-

tor magnets (SM1–8) are shown. The sector magnets are of C-type with the open edge directed toward the center. The beam comes in from the left, is injected into the inner edge of SM3 at the upper right and extracted from AHB at the upper right. Photo courtesy of Paul Scherrer Institute [328].

which a 2 mA DC beam can be extracted [328]. The highest energy is limited by the cost of making giant magnets.

12.3.2

Synchrotron

As is shown by Eq. (12.11), if one wants to increase the energy, either B or R or both has to be increased. The synchrotron keeps the radius of the circle constant by increasing the strength of the magnetic field. If one wants to go higher than, say, 1 GeV energy, it is the most practical method. It is accompanied by one or more pre-accelerators because a certain amount of energy is necessary to keep particles on a circular orbit of a definite radius. Pre-accelerators usually consist of a Cockroft–Walton accelerator followed by a linear accelerator. Sometimes two synchrotrons

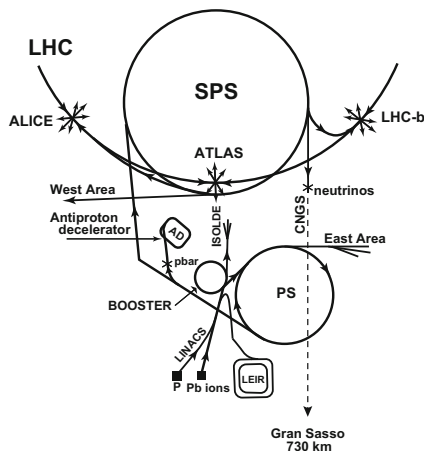


Figure 12.5 CERN accelerator complex: Protons are accelerated first to 50 MeV by a linac, then to 1.4 GeV by the booster, to 25 GeV by the PS (Proton Synchrotron), to 450 GeV by the SPS (Super PS) [133] and finally to 7 TeV by the LHC main ring. ISOLDE is used to study unstable nuclei. Part of the

beam is used for fixed-target experiments. For instance, CNGS (CERN Neutrinos to Gran Sasso) is a neutrino beam for neutrino oscillation experiments carried out at LNGS (the Gran Sasso National Laboratory) of INFN (Istituto Nazionale de Fisica Nucleare) in Italy 730 km away. Taken from [93].

are used in cascade to bring the energy even higher. This is because the acceptance of the injected beam can be made higher if the ratio of injected to maximum energy is lower, and there is a net gain in intensity. For this reason the first accelerator is called a booster and the last the main ring. Figure 12.5 shows CERN's⁷⁾ accelerator complex [93]. Protons are first accelerated to 50 MeV by a linac, then to 1.4 GeV by the booster, to 28 GeV by the PS (Proton Synchrotron), to 450 GeV by the SPS (Super PS) and finally to 7 TeV by the LHC main ring. PS was the main ring between 1960 and 1980 and is still used for a variety of experiments. SPS was built originally as an antiproton–proton collider and achieved the historical discovery of W and Z^0 bosons. It also served as an intermediate accelerator for electrons and positrons to be injected into the Large Electron–Positron Collider (LEP). LEP also used to be the LHC main ring; it operated from 1989 to 2000 at $\sqrt{s} = m_Z$ and later at $\sqrt{s} = 200$ GeV. It was the largest electron collider and instrumental in establishing the Standard Model of particle physics.

Elements of the Synchrotron The magnets that are used to keep the particles on their circular trajectories are of dipole type, like the one shown in Fig. 12.6a, or C-type, obtained by cutting the magnet at the center.

The accelerating tube is a RF cavity and is placed at a point on the circle. The frequency is synchronized so that it accelerates the particles every time they pass the

7) European Organization for Nuclear Research. The acronym CERN originally stood for the earlier French name Conseil Européen pour la Recherche Nucléaire. See for example [93] for more details.

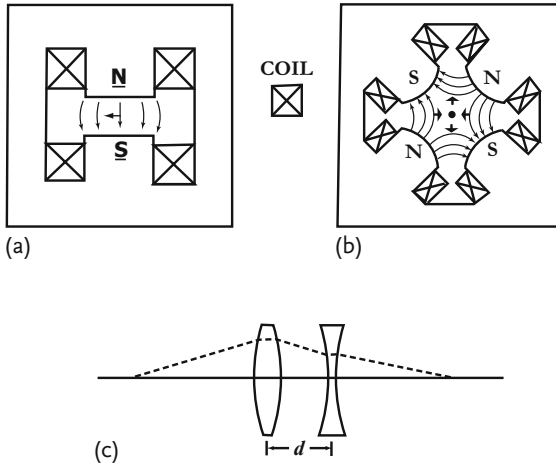


Figure 12.6 Cross sections transverse to the particle path of a typical dipole (a) and a quadrupole magnet (b). Thin lines with arrows denote magnetic field directions and thick arrows denote magnetic forces. In the dipole, the force is one directional, but in the

quadrupole, it is inward (converging) horizontally but outward (diverging) vertically.

(c) A pair of dipoles can make the particle ray convergent in both the x and y directions, acting like a combination of convex and concave lenses.

RF cavity. With present technologies, the maximum field strength of the magnet is limited to ~ 1.4 T (normal conducting) or 3 T (Tevatron) to 10 T (LHC) (superconducting).⁸⁾ Let us consider the cases of the Tevatron (980 + 980 GeV) and LHC (7 + 7 TeV). Inserting $B = 3$ T (3.8 T), $R = 1$ km (2.8 km), Eq. (12.11) tells us that $E \sim pc \sim 900$ GeV (7 TeV). Namely, the size of the accelerator becomes larger in proportion to the maximum energy. The revolution frequency at the Tevatron (LHC) is $\sim 10^{-5}$ (10^{-4}) and the accelerating voltage is ~ 1 MeV; it takes ~ 10 s (~ 700 s) to reach the maximum energy. When a continuous beam is injected into the linac, it becomes bunched in the synchrotron with a cluster length of ~ 1 mm.

Beam Stability Usually particles in the beam have dispersion (spread) in their velocity as well as their position. The accelerator will work only when the particles that are not exactly synchronous with those in the central orbit tend to converge or oscillate around the center. If their motion diverges, they will not be accelerated stably.

i) Betatron Oscillation Stability Transverse oscillation relative to the orbit is called betatron oscillation. To control the betatron oscillation and keep the particles on the orbit, quadrupole-type magnets are used. They have the structure shown in Fig. 12.6b. With the configuration depicted in the figure, the magnetic field exerts a converging force (focus, F) on the particles in the x direction and diverging force

8) In laboratories, if only a small volume is activated for a short time, magnetic field strengths of several hundred tesla have been achieved. Here, we are talking of industrial level magnets that are of large size and capable of stable operation.

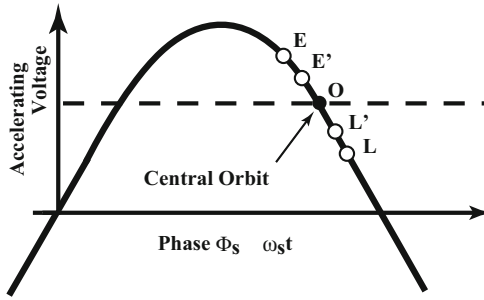


Figure 12.7 Synchrotron oscillation and phase stability. Point O denotes where the revolution frequency and the acceleration frequency synchronize. A particle which arrives early (E) (late, L) receives larger (smaller) acceleration

and its orbital radius becomes larger (smaller). Therefore, it takes more (less) time to go around and arrives later (E') (earlier, L') at the next period.

(defocus, D) in the y direction. The quadrupole magnet is a kind of convex lens and concave lens in optics in one. Just as a combination of concave and convex lenses placed at a distance d has a combined focal length F , which is given by

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2} \quad (12.13)$$

a pair of quadrupole magnets acts as a focusing device. If we take, for example, $f_1 = -f_2 = f$, then $F = d/f^2$ and the pair can exert a converging force in both the x and y directions (Fig. 12.6c). In the synchrotron dipole magnets are placed at equal distances to keep the particles moving on the circular orbit and quadrupoles are distributed between the dipoles in a configuration (FDFDFD...). The use of many FD combinations is called strong focusing.

ii) Phase Stability Longitudinal oscillation in the beam direction around the central point synchronized with the applied RF voltage is called synchrotron oscillation. Suppose the circular motion of particles at point O (see Fig. 12.7) is synchronous with the RF frequency and they come exactly back to point O after a period. The phase of point O is called synchronous phase. Particles at E, which arrive earlier, experience higher acceleration, which causes their orbit to expand and consequently they take more time to circulate, arriving later at E' in the next period. On the other hand, particles at L, arriving later, receive less acceleration and arrive earlier, at L', next time. Therefore the particles' phase tends to converge to the synchronous phase and thus the beam can be accelerated stably. The convergence of the longitudinal motion in the particles' orbit is referred to as phase stability.

Synchrotron radiation, which is described in the next paragraph, attenuates both the betatron and synchrotron oscillation and thus has the effect of stabilizing the beam.

Synchrotron Radiation Synchrotron radiation is the phenomenon that a particle emits light as its path is bent by the magnetic field. Because of this, the energy

cannot be made arbitrarily large proportional to the radius. At some point the energy loss due to synchrotron radiation becomes so large that the compensation of the energy becomes technically formidable. The energy loss ΔE per turn of the synchrotron radiation is given by

$$\Delta E = \frac{4\pi\hbar c\beta^3\gamma^4}{3R} \quad (12.14)$$

$$\Delta E[\text{MeV}] \sim \frac{0.0885(E[\text{GeV}])^4}{R[\text{m}]} \quad : \quad \text{for the electron}$$

At LEP, $R = 2804 \text{ m}$ and $E = 100 \text{ GeV}$, which gives $\Delta E = 3.2 \text{ GeV/turn}$. Power/RING = 16 MW. At LHC, $E = 7 \text{ TeV}$, which gives $\Delta E = 6.7 \text{ keV/turn}$.

Synchrotron radiation is almost monochromatic and collimated, which makes a high quality light source but is a loss as an accelerator. The proton being 2000 times heavier than the electron, the synchrotron radiation loss in the proton machine is $(2000)^4 \sim 10^{13}$ smaller and hence has never been a problem. LHC is the first machine to consider the loss seriously, but at $E = 7 \text{ TeV}$, $R = 2804 \text{ m}$ the energy loss is a mere 6.7 keV/turn . This is the main reason for the technologically available maximum energy of the electron accelerator being roughly a tenth that of the proton machine.

12.3.3

Linear Collider

As the energy loss due to synchrotron radiation increases rapidly, like $\sim E^4$, and because of the already formidable energy loss, a circular electron machine beyond the size of the LEP is considered impractical. This is why a linear collider is being considered as the future machine for electron–positron collisions. The first working linear collider was SLC (Stanford Linear Collider), which operated at $\sqrt{s} = 90 \text{ GeV}$. The technology of the linear collider is not mature and SLC could barely compete with LEP in producing physics results. However, for precision physics beyond the TeV region, it is considered a must and a team of collaborators from all over the world is working to realize the ILC (International Linear Collider). A starting energy of $\sqrt{s} = 500 \text{ GeV}$ is envisaged, later to be increased to $1 \sim 2 \text{ TeV}$. The main difficulties of the linear collider are:

1. Unlike circular accelerators, an accelerating tube can accelerate particles only once on a trajectory. Therefore, an increase of the acceleration voltage per unit length is necessary from the present level of $\sim 1 \text{ MeV/m}$ to $\sim 100 \text{ MeV/m}$. If this is achieved, the total length of the accelerator becomes $1 \text{ TeV}/100 \text{ MeV} \sim 10 \text{ km}$.
2. As two beams can collide only once, the collision probability has to be increased, too. This requires the colliding beam size to be on order of $\sim 0.1 \mu\text{m}$ or less compared to $\sim 10 \mu\text{m}$ at present.
3. As a corollary to the second item, technologies have to be developed to maintain and control the accuracy of $0.1 \mu\text{m}$ over a distance of 10 km .

12.4

Particle Interactions with Matter⁹⁾

12.4.1

Some Basic Concepts

A particle detector is a device that determines the energy, momentum, spatial position or time response of an incoming particle by inducing interactions with specific materials suitable for the purpose. It is useful to understand basic features of interactions that are used for measurement before discussing individual instruments. Particles are classified roughly into two classes depending on their interactions with matter: hadrons, which interact strongly, and electrons, photons and muons, which interact only electromagnetically. From the detector point of view, weak interactions (and gravity, too) can be neglected. Neutrinos are usually recognized as “missing energy” except for specific devices to detect neutrino interactions directly. Suppose a particle with definite energy and direction passes through a material. It interacts either electromagnetically (Fig. 12.8a, b) or strongly (c, d). A strongly interacting particle is usually elastically scattered if the energy is low ($\lesssim$ a few hundred MeV), but at high energy ($E \gg m_p c^2 \sim 1$ GeV) its reaction is inelastic, producing one or more particles. From the incident particle point of view, the energy is regarded as lost or absorbed by the material. In the case of electromagnetic interactions, the incident particle loses its energy ($E \rightarrow E - \Delta E$) by scattering electrons (namely ionizing atoms). It also undergoes multiple scattering, a series of small-angle scatterings by the atomic Coulomb potential, and changes its angle $\theta \rightarrow \theta + \Delta\theta$. If a particle is energetic enough, secondary particles produced by strong or electromagnetic interactions induce a series of inelastic reactions, making a shower of particles. Among the secondary particles, neutral pions decay im-

9) [138, 261, 311].

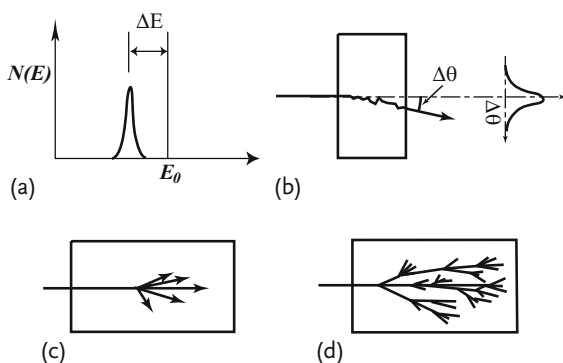


Figure 12.8 Particles' interactions with matter. Electromagnetic interaction: (a) The particle loses energy ΔE by ionizing atoms. (b) The particle changes direction by multiple scatter-

ing. Strong interaction: (c) Reaction with (absorption by) nuclei. (d) Shower phenomenon: a combination of strong and EM interaction.

mediately to two energetic gamma rays and by a combination of pair production and bremsstrahlung produces an electromagnetic shower. Therefore the shower phenomenon induced by hadrons is a combination of strong and electromagnetic interactions.

Collision Length and Absorption Coefficient We define a few quantities to give a measure of reaction probability.

Cross Section The cross section σ is defined as the probability of reaction when on average one particle per unit area comes in and hits a target whose density is one particle per unit volume. Then when N_{in} particles uniformly distributed over an area A hit a target consisting of N_{tgt} particles uniformly distributed in a volume $V = AL$, the number of reactions N is given by

$$N = A \times \frac{N_{\text{in}}}{A} \times \frac{N_{\text{tgt}}}{V} \times L \times \sigma = \frac{N_{\text{in}} N_{\text{tgt}}}{A} \times \sigma \quad (12.15)$$

The cross section has the dimension of an area and may be considered as the target's effective area size. The standard unit is

$$1 \text{ barn} \equiv 1 \text{ b} = 10^{-24} \text{ cm}^2 \quad (12.16)$$

but also

$$\begin{aligned} 1 \text{ mb (millibarn)} &= 10^{-3} \text{ b} &= 10^{-27} \text{ cm}^2 \\ 1 \text{ } \mu\text{b (microbarn)} &= 10^{-6} \text{ b} &= 10^{-30} \text{ cm}^2 \\ 1 \text{ nb (nanobarn)} &= 10^{-9} \text{ b} &= 10^{-33} \text{ cm}^2 \\ 1 \text{ pb (picobarn)} &= 10^{-12} \text{ b} &= 10^{-36} \text{ cm}^2 \\ 1 \text{ fb (femtobarn)} &= 10^{-15} \text{ b} &= 10^{-39} \text{ cm}^2 \\ 1 \text{ at (attobarn)} &= 10^{-18} \text{ b} &= 10^{-42} \text{ cm}^2 \end{aligned} \quad (12.17)$$

are used.

A process where particles A and B collide and become particles C (=A) and D (=B) is called elastic scattering. Sometimes when the particle species change ($C \neq A$, $D \neq B$, for example $\pi^- p \rightarrow K^0 \Lambda$), it is called quasielastic. If the final state consists of three or more particles C, D, E ..., it is called inelastic scattering. Let the fraction of (C + D) production of all the reactions be $BR(\text{CD})$ (branching ratio) and the total cross section σ_{TOT} , then the cross section of the reaction ($A + B \rightarrow C + D$) is given by

$$\sigma(A + B \rightarrow C + D) = \sigma_{\text{TOT}} BR(\text{CD}) \quad (12.18)$$

The angular distribution of secondary particles contains important information. The fraction of the cross section in which the secondary particle is scattered into solid angle $d\Omega$ at angle (θ, ϕ) (Fig. 12.9) is called the differential cross section and is defined as

$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega = \int \frac{d\sigma}{d\Omega} \sin \theta d\theta d\phi \quad (12.19)$$

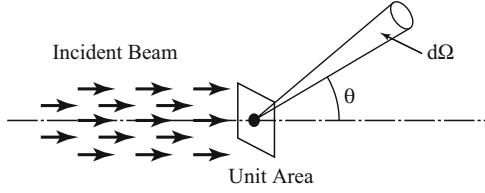


Figure 12.9 When a flux of particles whose intensity is one particle per unit area per unit time is incident on a target consisting of a single particle and dN particles per unit time are

scattered into the solid angle $d\Omega$ at (θ, ϕ) , the differential cross section is defined by $\frac{d\sigma}{d\Omega}(\theta, \phi) = dN$.

When a particle moves a distance d in matter, the number of collisions it experiences is given by Eq. (12.15) by putting $N_{\text{in}} = 1$, $A = \text{unit area}$, $V = d$

$$N_{\text{coll}} = \sigma n d \quad (12.20)$$

where n is the number density of target particles in the matter. Therefore the average length before a particle experiences a collision is given by

$$\lambda = \frac{1}{\sigma n N_A} \quad (12.21)$$

where ρ is the matter density [g/cm^3] and $N_A = 6.023 \times 10^{23} \text{ g}^{-1}$ is the Avogadro number. λ is referred to as the collision length or the absorption length when only inelastic collisions are counted.

The number of particles passing through the matter decreases as they interact. If the length is sufficiently short, the rate can be considered to be proportional to the length. As the probability that a particle is lost while moving a distance dx is given by dx/λ , the number of particles lost dN out of N particles is expressed as

$$\frac{dN}{N} = -\frac{dx}{\lambda} \equiv -k dx \quad (12.22)$$

k is called the absorption coefficient. Solving this equation gives

$$N = N_0 \exp(-kx) = N_0 \exp(-x/\lambda) \quad (12.23)$$

where N_0 is the initial number of particles. It is often more convenient to use not x , the length, but the weight per unit length ($\xi = \rho x$). Denoting the flux (the number of particles passing unit area per unit time) as I

$$I = I_0 \exp(-\mu \xi) \quad (12.24)$$

$$\xi = \rho x, \quad \mu = \frac{1}{\lambda \rho}$$

μ is referred to as the mass absorption coefficient.

When the energy of strongly interacting particles, such as protons and pions, exceeds ~ 10 GeV, the total cross section is ~ 25 mb and the collision length in matter with density ~ 1 g/cm³ is ~ 67 cm. The cross section of electromagnetic scattering by electromagnetic interaction (not ionization), for instance, Thomson scattering (elastic scattering) by protons is $\sigma \sim (8\pi/3)\alpha^2(\hbar/m_p c)^2 \sim 0.2$ μ b, which gives the mean collision length ~ 100 km. The cross section of weak interactions is, taking as an example neutrinos from the sun ($E \sim 1$ MeV), $\sigma \sim 10^{-44}$ cm² and the collision length is ~ 1 light year in matter with $\rho \sim 1$ g/cm³. Thus the collision length gives a measure to distinguish one type of interaction from another. Examples that are often mentioned in the instrumentation, for instance in the design of a calorimeter, are the collision and absorption lengths of iron (Fe): $\lambda \simeq 10.5$ cm and 16.8 cm, respectively.

A good guess at the strong interaction rate is given as follows. Hadrons such as nucleons constantly emit and reabsorb pions and are considered as an extended object whose size is of the order of the pion Compton wavelength ($\lambda_\pi = \hbar/m_\pi c \simeq 1.4 \times 10^{-13}$ cm). As a nucleus with atomic mass A can be considered as densely packed nucleons, it is a good approximation to treat the nucleus as a sphere with its radius approximated by $r_A = A^{1/3}\lambda_\pi$. If a hadron hits a nucleus, it reacts immediately at the surface and the nucleus looks like a totally absorbing black disk. The area of the disk is given by

$$\sigma_A = 2\pi \left(\frac{\lambda_\pi}{2} \right)^2 A^{2/3} \simeq 3.2 \times 10^{-26} A^{2/3} \text{ cm}^2 \quad (12.25)$$

This is called the geometrical cross section. In practice, adopting the measured cross section of πp scattering, the empirical formula

$$\sigma_A = 2.5 \times 10^{-26} A^{2/3} \text{ cm}^2 \quad (12.26)$$

may be used as a measure.

12.4.2

Ionization Loss

Ionization Loss per Unit Weight ($dE/d\xi$) The process of losing energy by ionizing atoms while passing through matter is the most fundamental phenomenon applied for measurement of charged particles [15]. It is important to understand the phenomenon in detail. Electrons are bound in atoms and although they are not at rest their motion is sufficiently slow that they can be approximated as a static target. When a particle with charge ze passes through matter with velocity v , and if its mass M is large compared with that (m) of electrons ($M \gg m$), the recoil of the particle by scattering electrons can be ignored and the particle may be considered to follow a straight line track. In the following, we try to calculate the energy loss

10) Quantum mechanically, the total cross section of a black disk is $2\pi R^2$, taking into account shadow scatterings.

using the quantum field theories we learned in Chap. 7 as much as possible. During the calculation we use natural units but recover $\hbar$ and c when expressing the final results. Let j_e^μ be the current the electrons make and j_m^μ that of the incoming charged particle. Then the scattering amplitude is given by [see Eq. (7.3)]

$$\begin{aligned} i T_{fi} &= \int d^4x d^4y \langle f | T [\{-i e j_e^\mu(x) A_\mu(x)\} \{-i z e j_m^\nu(y) A_\nu(y)\}] | i \rangle \\ &= -z e^2 \int d^4x d^4y \langle f | j_e^\mu(x) | i \rangle \langle 0 | T [A_\mu(x) A_\nu(y)] | 0 \rangle \langle f | j_m^\nu(y) | i \rangle \end{aligned} \quad (12.27)$$

Putting the energy-momentum of the initial and final state $p_i = (m, \mathbf{0})$, $p_f = (E_f, \mathbf{p}_f) \equiv (E, \mathbf{p})$, we find the matrix element of the electron current becomes

$$\langle f | j_e^\mu(x) | i \rangle = \bar{u}(p_f) \gamma^\mu u(p_i) e^{-i(p_i - p_f) \cdot x} \quad (12.28)$$

$\langle 0 | T [A_\mu(x) A_\nu(y)] | 0 \rangle$ is the Feynman propagator given by Eq. (6.115):

$$\langle 0 | T [A_\mu(x) A_\nu(y)] | 0 \rangle = -\frac{i g_{\mu\nu}}{(2\pi)^4} \int d^4q \frac{e^{-iq \cdot (x-y)}}{q^2 + i\epsilon} \quad (12.29)$$

If we neglect the recoil and energy loss of the charged particle, it can be treated classically. Assuming the particle follows a straight track along the z -axis, we have

$$\begin{aligned} j_m^\nu(y) &= \int_{-\infty}^{\infty} d\tau u^\nu \delta^4(y^\mu - u^\mu \tau) = v^\nu \delta(y^1) \delta(y^2) \delta(y^3 - vt) \\ v^\nu &= (1, 0, 0, v) \end{aligned} \quad (12.30)$$

where $u^\mu = (\gamma, \gamma \mathbf{v})$ is the particle's four-velocity and $\tau (= t/\gamma)$ is the proper time. Integration of x gives $(2\pi)^4 \delta^4(q + p_i - p_f)$, and further integration of q makes it $(2\pi)^4$. Integration of y gives the Fourier transform of j_m^ν :

$$\begin{aligned} j_\omega^\nu &\equiv \int d^4y j_m^\nu(y) e^{iq \cdot y} = \frac{v^\nu}{v} \int_{-L/2}^{L/2} dz e^{i(\omega/v - q_z)z} = \frac{v^\nu}{v} 2\pi \delta\left(q_z - \frac{\omega}{v}\right) \\ q^\mu &= (\omega, \mathbf{q}) = p_f^\mu - p_i^\mu = (E - m, \mathbf{p}) \equiv (T, \mathbf{p}) \end{aligned} \quad (12.31)$$

where we have explicitly shown the integration limits of z to remind us that the length the particle travels is infinite microscopically but finite macroscopically. Inserting these equations in Eq. (12.27), we obtain

$$T_{fi} = \frac{2\pi}{v} \delta\left(p_z - \frac{T}{v}\right) \frac{ze^2}{q^2} v_\mu \bar{u}(p_f) \gamma^\mu u(p_i) \quad (12.32)$$

The scattering probability is given by squaring T_{fi} , multiplying by the final state density, dividing by the initial state density $2E_i$ and taking the average and sum of the electron spin state:

$$\begin{aligned} dw_{fi} &= \frac{1}{2E_i} \left(\frac{1}{2} \sum_{\text{spin}} |T_{fi}|^2 \right) \frac{d^3p_f}{(2\pi)^3 2E_f} \\ &= \frac{z^2 e^4 L}{16\pi E_i E_f v^2} \delta\left(p_z - \frac{T}{v}\right) \frac{1}{q^4} v_\mu v_\nu L^{\mu\nu} d p_T^2 d p_z \end{aligned} \quad (12.33)$$

Here we have used $|2\pi\delta(p_z - T/v)|^2 = 2\pi\delta(p_z - T/v) \int dz = 2\pi L\delta(p_z - T/v)$. $L^{\mu\nu}$ is given by Eq. (6.143). Using

$$\begin{aligned}
 L^{\mu\nu} &= \frac{1}{2} \sum_{r,s} [\bar{u}_s(p_f) \gamma^\mu u_r(p_i)] [\bar{u}_s(p_f) \gamma^\nu u_r(p_i)]^* \\
 &= 2 \left\{ p_i^\mu p_f^\nu + p_i^\nu p_f^\mu + g^{\mu\nu} \frac{q^2}{2} \right\} \\
 v_\mu p_i^\mu &= m, \quad v_\mu p_f^\mu = E_f - v p_{fz} = E - T = m \\
 v_\mu v_\nu g^{\mu\nu} &= 1 - v^2 = \frac{1}{\gamma^2}
 \end{aligned} \tag{12.34}$$

$v_\mu v_\nu L^{\mu\nu}$ is expressed as

$$v_\mu v_\nu L^{\mu\nu} = 2 \left\{ 2m^2 + \frac{q^2}{2\gamma^2} \right\} = 4m^2 \left(1 + \frac{q^2}{4m^2\gamma^2} \right) \tag{12.35}$$

Inserting Eq. (12.35) in Eq. (12.33), the ionization probability is expressed as

$$dw_{fi} = \frac{4\pi z^2 \alpha^2 L}{v^2} \frac{m^2}{E_i E_f} \left[\frac{1}{q^4} + \frac{1/4 m^2 \gamma^2}{q^2} \right] dp_T^2 \tag{12.36}$$

The energy loss per unit length is obtained by multiplying Eq. (12.36) by the energy T given to the electron and its number density n_e , dividing by L and integrating over p_T^2 :

$$\frac{dE}{dx} = \frac{4\pi z^2 \alpha^2 n_e}{m v^2} \left[m^2 \int \frac{T}{m+T} \left\{ \frac{1}{q^4} + \frac{1/4 m^2 \gamma^2}{q^2} \right\} dp_T^2 \right] \tag{12.37}$$

The momentum transfer can be expressed in terms of p_T and T using $p_z = T/v$:

$$-q^2 = |\mathbf{q}|^2 - \omega^2 = p_T^2 + \left(\frac{T}{v} \right)^2 - T^2 = p_T^2 + \left(\frac{T}{\beta\gamma} \right)^2 \tag{12.38a}$$

On the other hand

$$\begin{aligned}
 -q^2 &= -(E_f - E_i)^2 + (\mathbf{p}_f - \mathbf{p}_i)^2 = -(E_f - m)^2 + \mathbf{p}_f^2 \\
 &= 2m(E_f - m) = 2mT
 \end{aligned} \tag{12.38b}$$

which is an exact relation with the mass of the charged particle taken into account.

As the energy loss T given to the electron is mostly in the nonrelativistic region, we may approximate

$$E_f = m + T \simeq m \quad (12.39a)$$

$$Q^2 \equiv -q^2 \sim p_T^2 \quad (12.39b)$$

Then Eq. (12.37) can be approximated by

$$\frac{dE}{dx} = \frac{4\pi z^2 \alpha^2 n_e}{mv^2} \left[\frac{1}{2} \int_{Q_{\min}^2}^{Q_{\max}^2} \left\{ \frac{1}{Q^2} - \frac{1}{4m^2\gamma^2} \right\} dQ^2 \right] \quad (12.40)$$

The maximum value of $Q_{\max}^2 = 2mT_{\max}$ occurs when the charged particle hits head on and scatters the electron in the forward direction. The value of T is given by

$$T = \frac{2m\beta^2\gamma^2 \cos^2 \theta}{(m/M + \gamma)^2 - \beta^2\gamma^2 \cos^2 \theta} \quad (12.41)$$

where θ is the electron scattering angle and M the mass of the charged particle. From this $T_{\max} \simeq 2m\beta^2\gamma^2$. $Q_{\min}^2$ is $(T_{\min}/\beta\gamma)^2$ from Eq. (12.38a). T must be at least larger than the ionization energy I of the atom to be ionized, which means $Q_{\min}^2 = (I/\beta\gamma)^2$. Thus the expression for dE/dx is

$$\frac{dE}{dx} = \frac{2\pi z^2 \alpha^2 n_e}{mv^2} \left[\ln \frac{2m\beta^2\gamma^2 T_{\max}}{I^2} - \beta^2 \right] \quad (12.42)$$

which is referred to as the Bethe–Bloch formula [137].

For practical applications, it is more convenient to give the ionization loss per unit weight ($\xi = \rho x$). Inserting $\hbar$ and c to recover the right dimensions and taking into account the density effect ($\delta/2$, which will be described soon), the ionization loss per gram-centimeter according to the Particle Data Group [311] is given by

$$\frac{dE}{d\xi} = K \frac{Z}{A} \left(\frac{z}{\beta} \right)^2 \left[\frac{1}{2} \ln \left(\frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} \right) - \beta^2 - \frac{\delta}{2} \right] \quad (12.43)$$

$d\xi$	$= \rho dx$	ρ : density of matter
K	$= 4\pi N_A r_e^2 m_e c^2$	$= 0.3071 \text{ MeV cm}^2/\text{g}$
N_A	$= 6.022 \times 10^{23} \text{ g}^{-1}$	$= \text{Avogadro number}$
r_e	$= e^2/(4\pi\epsilon_0 m_e c^2)$	$\simeq 2.82 \times 10^{-13} \text{ cm} = \text{classical electron radius}$
Z, A		$= \text{atomic number and atomic mass of matter}$
I	$\simeq 16Z^{0.9} \text{ eV}$	$= \text{average ionization potential of matter}$
δ		$= \text{density effect}$

From the formula, some characteristics of the ionization loss can be read off. The ionization loss per unit weight $dE/d\xi$ is

1. proportional to Z/A and oblivious of matter species (weak dependence on $\ln I$) and
2. a function of the velocity of the incoming particle.
3. At low $\beta\gamma \lesssim 1$, the curve has $1/\beta^2$ dependence, while at large $\beta\gamma \gg 1$ has a logarithmic rise ($\sim \ln \gamma^2$).

Figure 12.10 shows the energy loss of the positively charged muon in copper (Cu). The Bethe–Bloch formula is relevant in the region $\beta\gamma = 0.1$ to 1000. At higher energies, radiative losses due to bremsstrahlung and pair creation become important.

For hadrons (π , p , etc.), strong interactions become dominant at high energies before the radiative processes come in. As can be seen in Fig. 12.10, the ionization rises slowly ($\sim \ln \gamma$) and for a wide range of $\beta\gamma$ the ionization loss is more or less constant. This is referred to in the literature as minimum ionization loss. It is convenient to remember the fact that

Minimum Ionization Loss: The energy loss per gram of relativistic particles is about 1.8 MeV and barely depends on the energy.

The ionization loss is a statistical phenomenon. The measured data fluctuates around the value given by Eq. (12.43). The distribution is Gaussian, but with thin materials the statistical occurrence of ionization just once or twice with a big kick to the electron is sizeable, producing an asymmetric distribution with a long tail on the high-energy side, which is called the Landau distribution (Fig. 12.11a).

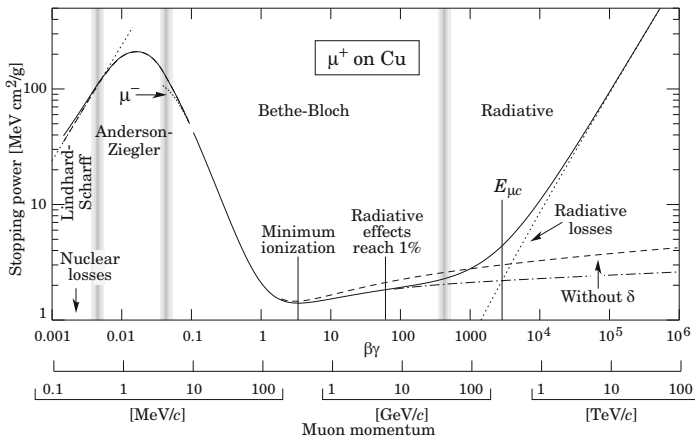


Figure 12.10 dE/dx from positive muons in copper as a function of $\beta\gamma = p/Mc$. Figure taken from [311].

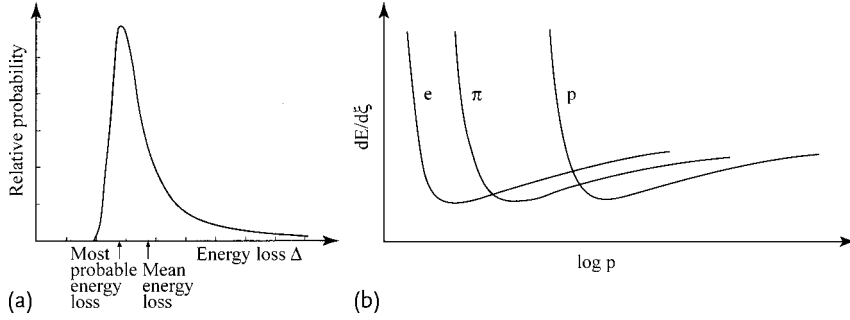


Figure 12.11 (a) The Landau distribution, which is a typical energy loss in a thin absorber. Note that it is asymmetric with a long high-energy tail. Figure from [311]. (b)

$dE/d\xi$ curves as a function of momentum $p = M\beta\gamma$. As $dE/d\xi$ is a function of $\beta\gamma$ only, the curves separate depending on the particle's mass.

In practical applications, information on momentum ($p = M\beta\gamma$) rather than velocity is easily obtained. Then the ionization curve is different for different masses (Fig. 12.11b). Thus combined knowledge of p and $dE/d\xi$ can be used to identify particle species.

Density Effect The logarithmic rise $\ln \gamma^2$ in the ionization loss is due to the relativistic effect of contraction of the distance from the passing particle to the atoms being ionized (or transverse electric field $E_{\perp} \propto \gamma$). However, if many atoms exist in this region, the screening effect due to dielectric polarization attenuates the effect, which is referred to as the density effect.

Let us consider how the ionization formula is affected. In matter, the photon satisfies the dispersion relation

$$\omega^2 = \frac{|q|^2 c^2}{\epsilon} = \frac{|q|^2 c^2}{n^2} \quad (12.44)$$

where ω is the wave frequency, q the wave number, ϵ (normalized as $\epsilon_0 = 1$) the dielectric constant and n the refractive index. Denoting the phase velocity of light in matter by $c_m (= c/\sqrt{\epsilon})$, the transverse wave number by q_T and using $q_z = \omega/v$, we have

$$q_T = \frac{\omega}{v} \sqrt{\epsilon \frac{v^2}{c^2} - 1} \quad (12.45)$$

For $\epsilon > 1$, a real photon (Cherenkov light) is emitted if $v > c_m$ (described later in Sect. 12.4.4).

Figure 12.12 shows argon's dielectric constant as a function of photon energy. The region of interest is the absorption region ($10 \text{ eV} \lesssim \hbar\omega \lesssim 10 \text{ keV}$), where the

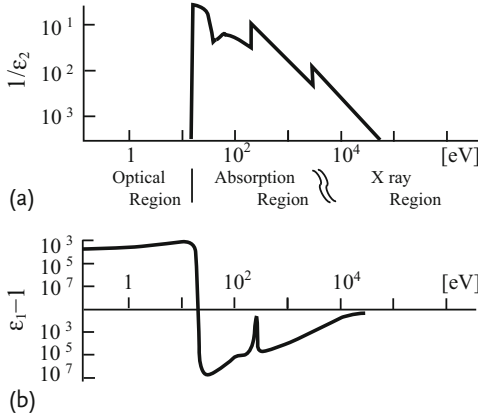


Figure 12.12 Dielectric constant ($\epsilon = \epsilon_1 + i\epsilon_2$) of argon gas at 1 atmosphere. (a) ϵ_2 is displayed as photon absorption length in meters ($\propto 1/\epsilon_2$). (b) $\epsilon_1 - 1$. Abscissa is photon energy in eV. In the optical region, matter

is transparent with $\epsilon_2 = 0$, $\epsilon_1 > 1$. This is the region in which Cherenkov light can be emitted. The absorption region has an effect on the ionization loss. The transition radiation occurs in the X-ray region. Figure from [14].

dielectric constant is expressed as

$$\begin{aligned}\epsilon &= \epsilon_1 + i\epsilon_2 \\ \epsilon_1 &\simeq 1 - \frac{\omega_p^2}{\omega^2} < 1 \\ \hbar\omega_p &= \hbar \sqrt{\frac{e^2 n_e}{\epsilon_0 m_e}} \simeq \left\{ 2 \frac{Z}{A} \rho [\text{g/cm}^3] \right\}^{1/2} \times 21 \text{ eV}\end{aligned}\quad (12.46)$$

where n_e is the electron number density and ω_p is the plasma frequency. In the absorption region, the light is partially absorbed and the dielectric constant is complex. In this region, q_T is imaginary and the transverse electric field decreases exponentially. Writing $q_T = i\kappa$,

$$\begin{aligned}\kappa &= \frac{q}{\beta' \gamma'} = \frac{\omega}{v} \sqrt{1 - \epsilon \left(\frac{v}{c} \right)^2} \simeq \frac{\omega}{c} \sqrt{\frac{1}{(\beta \gamma)^2} + \frac{\omega_p^2}{\omega^2}} \\ &\xrightarrow{\beta \gamma \rightarrow \infty} \frac{\omega_p}{c} \\ \beta' &= n\beta = \frac{v}{c_m}, \quad \gamma' = \frac{1}{\sqrt{1 - \beta'^2}}\end{aligned}\quad (12.47)$$

where c_m is the velocity of light in matter. The above formula tells us that when saturation due to the density is effective, $Q_{\min}$ is not given by $I/(c\beta\gamma)$ but by $\hbar\omega_p/c$. Formally, we only need to replace e by e/n and c by $c_m = c/n$ in the formula to give the ionization loss in matter. As a result of this the second term in the last equality of Eq. (12.38a) is replaced by κ^2 . In the language of quantized field theory, the photon has acquired mass due to the polarization effect of matter. From the

relation $(\omega/c)^2 = q^2/n^2 \simeq q^2(1 - \delta n^2) = q^2 + (\omega_p/c)^2$, the acquired mass is given by $\hbar\omega_p/c^2$.

Let us estimate at what value of $\beta\gamma$ the density effect becomes sizeable. The maximum transverse distance $\gamma_{\max}$ at which the charged particle can have influence without considering the density effect is given roughly by

$$\gamma_{\max} \simeq \frac{\hbar}{p_{T\max}} = \frac{\hbar c \beta \gamma}{I} \simeq \left(\frac{20 \text{ eV}}{I} \right) \beta \gamma \times 10^{-6} \text{ cm} \quad (12.48)$$

The size 10^{-6} cm is of the order of the average intermolecular distance of gas at one atmosphere. On the other hand, the maximum distance of the density effect at maximum saturation is given by

$$\gamma_{m,\max} = \frac{c}{\omega_p} \simeq \frac{1}{\sqrt{\rho}} \times 10^{-6} \text{ cm} \quad (12.49)$$

Comparing the two, the value of $\beta\gamma$ to reach saturation is given by $\beta\gamma \sim 1/\sqrt{\rho}$, which is ~ 100 for gas. But for a liquid or solid saturation is reached at small values of $\beta\gamma$. When the density effect is sizeable, the ionization loss is smaller by $\delta/2$

$$\frac{\delta}{2} = \ln \frac{\gamma_{\max}}{\gamma_{m,\max}} = \ln \frac{\hbar\omega_p}{I} + \ln \beta\gamma \quad (12.50)$$

Comparing this equation with the original formula, $dE/d\xi$ increases only as $\sim \ln \beta\gamma$ in contrast to the original $\sim \ln(\beta\gamma)^2$ (see the curve in Fig. 12.10 with and without the density effect).

Range When a charged particle enters a thick layer of matter, it loses all its energy by ionization loss and stops in the matter. The distance it travels is called the flight distance or range. Let E be the original energy of the particle and R the range, then

$$R = \int^R dx = \int^E \frac{dE}{dE/dx} \quad (12.51)$$

As $E/Mc^2 = \gamma = 1/\sqrt{1-\beta^2}$ and dE/dx are both functions of β only, R/M is also a function of β only. Referring to Eq. (12.43), the expression for R can be given as

$$R = \frac{M}{z^2} G(\beta) \quad (12.52)$$

Namely, the range is proportional to the mass of the incoming particle, and inversely proportional to the charge squared. The concept of the range is useful only when energy loss other than ionization loss can be neglected. This is the case for muons with energy less than a few hundred GeV and low-energy hadrons ($R < \lambda$, where λ is the nuclear interaction length or roughly those with $E - M \lesssim$ a few hundred MeV). Energetic hadrons ($E - M \gtrsim 1 \text{ GeV}$) usually interact strongly with matter before losing all their energy and induce shower phenomena by cascade interactions. The calorimeter utilizes the shower phenomenon to measure the hadronic energy.

12.4.3

Multiple Scattering

In the consideration of the ionization loss, recoils the traversing particle receives from electrons were totally neglected. However, scattering of a particle by nuclei is nonnegligible because they are heavy. Because the probability of Rutherford scattering is inversely proportional to $\sin^4 \theta/2$ (see Sect. 6.3.1), most of the scattering occurs at small angles. Therefore, when a particle traverses a thick material, the effect of Rutherford scattering is a superposition of statistically independent small-angle scatterings. Namely

$$\langle \theta^2 \rangle = \sum \theta_i^2 \quad (12.53)$$

which shows that the average scattering angle increases as the particle proceeds. Since it is a statistical phenomenon, the distribution is Gaussian. Namely, the angular distribution is given by

$$\frac{1}{2\pi} \exp\left(-\frac{\theta^2}{2\theta_0^2}\right) d\Omega \quad (12.54)$$

Let us calculate the average scattering angle using the Mott formula given in Sect. 6.7. Since the angle is small, the spin effect can be neglected. Using $|q|^2 = 2p^2(1 - \cos \theta) \simeq p^2\theta^2$,

$$d\sigma = \frac{4Z^2\alpha^2 E^2}{q^4} \left(1 - v^2 \sin^2 \frac{\theta}{2}\right) d\Omega \simeq \frac{4\pi Z^2 \alpha^2}{v^2} \frac{dq^2}{q^4} \quad (12.55)$$

The average scattering angle per collision can be found from

$$\begin{aligned} \int \theta^2 d\sigma &= \frac{4\pi Z^2 \alpha^2}{v^2 p^2} \int \frac{dq^2}{q^2} = \frac{8\pi Z^2 \alpha^2}{v^2 p^2} \ln \frac{q_{\min}}{q_{\max}} \\ &= \frac{8\pi Z^2 \alpha^2}{v^2 p^2} \ln \frac{r_{\max}}{r_{\min}} \end{aligned} \quad (12.56)$$

divided by $\int d\sigma$. Here, $r_{\max} = 1/q_{\min}$ is the distance at which the Coulomb field of the nuclei is screened by electron clouds. We take $r_{\max}$ as the Thomas–Fermi atomic radius ($r_{\text{atm}} = 1.4a_B Z^{-1/3} = 7.40 \times 10^{-9} Z^{-1/3}$), where $a_B = \hbar/\alpha m_e c$ is the Bohr radius. For $r_{\min}$, we adopt the nuclear radius $r_A = A^{1/3} \hbar/m_\pi c$. When a particle traverses a distance X cm, it is scattered N ($N = X \int d\sigma \rho N_A/A$) times on average. Then recovering $\hbar$ and c , the average scattering angle θ_0 is given by

$$\langle \theta_0^2 \rangle = N \langle \theta_i^2 \rangle = X \frac{16\pi Z^2 (\alpha \hbar c)^2 \rho N_A/A}{v^2 p^2} \ln \left(\frac{r_{\max}}{r_{\min}} \right)^{1/2} \equiv \frac{E_s^2}{v^2 p^2} \frac{X}{X_0} \quad (12.57a)$$

$$E_s = m_e c^2 \sqrt{\frac{4\pi}{\alpha}} = 21.2 \text{ MeV} \quad (12.57b)$$

$$X_0^{-1} = 4\alpha r_e^2 \frac{N_A Z^2 \rho}{A} \ln \frac{183}{Z^{1/3}} \quad (12.57c)$$

E_s is referred to as the scale energy and X_0 as the radiation length. It is customary to express the radiation length in g/cm^2 , in which case $X_0 \rightarrow X'_0 = \rho X_0$ is to be understood, and we follow the custom in the following. According to more detailed calculation or fits to the data [138, 311], the average scattering angle is given by

$$\theta_0 \equiv \sqrt{\langle \theta_0^2 \rangle} = z_{\text{in}} \left(\frac{E_s}{v p} \right) \sqrt{\frac{X}{X_0}} \left[1 + 0.038 \ln \frac{X}{X_0} \right]^{1/2} \quad (12.58a)$$

$$X_0 = \left[4\alpha r_e^2 \frac{N_A Z(Z+1)}{A} \ln \frac{287}{Z^{1/2}} \right]^{-1} = \frac{(716.4 \text{ g cm}^{-2}) A}{Z(Z+1) \ln(287/\sqrt{Z})} \quad (12.58b)$$

The accuracy of measuring the spatial position of a particle track is considerably affected by multiple scattering, especially at low energies. In fact, in deciding the particle momentum using the curvature of a charged track in the magnet, the limit of accuracy very often comes from the indefiniteness due to multiple scattering.

12.4.4

Cherenkov and Transition Radiation

A charged particle flying at constant speed does not radiate in vacuum, but can emit photons in matter. Physically, this can be interpreted as dipole radiation of polarized matter that the moving charge has created. If the particle velocity is smaller than the velocity of light in the matter $v < c/n$, the polarization is symmetric and no light is emitted because of destructive interference (Fig. 12.13a).

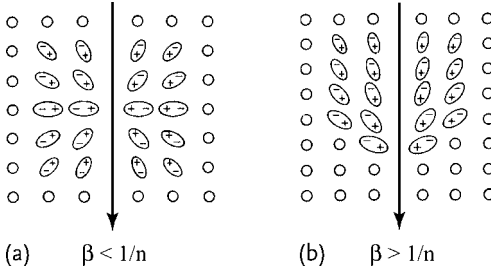


Figure 12.13 Polarization of a medium as a charged particle traverses it. (a) When the particle velocity v is smaller than the light velocity in the matter c/n , the polarization is symmetric and no light is emitted. (b) When $v > c/n$,

the polarization is asymmetric and light is emitted (Cherenkov light). If the medium is not uniform, the polarization becomes asymmetric and light is emitted even if $v < c/n$ (transition radiation).

If $v > c/n$, the polarization is asymmetric and constructive interference results in light emission (Cherenkov light). If the medium is not uniform, the polarization distribution is asymmetric and light is emitted even if $v < c/n$.

11) $\theta = \sqrt{\theta_x^2 + \theta_y^2}$ is a two-dimensional angle. If a one-dimensional projected θ_x is desired, use $E_s \rightarrow E_s/\sqrt{2} = 13.6 \text{ MeV}$.

The transition amplitude for radiation by a particle is given by

$$i T_{fi} = \langle f | -i \int H_{\text{int}} d^4x | i \rangle = -i e \int d^4x \langle f | A_\mu(x) j_m^\mu(x) | i \rangle \quad (12.59)$$

$j_m^\mu(x)$ is a classical current created by the incident charged particle and is given by Eq. (12.30). Treating photons as a quantized field

$$\langle f | A_\mu(x) | i \rangle = \varepsilon_\mu^*(\lambda) e^{ikx} \quad (12.60a)$$

Integration over x gives

$$j_m^\mu(x) \rightarrow \frac{v^\mu}{v} j_\omega \quad (12.60b)$$

$$j_\omega = \int \delta(x) \delta(y) \delta(z - vt) e^{ik \cdot x} d^4x = \int_{-L/2}^{L/2} dz e^{i(\omega/v - k \cos \theta)z}$$

where L is the length of material that the particle goes through. The transition amplitude becomes

$$dw_{fi} = \sum_{\lambda=\pm} |T_{fi}|^2 \frac{d^3k}{(2\pi)^3 2\omega} = \sum_{\lambda=\pm} e^2 \left| \frac{\varepsilon_\mu v^\mu}{v} \right|^2 |j_\omega|^2 \frac{d^3k}{(2\pi)^3 2\omega} \quad (12.61)$$

The polarization sum is given by

$$\sum_{\lambda=\pm} \left| \varepsilon_\mu \frac{v^\mu}{v} \right|^2 = \left(\delta_{ij} - \frac{k_i k_j}{|k|^2} \right) \frac{v_i v_j}{v^2} = \sin^2 \theta^2 \quad (12.62)$$

The above equations apply to vacuum. The electromagnetic field (photon) in matter can be obtained by replacing

$$e \rightarrow \frac{e}{\sqrt{\varepsilon}}, \quad k \rightarrow nk \quad (\text{or } c \rightarrow c_m = c/\sqrt{\varepsilon}), \quad \varepsilon = n^2 \quad (12.63)$$

Inserting Eqs. (12.60), (12.62) and (12.63) in Eq. (12.61), the number of emitted photons per unit frequency and solid angle $d\Omega$ is expressed as follows (recovering physical dimensions with $\hbar$ and c):

$$\frac{d^2 N}{d\omega d\Omega} = \frac{\alpha \omega n}{4\pi^2 c^2} \left| \int_{-L/2}^{L/2} dz (\hat{k} \times \hat{v}) e^{iz/Z_F} \right|^2 \quad (12.64a)$$

$$\hat{k} = \frac{\mathbf{k}}{|\mathbf{k}|}, \quad \hat{v} = \frac{\mathbf{v}}{|\mathbf{v}|}$$

$$Z_F = \left[\frac{\omega}{v} - nk \cos \theta \right]^{-1} \quad (12.64b)$$

Z_F is referred to as the formation zone.

Cherenkov Radiation If the medium is uniform, the integration gives a δ -function and nonzero value only when the phase vanishes, namely

$$\cos \theta_c = \frac{1}{n\beta} = \frac{1}{v/c_m}, \quad c_m = \frac{c}{n} \quad (12.65)$$

Therefore, only when the particle velocity exceeds that of light in the matter does the traveling particle emit photons, which are called Cherenkov light. Carrying out the integration, using $|2\pi\delta(\dots)|^2 = 2\pi\delta(\dots) \int dz = 2\pi\delta(\dots)L$ and dividing by L , we find the number of emitted photons per unit length and per unit energy interval ($dE = d\hbar\omega$) is

$$\frac{d^2 N}{dE dx} = \frac{\alpha}{\hbar c} \left(1 - \frac{1}{n^2 \beta^2}\right) = \frac{\alpha}{\hbar c} \sin^2 \theta_c \approx 370 \sin^2 \theta_c \text{ eV}^{-1} \text{ cm}^{-1} \quad (12.66)$$

Note, the index of refraction is a function of the photon energy ($E = \hbar\omega$).

Transition Radiation If the velocity of the charged particle does not exceed the velocity of light in the material, the phase of Eq. (12.64) does not vanish and no radiation is emitted in a uniform medium. However, if there exists nonuniformity in the medium, for instance, if there is a thin dielectric foil in vacuum with thickness L at position $-L/2 \leq z \leq L/2$ (see Fig. 12.14), they give a finite contribution to the integral and the photon is emitted. This is called transition radiation. As is clear from the above arguments, transition radiation and Cherenkov radiation have a common origin.

For this configuration, the number of photons per unit energy interval and per unit solid angle is obtained from Eq. (12.64) by integration. The integral is of the form

$$\int_{-L/2}^{L/2} f(n_1) + \left(\int_{-\infty}^{-L/2} + \int_{L/2}^{\infty} \right) f(n_2) = \int_{-L/2}^{L/2} \{f(n_1) - f(n_2)\} + \int_{-\infty}^{\infty} f(n_2) \quad (12.67)$$

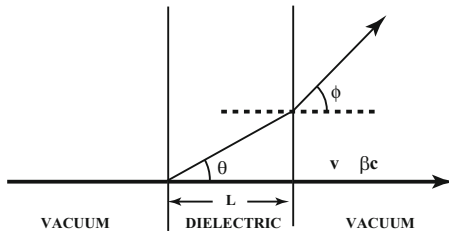


Figure 12.14 Emission of transition radiation in a thin slab.

where $n_2 = 1$ is the index of refraction in vacuum. The second term gives a δ -function and vanishes in vacuum. Therefore,

$$\begin{aligned} \frac{d^2 N}{dE d\Omega} &= \frac{\alpha \omega}{4\pi^2 c^2} n \sin^2 \theta \left[\frac{\sin[(L/2)(\omega/v - n_1 k \cos \theta)]}{\omega/v - n_1 k \cos \theta} - (n_1 \rightarrow n_2) \right]^2 \\ &= \frac{\alpha}{4\pi^2 \omega} n \sin^2 \theta \left[\frac{\sin[(\omega L/2c)(1/\beta - n_1 \cos \theta)]}{1/\beta - n_1 \cos \theta} - (n_1 \rightarrow n_2) \right]^2 \end{aligned} \quad (12.68)$$

In the region $|z| \gg Z_F$, the exponential function oscillates violently and does not contribute, but the integral over the formation zone ($|z| \lesssim Z_F$) can give a finite contribution. This means the foil thickness L must be larger than the size of the formation zone, which in practical situations described later is tens of micrometers. There are no constraints on the index of refraction n and transition radiation occurs for a wide range of ω . For ω in the optical region (see Fig. 12.12), it gives rise to optical transition radiation at wide angles at modest velocities, that is for β below the Cherenkov threshold. However, we are interested in applications for very relativistic particles with, say, $\gamma \gtrsim 1000$. In the X-ray region, the index of refraction is smaller than and very close to 1. It can be written as $n = \sqrt{\epsilon} \simeq 1 - \omega_p^2/2\omega^2$ [see Eq. (12.46)]. The intensity of emitted photons is high only where the denominator in Eq. (12.68) nearly vanishes, which occurs at very small angles of $\theta \simeq 1/\gamma$. Then using $\beta \simeq 1 - 1/2\gamma^2$, the denominator can be approximated by

$$\frac{1}{\beta} - n \cos \theta \simeq \frac{1}{2} \left(\frac{1}{\gamma^2} + \frac{\omega_p^2}{\omega^2} + \theta^2 \right) \quad (12.69)$$

where we have put $n_1 = 1 - \omega_p^2/2\omega^2$ and $n_2 = 1$. The emission rate becomes

$$\begin{aligned} \frac{d^2 N}{dE d\Omega} &\simeq \frac{\alpha}{\pi^2 \omega} \phi^2 \sin^2 \left[\frac{\omega L}{2c} \left(\frac{1}{\gamma^2} + \frac{\omega_p^2}{\omega^2} + \phi^2 \right) \right] \\ &\times \left[\frac{1}{1/\gamma^2 + \omega_p^2/\omega^2 + \phi^2} - \frac{1}{1/\gamma^2 + \phi^2} \right]^2 \end{aligned} \quad (12.70)$$

The two numerators taken approximately equal were factored out and we have used

$$n \sin^2 \theta^2 \simeq (n \sin \theta)^2 = \sin \phi^2 \simeq \phi^2 \quad (12.71)$$

where ϕ is the emission angle in vacuum (see Fig. 12.14). Equation (12.70) represents the interference between the two equal and opposite amplitudes with relative phase given by the argument of the $\sin^2$ factor. The interference can be important, but if there is no interference, the flux is

$$\frac{d^2 N}{dE d\Omega} \simeq \frac{\alpha}{\pi^2 \omega} \phi^2 \left[\frac{1}{1/\gamma^2 + \omega_p^2/\omega^2 + \phi^2} - \frac{1}{1/\gamma^2 + \phi^2} \right]^2 \quad (12.72)$$

Integration over solid angle $d\Omega = \pi d\phi^2$ gives

$$\begin{aligned} \frac{dI}{dE} &= \hbar\omega \frac{dN}{dE} = \frac{\alpha}{\pi} \left[\left\{ 1 + 2 \left(\frac{\omega}{\gamma\omega_p} \right)^2 \right\} \ln \left\{ 1 + \left(\frac{\gamma\omega_p}{\omega} \right)^2 \right\} - 2 \right] \\ &= \frac{\alpha}{\pi} \begin{cases} 2 \ln(\gamma\omega_p/\omega) & \omega/\gamma\omega_p \ll 1 \\ \frac{1}{6} (\gamma\omega_p/\omega)^4 & \omega/\gamma\omega_p \gg 1 \end{cases} \end{aligned} \quad (12.73)$$

The spectrum is logarithmically divergent at low energies and decreases rapidly for $\omega > \gamma\omega_p$. Integration of Eq. (12.73), however, is finite and gives a total energy flux $\alpha\gamma\hbar\omega_p/3$ and a typical photon energy of order $\gamma\hbar\omega_p/4$ [216, 311], where $\hbar\omega_p \approx 20$ eV. Therefore the radiated photons are in the soft X-ray range 2–20 keV. Since the emitted energy is proportional to γ , the transition radiation is inherently suitable for measurements of very energetic particles. However, only $\sim \alpha$ photons are emitted with one thin foil. The yield can be increased by using a stack of foils with gaps between them.

12.4.5

Interactions of Electrons and Photons with Matter

Bremsstrahlung The interactions of electrons and photons differ considerably from those of other particles and need special treatments. The reason is as follows. An energetic charged particle loses its energy by bremsstrahlung under the influence of the strong electric field made by the nuclei. According to classical electrodynamics, the loss is proportional to $(2/3)\alpha|\ddot{\mathbf{r}}|^2$, where $\ddot{\mathbf{r}}$ is the acceleration the particle receives. Since $|\ddot{\mathbf{r}}| \propto Z/m$, the loss rate is proportional to $(Z/m)^2$, which is large for light particles, especially electrons.

The bremsstrahlung cross section is given by the Bethe–Heitler formula Eq. (7.71). The cross section for emitting a photon per photon energy interval is obtained by integrating over the electron final state and photon solid angle. The calculation is tedious. Besides, target nuclei are not isolated but surrounded by electrons, which effectively screen the nuclear electric field. Therefore, we only quote a result in the relativistic limit with the screening effect taken into account [61, 203]:

$$\begin{aligned} d\sigma &= 4Z^2\alpha r_e^2 \frac{E_f}{E_i} \frac{d\omega}{\omega} \left[\left(\frac{E_i^2 + E_f^2}{E_i E_f} - \frac{2}{3} \right) \ln \frac{183}{Z^{1/3}} + \frac{1}{9} \right] \\ \text{for } \frac{E_i E_f}{m_e \omega} &\gg \alpha^{-1} Z^{-1/3} \end{aligned} \quad (12.74)$$

where $E_i, E_f = E_i - \omega$, ω are the energies of the initial and final electron and photon, respectively, $r_e = \alpha\hbar/(m_e c)$ being the classical electron radius. If we neglect

Table 12.1 Radiation length and critical energy

Material	X_0 [g]	E_c [MeV]
H ₂ (liquid)	69.3	382
C	44.4	102
Al	24.3	47
Fe	13.8	23.9
Pb	5.83	6.9

the second term in the square brackets of Eq. (12.74), which is on the order of a few percent, the equation can be rewritten as

$$\frac{d\sigma}{d\omega} = \frac{A}{X_0 N_A} \frac{1}{\omega} \left(\frac{4}{3} - \frac{4}{3} \gamma + \gamma^2 \right) \quad (12.75a)$$

$$\gamma = \frac{\omega}{E_i} \quad (12.75b)$$

where the radiation length X_0 is given by Eq. (12.57c).

A characteristic of bremsstrahlung is that the number of emitted photons is proportional to $1/\omega$ and hence the energy spectrum is nearly flat. The energy loss is obtained by

$$-\frac{dE}{dx} = n_A \int_0^{\hbar\omega_{\max}} \omega \frac{d\sigma}{d\omega} d\omega \quad (12.76)$$

where n_A is the number density of atoms. The result is an easy to remember formula:

$$-\frac{dE}{dx} = \frac{E}{X_0}, \quad \rightarrow \quad E = E_0 e^{-x/X_0} \quad (12.77)$$

Actually X_0 is defined as the length where the energy loss by bremsstrahlung is equal to that by ionization. Figure 12.15 shows Eq. (12.77) is a good approximation for energies higher than a few hundred MeV. As $1/X_0 \propto Z^2$, $(dE/d\xi)_{\text{brems}} \propto Z^2$ and $(dE/d\xi)_{\text{ionization}} \propto Z$,

$$\frac{(dE/d\xi)_{\text{brems}}}{(dE/d\xi)_{\text{ionization}}} \simeq \frac{E[\text{MeV}]}{800/(Z + 1.2)} \equiv \frac{E}{E_c} \quad (12.78)$$

$E_c = 800 [\text{MeV}]/(Z + 1.2)$ is called the critical energy.¹²⁾ At $E > E_c$ the energy loss due to bremsstrahlung is larger than that due to ionization and at $E < E_c$ the converse is true. The critical energy is an important parameter in designing a calorimeter. Table 12.1 lists radiation lengths and critical energies of a few examples.

12) There are other definitions of E_c and refined approximations. See [311].

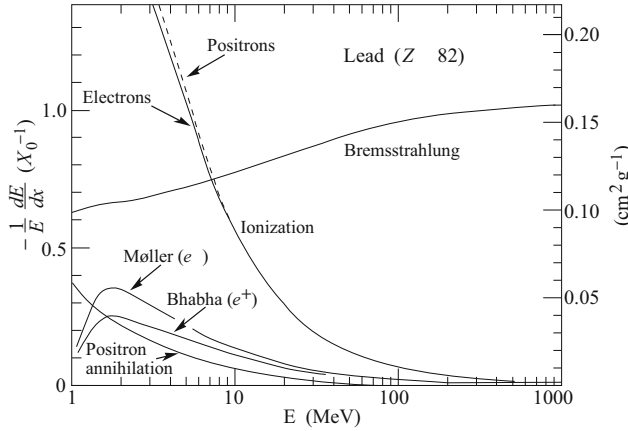


Figure 12.15 Radiation loss of the electron. Figure taken from [311].

Photon Absorption Photons lose energy through three processes: the photoelectric effect, Compton scattering and pair creation. For $E_\gamma \gg$ a few MeV, pair creation is the dominant process.

Pair creation is closely related to bremsstrahlung. It can be considered as bremsstrahlung by an electron with negative energy and that the electron itself is transferred to a positive energy state. The cross section is almost identical to that of bremsstrahlung and hence shares many common features. Like electrons, the photon intensity (energy $\times$ number per unit area per unit time) as it passes through matter is approximated by

$$\begin{aligned} dI &= -\mu I dx \\ \therefore I &= I_0 e^{-\mu x} \equiv I_0 e^{-x/\lambda} \end{aligned} \quad (12.79)$$

μ is called the mass absorption coefficient if x is expressed in g/cm^2 . The (mass) attenuation length $\lambda = 1/\mu$ of the photon for various materials is given in Fig. 12.16. At high energies ($E_\gamma \geq 100$ MeV), they become constant and Eq. (12.79) is a good approximation in the high-energy region. The attenuation length is conceptually analogous to the radiation length. Above $E_\gamma > 1$ GeV

$$\lambda = \frac{9}{7} X_0 \quad (12.80)$$

holds to within a few percent.

Cascade Shower

Longitudinal Profile A very high energy electron loses its energy by bremsstrahlung but the radiated photon produces a pair of electrons, which produce photons again.

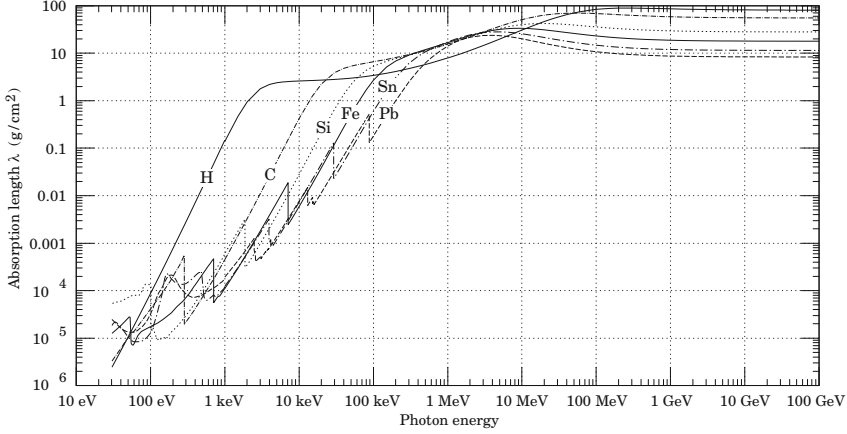


Figure 12.16 Photon attenuation length. Figure taken from [311].

Thus a great number of electrons and photons are created in a cascade by an energetic electron or photon. The multiplication continues until the electron energy reaches the critical energy, after which the ionization effect takes over. If the material length is measured in units of radiation length and the ionization and Compton effect are neglected (Rossi's approximation A [336]), the shower profile is independent of material. If only the Compton effect is neglected (approximation B) and if the energy is measured in units of critical energy, the result is also independent of material.

A Simple Model Basic characteristics of the cascade shower can be understood by considering a very simple model [203]. Let us assume that the electron as well as the γ -ray lose half their energy as they pass a radiation length X_0 . The electron emits a γ -ray and the γ -ray creates a positron–electron pair. Then at the depth of $2X_0$, the incident electron that had E_0 is converted to two electrons, one positron and one photon, each having energy $E_0/4$. Proceeding this way, at the depth t in units of X_0 , the numbers of electrons and photons are given by

$$n_e = \frac{2^{t+1} + (-1)^t}{3} \simeq \frac{2}{3}2^t \quad (12.81a)$$

$$n_\gamma = \frac{2^t - (-1)^t}{3} \simeq \frac{1}{3}2^t \quad (12.81b)$$

and the energy of each electron is given by $E = E_0/2^t$. Assuming the shower stops at $E = E_c$, then $E_c = E_0/2^{t_{\max}}$. Therefore

$$t_{\max} = \frac{\ln(E_0/E_c)}{\ln 2} \quad (12.82a)$$

$$N_{\max} = \frac{2}{3}2^{t_{\max}} = \frac{2}{3} \frac{E_0}{E_c} \quad (12.82b)$$

$$\begin{aligned}
 N(> E) &= \int^{t(E)} dt N = \frac{2}{3} \int^{t(E)} dt \exp(t \ln 2) \\
 &\simeq \frac{2}{3} \frac{e^{t(E) \ln 2}}{\ln 2} = \frac{2}{3} \frac{E_0}{E \cdot \ln 2} \simeq \frac{E_0}{E}
 \end{aligned} \tag{12.82c}$$

$$\therefore \frac{dN}{dE} = -\frac{2}{3} \frac{E_0}{E^2} \cdot \frac{1}{E^2} \tag{12.82d}$$

As there are two electrons for each photon, the length per track is $(2/3) X_0$, which gives the total length traveled by the charged particles

$$L = \frac{2}{3} \int^{t_{\max}} dt N = \left(\frac{2}{3}\right)^2 \int dt e^{t \ln 2} \simeq \frac{4}{9 \ln 2} 2^{t_{\max}} \simeq \frac{E_0}{E_c} \tag{12.83}$$

As the ionization loss is approximately constant regardless of the particle's energy, the total energy of the incident particle transferred to the cascade shower is given by $(dE/dx) \times L \propto E_0$. The depth at shower maximum is $t_{\max} \simeq \ln(E_0/E_c)$. This means the length of the shower increases logarithmically with the energy. The shower stops abruptly at the shower maximum owing to the simplicity of the model whereas the real shower decreases exponentially. However, basic characteristics of the model are confirmed qualitatively by both Monte Carlo calculations and data.

Analytical formulae are useful for understanding the characteristics of the shower, but with the advent of high speed computers Monte Carlo programs have come in handy for simulating showers. Figure 12.17 shows a longitudinal profile of the shower produced by a 30 GeV electron. The longitudinal shape is well reproduced

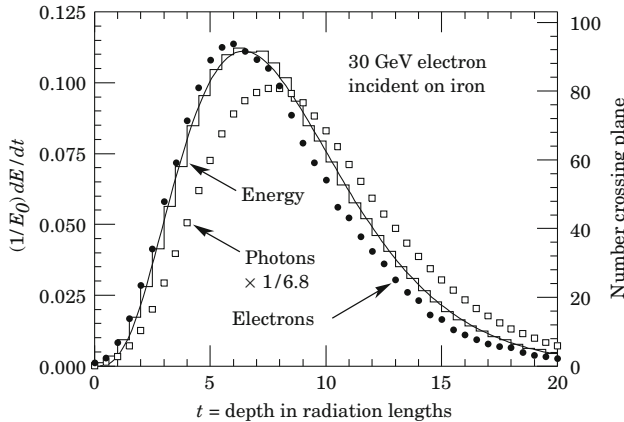


Figure 12.17 A profile of a 30 GeV electron shower generated by the EGS4 simulation program. The histogram shows fractional energy deposition per radiation length and the

curve is a gamma-function fit to the distribution. Circles and squares indicate the number of electrons and photons with total energy greater than 1.5 MeV. Figure taken from [311]

by a gamma function of the form

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)} \quad (12.84a)$$

$$t = x/X_0, \quad t_{\max} = \frac{a-1}{b} = \ln(E/E_c + C_i) \quad (12.84b)$$

$$b \simeq 0.5, \quad {}^{13)} \quad C_e = -0.5, \quad C_\gamma = 0.5 \quad (12.84c)$$

where $t_{\max}$ is the thickness at shower maximum and a is calculated from Eq. (12.84b).

Transverse Profile So far we have ignored transverse expansion of the shower. In actual situations, although both bremsstrahlung and pair creation are well characterized by the radiation length, Coulomb scattering becomes significant below the critical energy. A characteristic length representing the transverse expansion is given by the Molière unit

$$R_M \equiv \frac{E_s}{E_c} X_0 \simeq \frac{7A}{Z} [g/cm^2] \quad (12.85)$$

where E_s is the scale energy defined in Eq. (12.57b). It is known that 90% of the total energy is contained within a radius $r = R_M$ and 99% within $3.5 R_M$. A detector to measure the total energy of an incident particle by inducing showers is called a calorimeter.

LPM Effect In bremsstrahlung, both the scattering angle of electrons and the emission angle of photons are small. This means the kick given to the electron is almost transverse and the longitudinal momentum transfer is very small:

$$k_{\parallel} = p_i - p_f - \omega = \sqrt{E_i^2 - m_e^2} - \sqrt{E_f^2 - m_e^2} - \omega \simeq \frac{m_e^2 \omega}{2E_i E_f} \simeq \frac{\omega}{2\gamma^2} \quad (12.86)$$

Then, due to the Heisenberg uncertainty principle, the region of photon emission is spatially extended $\sim 1/k_{\parallel}$. This is reintroduction of the formation zone (Z_F) that we described in the section on transition radiation. The length of the formation zone is given by $\sim 2\gamma^2/\omega$, which is on the order of $10 \mu\text{m}$ for 100 MeV photons emitted from 25 GeV electrons. If the condition of the electron changes before it goes out of the zone, interference with the emitted photon occurs. For instance, if the electron receives a sizeable multiscattering due to many scattering centers inside the zone and if the net scattering angle becomes comparable to that of the emitted photon, the emission is attenuated. This is referred to as the LPM (Landau–Pomeranchuk–Migdal) effect [234, 250, 275]. When the emitted photon energy is small (i.e., has long wavelength), the density effect (see the paragraph “Density Effect” in Sect. 12.4.2) has to be taken into account, too. Here, due to dielectric effect of matter, the wave number becomes $k \rightarrow nk$, which induces a phaseshift $(n-1)kZ_F$ and hence destructive interference.

13) Exact numbers depend on material. See for instance, Figure 27.19 of [311].

In order to quantify the above argument, we note that the amplitude of photon emission is proportional to the formation zone Z_F . Using the approximations $1 - \beta \simeq 1/2\gamma^2$, $1 - n \simeq \omega_p^2/\omega^2$ and $1 - \cos \theta \simeq \theta^2/2$, Eq. (12.64b) can be written as

$$Z_F = \frac{2c\gamma^2}{\omega} \frac{1}{1 + \gamma^2\theta^2 + \gamma^2\omega_p^2/\omega^2} \quad (12.87)$$

Equation (12.87) shows that the attenuation of the photon emission occurs when the second and third terms in the denominator become nonnegligible compared to the first term. The LPM attenuation results from the second term and the density attenuation from the third term. Defining ω_{LPM} by the condition that the multiple scattering angle becomes equal to the emission angle

$$\gamma^2 \langle \theta_0^2 \rangle |_{X=Z_F(\gamma^2\theta^2=1)} = 1 \rightarrow \omega_{\text{LPM}} = \frac{E_e^2}{E_{\text{LPM}}} \quad (12.88a)$$

$$E_{\text{LPM}} [\text{TeV}] = \frac{(m_e c^2)^4 X_0}{\hbar c E_e^2} = 7.7 X_0 [\text{cm}] \quad (12.88b)$$

From the above formula we see that the photon emission is suppressed when

$$\hbar\omega < \hbar\omega_{\text{LPM}} \quad \text{or} \quad \frac{\hbar\omega}{E_e} < \frac{E_e}{E_{\text{LPM}}} \quad (12.89)$$

As E_{LPM} is given as 4.3 TeV in lead and 151 TeV in carbon [234], this is primarily an extremely high energy phenomenon. The quantitative estimation of the suppression is not easy, because the number of terms in the bracket of Eq. (12.68) increases. Qualitatively speaking, the ω dependence on the number of emitted photons shifts from $1/\omega$ to $1/\sqrt{\omega}$ at $\omega \ll E_e$. Figure 12.18a shows a comparison of the energy loss per radiation length of uranium by a 25 GeV electron with the standard Bethe–Heitler formula [26].

We may say that the ultrahigh energy electron behaves like a pion, because emission of photons and electrons is suppressed and the effective radiation length becomes larger.

The density attenuation comes in when the phase shift of the emitted photon becomes of the order of 1:

$$(n-1)kZ_F \simeq \gamma^2\omega_p^2/\omega^2 \simeq 1 \quad (12.90)$$

i.e. $\hbar\omega \sim \gamma\hbar\omega_p \sim \gamma\{\rho [\text{g/cm}^3]\}^{1/2} \times 21 \text{ eV}$

The suppression effect is due to the third term in the denominator of Eq. (12.87). For the case of electrons with 25 GeV energy ($\gamma = 5 \times 10^4$), the density effect comes in at $\omega \sim 5 \text{ MeV}$ in lead and $\sim 1 \text{ MeV}$ in carbon, which are both smaller than the LPM effect. The LPM effect exists also in pair creation by photons.

Although the LPM effect was pointed out long ago theoretically, it has been considered only in treating ultrahigh energy cosmic ray reactions [224, 278] and a quantitative treatment was difficult. One experiment was carried out using an electron

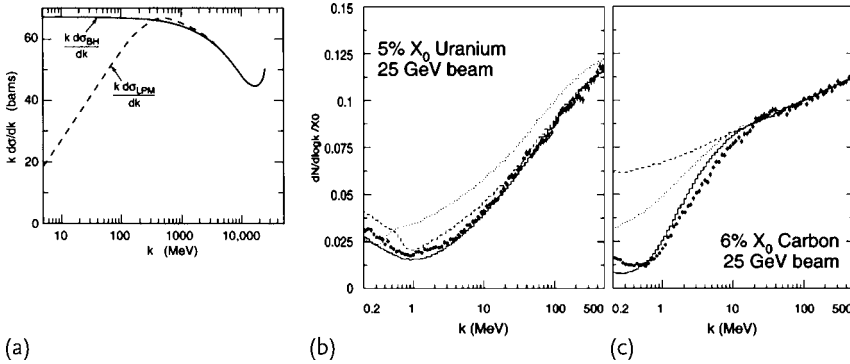


Figure 12.18 LPM and dielectric effect in bremsstrahlung [25, 26]. (a) Calculated energy spectrum of photons emitted by 25 GeV electrons incident on uranium. The LPM curve (dashed line) is compared with the Bethe-Heitler (BH) formula (solid line). Experimental data are the emitted photon spectra from 25 GeV electrons incident on

5% X_0 uranium (b) and 6% X_0 carbon (c). The solid line is a Monte Carlo calculation with LPM + DE (density effect) + TR (transition radiation). The dotted line (LPM + TR1) and the dashed line (BH + TR2) use different TR formulae. The deviation of BH from a flat line is due to multiple scattering.

beam at SLAC [26]. Figure 12.18 shows the result. Both the LPM and density effects are observed. The rise of the spectrum at the low energy end is due to transition radiation.

The LPM effect is sizeable only at extremely high energies. However, it may become a problem when designing a calorimeter for the future LHC or ILC accelerators in the tens of TeV region.

12.4.6

Hadronic Shower¹⁴⁾

Energetic hadrons incident on thick material also make a shower. If we scale by nuclear absorption length instead of radiation length, arguments from the electromagnetic shower are generally valid, too. Namely, the total length of all the tracks is proportional to the incident energy and the size of the shower grows as $\ln E$, etc. There are some characteristics peculiar to hadronic showers. We show typical shower configurations of the hadronic and electromagnetic showers side by side (Fig. 12.19) to demonstrate their similarity as well as their difference.

i) Absorption Length Usually the absorption length λ is much longer than the radiation length: $\lambda_{\text{Fe}} = 131.9$, $\lambda_{\text{Pb}} = 194$, $X_{0\text{Fe}} = 13.84$, $X_{0\text{Pb}} = 6.37$ (all in units of g/cm^2). The hadronic shower grows slowly compared with the electromagnetic shower. This fact can be used to separate them.

ii) e/h Compensation In nuclear reactions, a sizeable fraction of the energy disappears. Some fraction is used to break nuclei (binding energy) and some is carried

14) [135, 152, 336, 346].



Figure 12.19 Monte Carlo simulation of the different development of hadronic and electromagnetic cascades in the earth's atmosphere, induced by 100 GeV protons (a) and photons (b). [Z (vertical) axis range: 0–30.1 km, X/Y axis range: ± 5 km around shower core.] Image by [345].

away by neutrinos, which are the decay product of pions. Besides, slow neutrons, which are produced in plenty at nuclear breakup, are hard to observe. Consequently, the energy resolution of hadronic calorimeters is worse than that of electromagnetic calorimeters. Furthermore, the secondary particles are mostly pions, of which $\sim 1/3$ are neutral and promptly decay to two photons. They produce electromagnetic showers, which are localized relative to hadronic showers. If all the energy is measured this is not a problem, but a practical calorimeter is usually of sampling type and has a structure consisting of layers of absorbers with active detectors sandwiched between them. The electromagnetic part of the shower may not be sampled or, if it is, the measured energy is much higher because of its local nature. Both effects enhance the statistical fluctuation, especially that of the first reaction. The hadronic calorimeter has an intrinsic difference in the detection efficiency for the electronic and hadronic components and its ratio is dubbed the e/h ratio. The performance can be improved by enhancing the detection efficiency of the neutron and deliberately decreasing that of the electric component. As uranium captures slow neutrons with subsequent photon emission, the efficiency can be improved if it is used as the absorber material [134, 389]. However, as it is a nuclear fuel and highly radioactive, a conventional choice is to use a material with high hydrogen content (for example, organic scintillators) to enhance the detection efficiency for slow neutrons and to optimize the ratio of the absorber thickness to that of the detector material to adjust the e/h ratio at the expense of reduced energy resolution.

Balancing the responses of electron and hadron components improves linearity of the energy response and is effective in measuring the so-called missing energy.

12.5

Particle Detectors

12.5.1

Overview of Radioisotope Detectors

Particle detectors derive information from light or electrical signals that are produced by particles when they interact with material as they enter the detector. Historically the first detector was a fluorescent board, with which Röntgen discovered the X-ray. Rutherford used a ZnS scintillator to measure the α -ray scattering by a gold foil, which led to his celebrated atomic model. The measurement was a tedious observation by the human naked eye using a microscope. Later, ionization chambers, proportional counters and Geiger counters were introduced. Their signal is electrical. Then, around 1940, there was a major step forward with the invention of photomultipliers, which enabled amplification of fluorescent light combined with fast electronics. Another breakthrough was made by G. Charpak when he invented the multiwire proportional chamber (MWPC) in 1968, which laid the foundations of modern detector technology.

A Wilson cloud chamber was used to discover the positron and strange particles, emulsions to discover π mesons, and bubble chambers to discover a great number of resonances. They were of a type that covered the complete 4π solid angle. They were powerful in that they could reconstruct the entire topology of reactions, including many secondary particles, and played an important role historically. However, because of their long response and recovery times, they become obsolete as the need for rapid measurement and fast processing time increased.

Modern measuring devices use a combination of various detectors, including light-sensitive counters, gas chambers and radiation-hard solid-state detectors with computer processing, and achieve the functions of (1) complete-solid-angle coverage, (2) the ability to reconstruct the whole topology and (3) high-speed processing.

In the following, we first explain detectors classified according to their output signals, then describe their usage depending on what one wants to measure. Signals are of two kinds, photonic and electronic. Detectors with light outputs are scintillation counters, Cherenkov counters and transition-radiation counters. Detectors with electric signals are semiconductor detectors and various gaseous chambers. Roughly speaking, light signals are faster than electric signals by at least an order of magnitude, but the latter excel in precision.

12.5.2

Detectors that Use Light**a) Scintillation Counters**

i) Organic Scintillators Fluorescent materials such as anthracene ($C_{14}H_{10}$), *p*-terphenyl ($C_{18}H_{14}$), tetraphenyl butadiene ($C_{28}H_{22}$) and their mixtures are used as light-emitting scintillators. They come as liquid scintillators, when dissolved in an organic solvent such as toluene (C_7H_8), or solid plastic scintillators, when mixed in plastic bases.

Charged particles ionize atoms in the scintillator as they pass through. Ionized electrons excite molecules to excited energy levels and when they return to the ground state they emit UV (ultraviolet) photons, which are generally outside the sensitive regions (400–700 nm; $1\text{ nm} = 10^{-9}\text{ m}$) of photomultipliers. Therefore, usually wavelength shifters (another type of fluorescent material) are mixed to generate photons of longer wavelength (Fig. 12.20). Commonly used organic scintillators produce light pulses with attenuation time $\sim 10^{-9}\text{ s}$ and are suitable for fast counting.

ii) Inorganic Scintillators (NaI, CsI, BiGe, BGO (bismuth germanate), etc.) Historically NaI has been in long use but others share a common light-emitting mechanism. A small amount of thallium is mixed in NaI (NaI-Tl), which become photoactive centers, capture ionized electrons and emit light on falling from the excited state to the ground state. One advantage of inorganic scintillators is a large light output ($\gtrsim 10$ times) compared to organic scintillators, which gives much better resolution (a few percent at 1 MeV). Another advantage is their short radiation length (2.6 cm with NaI, 1.12 cm with BGO), which is used to make a compact electromagnetic calorimeter. Disadvantages are that they are slow (attenuation time $\sim 10^{-6}\text{ s}$) and generally expensive.

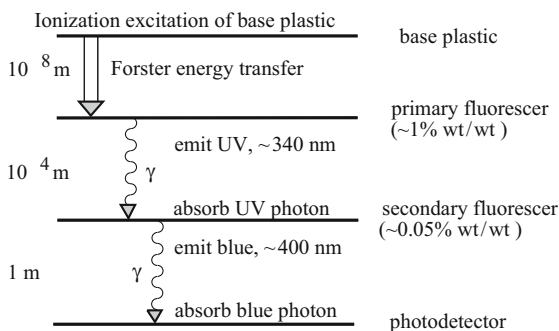


Figure 12.20 Scintillation ladder depicting the operating mechanism of plastic scintillators. Förster energy transfer is an energy transfer between two chromophores (part of molecules responsible for its color) in proxim-

ity ($\sim 10^{-8}\text{ m}$) without radiation. Approximate fluorscener concentrations and energy transfer distances for the separate subprocesses are shown [311].

iii) Structure of Plastic Scintillators Light from scintillators is guided to photon sensors (typically photomultipliers but photodiodes and other semiconductor devices are also used) through transparent plastic light guides. Figure 12.21 shows a few examples of the basic structure of plastic scintillators. Figure 12.21a is the most primitive; (b) is an improved (called adiabatic) type, which has better light collection efficiency. This was achieved by matching the areas of the light-emitting surface and receiving surface by using bent rectangular strips. Liouville's theorem of phase-space conservation works here. With the advent of collider detectors, a compactified multilayer structure without dead space looking into the interaction point has become important. Figure 12.21c uses optical fiber to achieve this at the expense of light collection efficiency and guides light to far away photomultipliers or other light-collecting devices.

A typical structure of a photomultiplier is shown in Fig. 12.22. When photons hit the photocathode, an electron is emitted with probability 10–20% or higher and hits the first dynode D_1 , emitting a few more electrons. They are attracted and accelerated by the higher electric potential of successive dynodes. The potential is supplied by the accompanying circuit shown at the top of the figure. This is repeated as many times as the number of dynodes and anode. For an incident electron, 3–4 electrons are emitted, thus multiplying the number of electrons in cascade. As they reach the anode, the number of electrons is multiplied to $(3 - 4)^{14} \simeq 10^8$, producing a strong electric pulse.

Plastic scintillation counters excel in time response and resolution and are used to detect the arrival time of particles and trigger other devices. They can also mea-

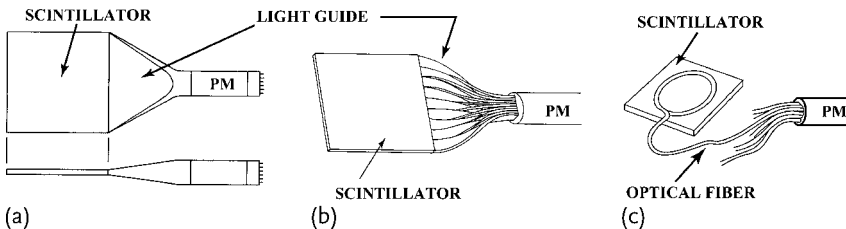


Figure 12.21 Scintillation counters with primitive (fishtail) (a) and improved (adiabatic) (b) light guides. (c) The use of optical fibers embedded in a groove as light guides is mostly

used for stacks of scintillators, typically in a calorimeter where compactness takes precedence over light collection efficiency. PM = photomultiplier.

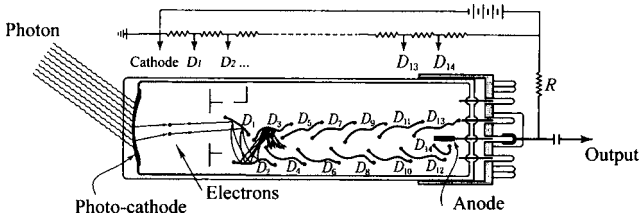


Figure 12.22 Schematic layout of a photomultiplier and voltage supply taken from [298].

sure the time interval that the particle takes to fly a certain length (TOF; time of flight) and determine the particle's velocity. Combined with the momentum information this can be used to determine the mass, i.e. particle species. A typical plastic scintillator has a density of 1 g/cm^3 . Therefore, when an energetic charged particle passes a scintillator of thickness 1 cm , it deposits $\sim 1.8 \text{ MeV}$ of energy in it. Assuming an average ionization energy of 30 eV , it creates $\sim 6 \times 10^4$ ion pairs and produces $\sim 10\,000$ photons. As the average energy of fluorescence is $\langle h\nu \rangle \sim 3 \text{ eV}$, about 1–2% of the energy is converted to photons. Assuming geometrical efficiency of 10% for the photons to reach the photomultiplier and assuming quantum efficiency $\sim 20\%$ for conversion to electrons, the number of emitted photoelectrons is about 200, which has statistical uncertainty of 10–20%, a typical resolution of plastic scintillators.

Problem 12.2

Assume that light signals with an attenuation time of 10 ns ($1 \text{ ns} = 10^{-9} \text{ s}$) are 10 ns long triangular pulses and are amplified by the photomultiplier given in Fig. 12.22. Assuming a gain (amplification factor) of 10^7 , $R = 50 \Omega$ ($RC \ll 10 \text{ ns}$, where C is stray capacitance of the circuit), what is the voltage of the pulse height obtained from 200 photoelectrons?

b) Cherenkov Counter

i) Basic concept The Cherenkov counter resembles the scintillation counter in that it extracts light from a transparent material. It differs in that it does not emit light if the particle velocity is below a threshold ($\beta = v/c = 1/n$; n = index of refraction), and also in that the emission is not isotropic but directional. The response time is also much faster, restricted only by the photomultiplier response. The emission angle θ with respect to the particle direction is determined by

$$\cos \theta = \frac{1}{n\beta} \quad (12.91)$$

The energy loss per unit length is given by Eq. (12.66). Denoting the frequency of the Cherenkov light as ν

$$\frac{dE}{dx} = \frac{2\pi z^2 \alpha}{c} \left(1 - \frac{1}{\beta^2 n^2} \right) h\nu d\nu \quad (12.92)$$

Here, ze is the charge of the incident particle. In terms of the wavelength it is proportional to $d\lambda/\lambda^3$ and the spectrum is dominated by photons in the violet to ultraviolet region. The number of photons per unit length is obtained by dividing the above formula by $h\nu$. Putting n approximately constant in the narrow region $(\nu_1 - \nu_2)$ where photomultipliers are sensitive, we obtain

$$\frac{dN}{dx} = \frac{2\pi z^2 \alpha}{c} (\nu_1 - \nu_2) \sin^2 \theta_c \simeq 500 z^2 \sin^2 \theta_c \text{ cm}^{-1} \quad (12.93)$$

As an example, putting $z = 1$, $\beta = 1$, $n = 1.33$ (water), $dN/dx \simeq 200 \text{ cm}^{-1}$. Further setting $\lambda = 600 \text{ nm}$, $dE/dx \simeq 400 \text{ eV/cm}$. As the ionization loss is $\sim 1.8 \text{ MeV/g}$, only 0.02% of the energy is used and the light emission efficiency is $\sim 1/100$ that of a scintillator.

ii) Threshold-Type Cherenkov Counter Counters of this type are used to identify particle species by utilizing the fact that light is emitted only when $\beta > 1/n$. For instance, when there is a mixed beam of electrons, muons and pions with fixed momentum, the velocities are $\beta_e > \beta_\mu > \beta_\pi$ because $m_e < m_\mu < m_\pi$. As the index of refraction of the gas changes in proportion to the pressure, two Cherenkov counters can be arranged so that one of them detects electrons only and the other electrons and muons but not pions. Then the two counters in tandem can separate three species of the particle.

iii) Differential-Type Cherenkov Counter Even if $\beta > 1/n$, the emission angle depends on the velocity. By picking up light emitted only in the angle interval from θ to $\theta + \Delta\theta$, particle identification is possible. Figure 12.23a shows a differential-type Cherenkov counter called DISC [265], which was used to identify particles in the hyperon beam (a 15 GeV secondary beam produced by a 24 GeV primary proton beam at CERN). The beam enters from the left and goes through a gas radiator. Emitted light is reflected back by a mirror, through a correction lens and slit, reflected through 90° by another mirror and collected by photomultipliers. Only light between θ and $\theta + \Delta\theta$ is detected by use of a slit. To achieve $\Delta\beta/\beta = 5 \times 10^{-5}$, a correction lens made of silica gel is used. Figure 12.23b shows a velocity curve of the Cherenkov counter. The chosen velocity is varied by changing the pressure of the nitrogen gas.

iv) RICH (Ring Image Cherenkov Counter): In this type of counter the emitted light is mapped on a plane to produce characteristic ring patterns. An example used as part of the collider detector complex at DELPHI [119] is shown in Fig. 12.24.

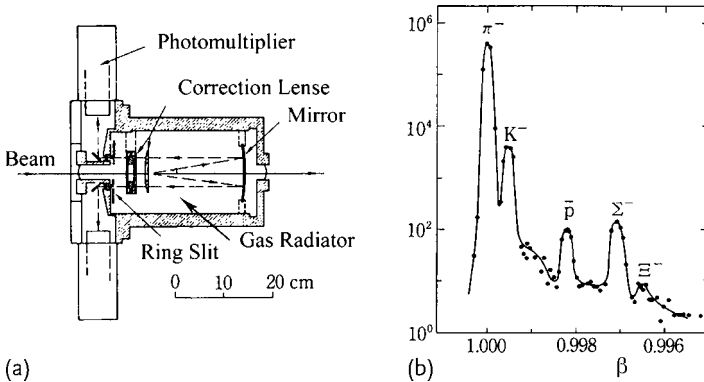


Figure 12.23 (a) A differential-type gas Cherenkov counter used to identify particles in the hyperon beam [265]. (b) A velocity curve of the Cherenkov counter to identify 15 GeV/c

particles produced by the 24 GeV primary proton beam. The velocity is varied by changing the pressure of the nitrogen gas.

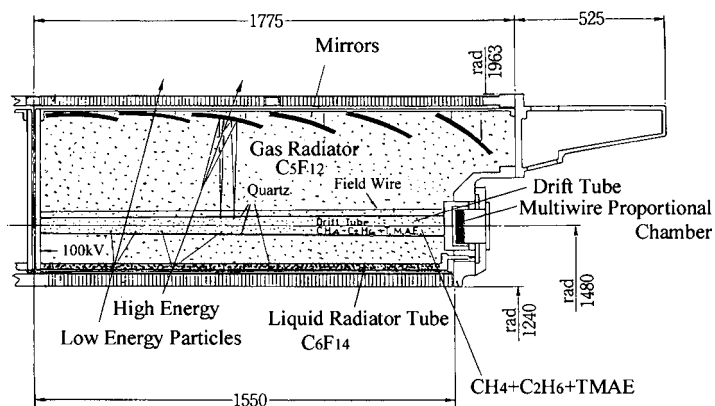


Figure 12.24 RICH counter. A part of DELPHI detector at LEP/CERN [119].

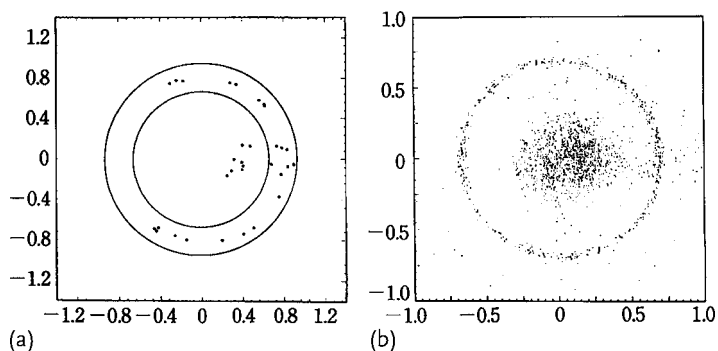


Figure 12.25 Cherenkov light detected by the RICH counter. (a,b) Cherenkov light from the liquid radiator projected on a plane perpendicular to the incident particles. (a) Photons emitted by a single particle. (b) Accumulated light from 246 cosmic ray muons.

Functionally, it is two counters in one. The front section uses the liquid radiator C_6F_{14} and the rear section uses the gas radiator C_5F_{12} . Cherenkov light from the gas radiator is reflected back by mirrors and detected by the same drift tube detector that detects the light from the liquid radiator. The drift tube (similar to the drift chamber to be described later) contains liquid $CH_4 + C_2H_6 + TMAE$, where TMAE is tetrakis(dimethylamino)ethylene. The signals are processed by computers. Figure 12.25 shows a pattern projected on a plane perpendicular to the incident particles.

Figure 12.25a is the response for a single particle and (b) is a collection of light signals obtained from 246 cosmic ray muons. The poor photon statistics of Cherenkov light is reflected in the small number of detected photons.

c) Transition Radiation Counter The transition radiation counter uses light emitted at the boundary of two materials with different dielectric constants. As described

before, the origin is dipole emission of polarized media induced by the passage of a particle. As the emitted photon energy is proportional to $\gamma = 1/\sqrt{1-\beta^2}$, it is most suitable for detection of extremely relativistic ($\gamma \gtrsim O(1000)$) particles. The differential emission spectrum is given by Eq. (12.72). The emission angle of the photon is concentrated in the forward region ($\theta \leq 1/\gamma$). The spectrum decreases rapidly above $\omega > \gamma \omega_p$ and the lower cutoff is determined by the detector condition. Taking the cutoff frequency $\omega_c = 0.15\gamma \omega_p$, the detected number of photons per layer is given roughly by

$$N_\gamma(\omega > 0.15\gamma \omega_p) \sim 0.5\alpha \quad (12.94)$$

As the number of photons emitted is extremely small, some several hundred layers with gaps between are necessary for practical use. In a practical design, the transition radiation counter consists of radiators and detectors. A dominant fraction of soft X-rays are absorbed by the radiator itself. To avoid absorption of emitted X-rays by the absorbers, they have to be thin. But if they are too thin the intensity is reduced by the interference effect of the formation zone ($Z_F \simeq$ a few tens of micrometers). Besides, for multiple boundary crossings interference leads to saturation at $\gamma_{\text{sat}} \simeq 0.6\omega_p \sqrt{\ell_1 \ell_2}/c$, where ℓ_1, ℓ_2 are the thickness of the radiators and the spacing, respectively [30, 155]. As the X-ray detectors are also sensitive to the parent particle's ionization, sufficient light to surpass it is necessary. A practical design has to balance the above factors. Radiators that have been used come in two kinds: a stack of thin foils or a distribution of small-diameter straw tubes. Figure 12.26 shows the latter type used in VENUS/TRISTAN [340]. The radiator consists of polypropylene straw tubes with 180 μm diameter placed in He gas. With cutoff at 11 keV, it achieved 80% efficiency for electrons with rejection factor greater than 10 against pions.

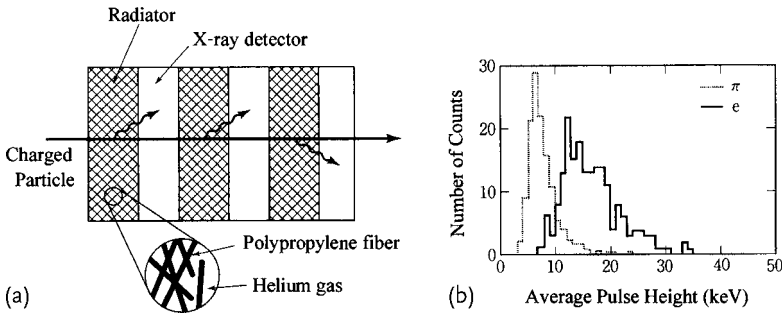


Figure 12.26 Transition radiation counter. (a) Schematic layout of one used in VENUS/TRISTAN. (b) Its radiation spectra [340, 341]. Note the pion's signal is due to ionization.

12.5.3

Detectors that Use Electric Signals**a) Gaseous Detectors**

Basic Structure Electrons and ions can move at high speed in gas and when the proper potential is applied they can be extracted as electric signals. Three types of gaseous detectors – ionization, proportional, and Geiger tubes – were used first historically. They share the same structure but have different operating voltages (Fig. 12.27). A rare gas such as argon (Ar) or xenon (Xe) is confined in a cylindrical tube with a metal wire stretched at the central position. The reason for using rare gas is so that ionized electrons do not easily recombine with positive ions. The outer wall of the cylinder (cathode) is grounded and high voltage is applied to the central wire (anode). The strength of the electric field is proportional to $1/r$. When a charged particle passes through the tube, electrons and positive ions are made along the track. The number of ion pairs is proportional to the energy loss of the particle. If no voltage is applied, they recombine and restore electric neutrality, but if a voltage is applied, electrons are attracted to the anode and positive ions are attracted to the cathode, creating an electric current. However, as positive ions are much heavier they are slow in reaching the cathode and the signals they create are usually discarded.

Operation Mode When the applied voltage is low, exactly the same number of free electrons as are created by ionization are collected by the anode. Detectors operating in this region (ionization mode) are called ionization tubes (or chambers, depending on the shape of the detector). The electric signal is small and used only when a large enough number of electrons are created. With higher voltage, the

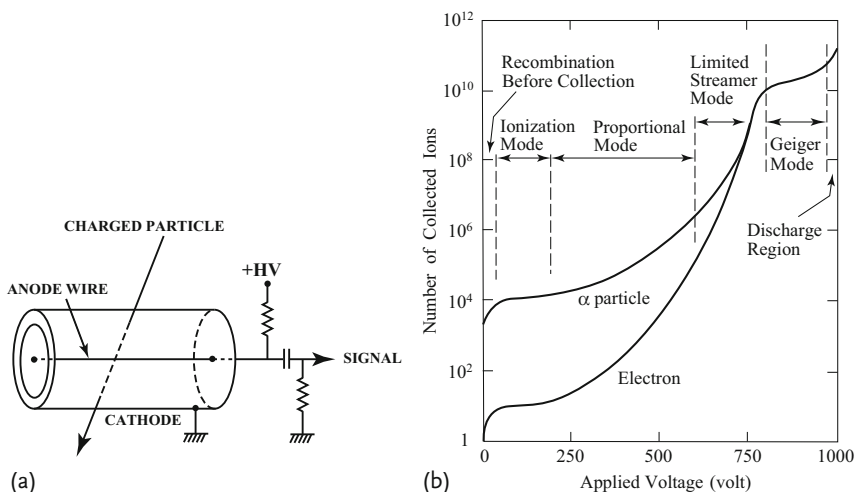


Figure 12.27 (a) Basic structure of a gaseous counter. (b) The number of collected ions in various detection modes depending on the applied voltage.

accelerated electrons knock out additional electrons from the gas molecules. By making the radius of the anode wire small, an electric field with large gradient is created near the anode, the secondary electrons can ionize the molecules and multiplication begins, leading eventually to a cascade. When the voltage is not too high, the number of electrons collected at the anode is proportional to the number produced by ionization (proportional mode). Detectors in this range (proportional tubes) can attain an amplification factor of 10^3 – 10^6 by varying the voltage.

With further increase of the applied voltage, the space charge effect due to the ion pairs distorts the electric field and the proportionality is lost (limited streamer mode). However, some designs take advantage of the high output, which eliminates the need for electronic amplifiers. The induced cascade is still local, but with even higher voltage it spreads all over the detector region (Geiger mode), giving a saturated output signal regardless of the input. The reason for the spreading of the cascades is that excited molecules emit light when going back to the original states and excite more molecules in the far distance. In this mode, a quenching gas to absorb light is needed to stop the discharge. The recovery time is long and it is not suitable for high-speed counting. Note that, depending on the inserted gas species, the region of limited proportionality may not appear. The relation between the voltage and the number of output electrons is depicted in Fig. 12.27b.

Ionization Chamber As the output is small in the ionization mode, it is used only for special purposes, such as a calorimeter, where all the energy is deposited in the detector. The use of liquid increases the density and enables the detector to contain showers in a small volume and also increase the number of ionized electrons per unit length. In liquids, molecules such as oxygen capture electrons and a small fraction of such impurities prevents electrons from reaching the anode. For this reason only rare gas liquids such as Ar and Xe are used. Even in this case, the concentration of oxygen impurities has to be less than a few ppm (parts per million = 10^{-6}). There are many advantages in using it in high-energy experiments, but the disadvantage is that the liquid state can be maintained only at liquid nitrogen temperature (77 K) and therefore necessitates a cooling vessel (cryostat), which makes a dead region. To avoid this, so-called warm liquids (organic liquids referred to as TMP and TME) are being developed.

MWPC (Multiwire Proportional Chamber) The MWPC, invented by Charpak [95] in 1968, made a big step forward in the high-speed capability of track detection. Capable of high speed and high accuracy, it has accelerated the progress of high-energy physics and has found widespread application as an X-ray imager in astronomy, crystallography, medicine, etc. For this contribution, Charpak was awarded the Nobel prize in physics in 1992. The basic structure is depicted in Fig. 12.28a. Many wires are placed at equal intervals between two cathode plates and high voltage is applied between the wires (anodes) and the cathodes to be operated in proportional mode. The electric field is shaped as in Fig. 12.28b and each wire acts as an independent proportional tube. When a particle traverses the plate perpendicularly, ionized electrons are collected by nearby wires, thus achieving position accuracy determined by the spacing of the anode wires.

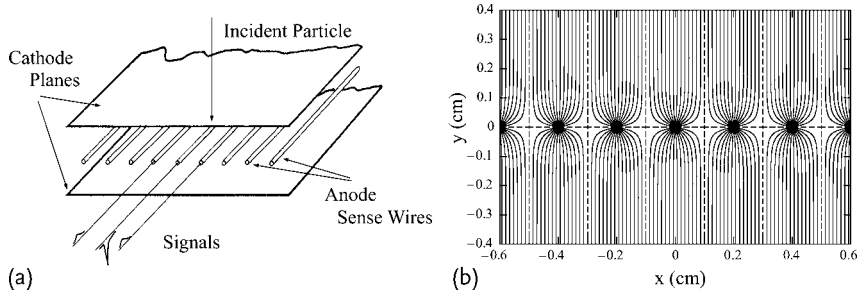


Figure 12.28 (a) A schematic of a MWPC and (b) lines of the electric field [312].

A typical example has wire spacing 0.5–2 mm, plate separation 1.5–2 cm, electric voltage 3–4 kV and radius of the anode wires 10–50 μm . The field is 10^4 – 10^5 V/cm and the amplification factor is $\sim 10^4$. A space resolution of 0.5 mm is obtained. For the gas in the chamber, Ar with 10% methane or so-called magic gas [Ar 75% + isobutane (C_4H_{10}) + Freon 13 (CF_3Br)] are commonly used.

Usually two MWPCs are used in combination to determine both x and y coordinates. Sometimes more planes with u or/and v coordinates (slanted relative to x, y) are used to eliminate multi-hit ambiguity. If strips of metal sheet instead of planes are used as cathodes, a single plane of the MWPC can obtain x and y simultaneously. The risetime of the MWPC is fast and the decay time is dictated by the ion mobility and the size of the chamber. If the wire spacing is a few millimeters, the decay time can be very fast (25–30 ns) and the MWPC can sometimes be used to obtain time information. Each wire can cope with a very high luminosity of $\sim 10^4 \text{ s}^{-1} \text{ mm}^{-1}$.

i) Drift Chamber (DC) The basic structure of a drift chamber is given in Fig. 12.29. Contrary to the proportional chamber, the field lines in the DC are horizontal between the anode and cathode wires. Besides, a number of field wires with graduated voltages are used to obtain a uniform electric field. Ions produced by incident particles drift towards the anode with constant velocity. The drift time together with the time information about the particle passage obtained separately determines the distance from the anode wire. The anode wires are in the proportional mode and the field distribution near the anode is similar to that in the MWPC. The uniform

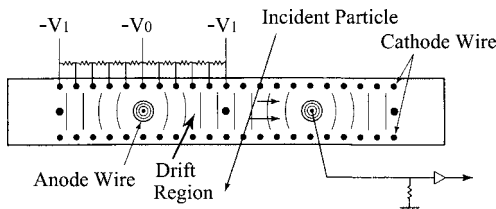


Figure 12.29 Basic structure of a drift chamber [77]. The field lines in the figure are horizontal between the anode and cathode wires. A number of field wires with graduated voltages are used to obtain a uniform electric field.

electric field has $\sim 1 \text{ kV/cm}$ extending 1–10 cm. As the drift velocity of the electrons is $4\text{--}5 \text{ cm}/\mu\text{s}$ and the accuracy of time measurement is $\sim 2 \text{ ns}$, it can give an accuracy of position determination $\sim 100 \mu\text{m}$. Drift chambers are often used in a magnetic field to determine the momentum of particles. In a magnetic field, electrons feel the Lorentz force and the voltage applied to the field wires is varied to compensate it. As the operational mode is the same as the MWPC, the response time and radiation hardness are similar. There are many variations of drift chambers, some of which will be introduced later.

ii) GEM (Gas Electron Multiplier) With the advent of semiconductor microchip technology, much smaller strips of conductive plates have come into realization. GEM is one such technology among many similar detectors (micromegas detector, etc., see [311]). Figure 12.30 shows the layout of the GEM detector. It has essentially the same structure as the MWPC, however, the anode intervals are much smaller. Multiplication of ions is carried out not by a cylindrical electric field but by a different potential across the hole and outgoing electrons are captured by nearby anodes.

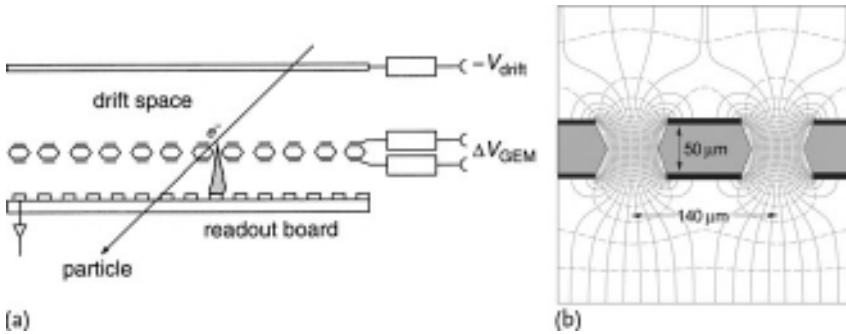


Figure 12.30 (a) Layout of a GEM detector. Ions in the drift space are attracted by the GEM, go through an amplification hole and are collected by nearby anodes. (b) Schematic view and typical dimensions of the hole structure in the GEM amplification cell. Electric field lines (solid) and equipotentials (dashed)

are shown. On application of a potential difference between the two metal layers, electrons released by ionization in the gas volume above the GEM are guided into the holes, where charge multiplication occurs in the high field [311].

b) SSDs (Solid State Detectors) When a small amount of an element with five valence electrons (P, As, Sb) is added to a semiconductor with four valence electrons such as silicon (Si) or germanium (Ge), excess electrons are easily separated to become free electrons (n-type semiconductors or n-SSD) and the SSD becomes slightly conductive. On the other hand if an element with three valence electrons (B, Al, Ga, In) is added, one electron is missing for each covalent bond and an electron is borrowed from somewhere else, resulting in excess positive freely moving holes (p-type semiconductor or p-SSD). The elements with five valence electrons are called donors and those with three, acceptors. Free electrons and holes

are called carriers. When p-SSD and n-SSD are in plane contact with each other (p-n junction), thermal motion causes electrons to diffuse into p-SSD and holes into n-SSD. The diffused electrons (holes) annihilate holes (electrons) on the other side, leaving both sides of the junction with ionized donors and acceptors. As these ions are bound in lattices, there remain distributed static charges, namely electric fields. The field works in the opposite direction to prevent conduction and eventually balances with the diffusing force. Carriers are absent in this region, which is called the depletion layer. The voltage across the depletion region is called the diffusion voltage. The outside region is neutral and the potential is nearly flat.

When one applies reverse bias, namely, positive (negative) voltage on the n (p) side of the junction (Fig. 12.31a), electrons and holes are attracted and extend the depletion region. No current flows. If, on the other hand, forward bias is applied (Fig. 12.31b), the diffusion potential is reduced, causing electrons on the n side and holes on the p side to diffuse beyond the depletion region, and a current flows.

If a charged particle passes through the depletion region, ionized electrons and ions along the track are attracted by the diffusion voltage and a pulse current flows. Ion pairs outside the depletion region diffuse slowly, giving a slow component that is mostly ignored. The p-n junction can be considered as a condenser with electrodes placed at the two edge of the depletion layer. Let C be its capacitance, C_A the effective capacitance of an amplifier connected to the junction to read out the electric signal, and Q the amount of all the charges collected. Then the magnitude of the pulse is given by $Q/(C + C_A)$. A characteristic of a semiconductor is that the energy levels of the impurity are very close to the conduction band, and it takes on average only 3.0 eV (Ge) or 3.7 eV (Si) to make pairs of ions. We have already stated that in ordinary matter it takes about 30 eV to ionize atoms. This means that the number of created ions in the SSD is ten times and statistical accuracy is greatly enhanced. An ionization loss of 1.8 MeV can produce $1.8 \times 10^6 / 3.7 \simeq 5 \times 10^5$ ions with statistical error of $1/\sqrt{N} \simeq 0.2\%$. If we consider the additional geometrical loss of

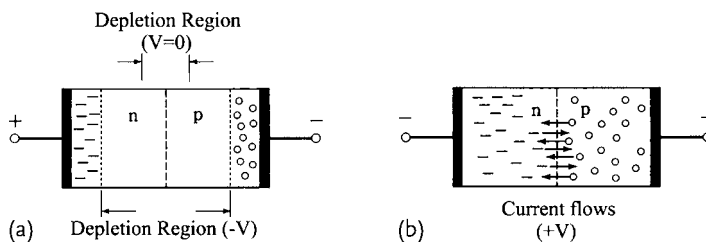


Figure 12.31 A p-n junction in reverse bias (a) and in forward bias (b). When there is no bias voltage across the junction, electrons from the n side and holes from the p side diffuse across the junction and vanish after recombination, creating a positively charged region on the n side and a negatively charged region on the p side. As a result, a region with no

conducting electrons and holes, referred to as the depletion region, develops. With reverse bias (a) the depletion region expands with little mobility. With forward bias (b), both electrons and holes are attracted and absorbed by the electrodes, which in turn supply more of them, and the current flows.

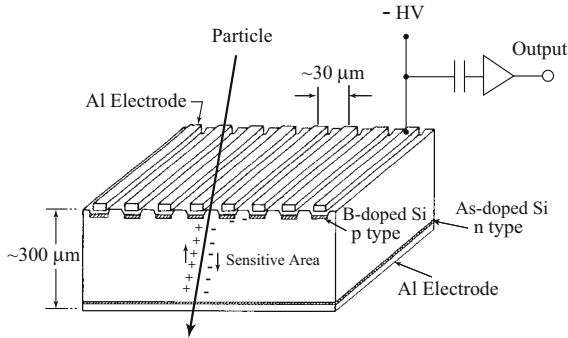


Figure 12.32 Schematic layout of a silicon microstrip detector. The strips are capacitively coupled.

scintillation counters, the resolution is at least $\times 10$ that of NaI and approximately $\times 100$ that of plastic scintillation counters.

If a SSD is made in the form of arrays of strips with a width of say $30\ \mu\text{m}$ using methods taken from the manufacture of integrated circuits (Fig. 12.32), it can function as a high precision position detector. Applications of SSDs as high-quality detectors are plentiful. The disadvantage is that it is difficult to manufacture large detectors and they are prone to radiation damage.

12.5.4

Functional Usage of Detectors

The information one wants to obtain using particle detectors is classified as follows.

(1) Time of particle passage: Detectors are used to separate events that occur within a short time interval or to determine the particle's velocity by measuring it at two different locations. The most widely used detectors are scintillation counters with time resolution $\Delta t \simeq 100$ to $200\ \mu\text{s}$ ($1\ \text{ps} = 10^{-12}\ \text{s}$).

(2) Position of particles: Detectors are used to reconstruct tracks, the direction of momentum and patterns of reactions. MWPCs (multiwire proportional chambers with position resolution $\Delta x \sim 1\ \text{mm}$), DCs (drift chambers with $\Delta x \simeq 100$ to $200\ \mu\text{m}$) and SSDs (solid state detectors with $\Delta x \simeq 10$ to $100\ \mu\text{m}$) are used. For some special experiments where detection of fast decaying particles is at stake emulsions are used to obtain $\Delta x \simeq 0.1$ to $0.5\ \mu\text{m}$. When the speed matters, hodoscopes with reduced position resolution in which strips of scintillation counters are arranged to make a plane come in handy. They are used for triggering (selecting desired events from frequently occurring reactions and initiating data taking) or for crude selection of events.

(3) Energy-momentum ($E, \mathbf{p}$) is obtained by dE/dx counters (which measure the ionization loss by a thin material), range counters or calorimeters (which measure the total energy deposited in a thick material). To obtain the momentum one measures the curvature of a particle's circular motion in a magnetic field.

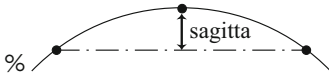
(4) Particle identification (PID) is carried out by measuring the mass of the particle. With combined knowledge of the momentum p and velocity v , the mass value can be extracted from $m = p/v\gamma$. A pattern of particle reactions contains much valuable information. A long track without interaction may be identified as a μ , an electromagnetic shower as γ or e (depending on the existence of a pre-track), a V-shaped decay pattern with long (typically a few cm $\sim$ 1 m) missing tracks as K^0 , Λ and with (typically 10 $\mu\text{m} \sim$ 1 mm) short missing tracks as τ leptons or B , C particles. In a many-body system, knowing the energies and momenta of particles can give an invariant mass $((mc^2)^2 = (\sum E_i)^2 - (\sum \mathbf{p}_i)^2)$ of the whole system.

In the following we list detectors classified by their function.

Momentum Measurement A particle with momentum p makes a circular motion in a magnetic field, and from its curvature the momentum can be determined [see Eq. (12.11)]:

$$p_T[\text{GeV}/c] = 0.3 B R [\text{T m}] \quad (12.95)$$

Note, however, that the measured momentum is a component (p_T) transverse to the magnetic field. The total momentum can be calculated if the angle with respect to the magnetic field is known. Let us assume that a track-measuring device is placed in the magnetic field and n points of a track with total projected length L within the magnet are measured. The momentum can be determined by measuring a quantity called the sagitta, which is defined as the distance from the middle point of the arc to a line made by two end points. For a not so large angle where $\theta \simeq \sin \theta$ holds

$$s \simeq \frac{L^2}{8R} = \frac{R}{8} \left(\frac{0.3 B L}{p_T} \right)^2 \quad \text{--- sagitta} \quad (12.96)$$


From the formula, one sees that the momentum resolution improves as the magnetic field and path length become larger and as the resolution of space measurement improves. The momentum resolution is given by

$$\frac{\Delta p_T}{p_T} = \frac{\Delta R}{R} = \frac{\Delta s}{s} = \Delta s \frac{8R}{L^2} = \frac{8}{0.3} \frac{p_T}{B L^2} \Delta s \quad (12.97)$$

One needs to measure at least three points to determine the sagitta and the resolution improves as the number of measured points increases. If we denote the resolution of a point measurement as Δx , the relation and the momentum resolution are given by

$$\frac{\Delta p_T}{p_T} = \begin{cases} \sqrt{\frac{3}{2}} \frac{8}{0.3} \frac{p_T \Delta x}{B L^2} & N = 3 \\ \frac{p_T \Delta x}{B L^2} \frac{1}{0.3} \sqrt{\frac{720}{N+4}} & N \gtrsim 10^{15} \end{cases} \quad (12.98)$$

In addition to this measurement error, one has to consider the error due to multiple scattering (see Sect. 12.4.3).

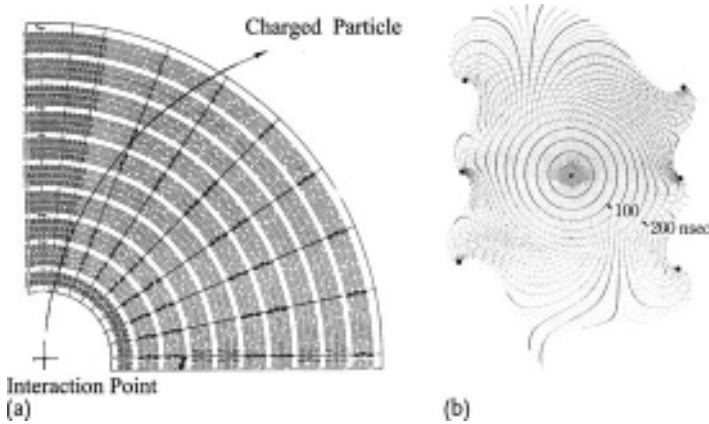


Figure 12.33 An example of momentum measurement by a drift chamber in the magnet. (a) Cross section perpendicular to beam axis of the Central Drift Chamber (CDC) in the VENUS detector [27]. (b) Electric field lines and equipotentials in a unit cell. CDC has 10 layers, each of which, except the innermost layer, is a combination of three planes of drift

chambers staggered by half a cell to each other. Time information from one cell alone suffers from left-right ambiguity but information from two staggered planes can determine the vector of the particle motion. Information from the third plane is used to obtain the z coordinate along the axis.

Figure 12.33 shows a cylindrical drift chamber to determine momenta of secondary particles built in the VENUS/TRISTAN detector [27]. The magnetic field is along the beam direction and the wires are also strung in the same direction. A unit cell is made of one anode wire and eight field wires. The equipotentials of the electric field are circular and the position accuracy does not depend on the angle of incidence of the particle.

Energy Measurement If a particle loses energy only by ionization, the range (Sect. 12.4.2) can determine the total kinetic energy of the particle. However, this method can be applied only to muons or low-energy hadrons below a few hundred MeV. The majority of particles interact in one way or another before they stop. A calorimeter can measure the total energy of practically all known particles except muons and neutrinos. A calorimeter can measure the energy of neutral particles, too. They have become an indispensable element of modern high-energy collider experiments. They are suitable for observation of jets, which have their origins in energetic quarks and gluons, and by making a calorimeter a modular structure, it is possible to cover a large solid angle. The depth of a calorimeter grows only logarithmically with energy.

i) Total Absorption Cherenkov Counter This is a Cherenkov counter of total absorption type. A transparent material that contains a heavy metal having large density and short radiation length is suitable for the purpose. Lead glass ($\rho \sim 5$, $X_0 \sim 2$ cm), for instance, with a depth of $15\text{--}20X_0$ can contain almost all electromagnetic showers and can detect Cherenkov light, whose output is proportional to

the total length of the shower tracks. Thus it serves as an electromagnetic calorimeter and can measure the total energy of an incident electron or a gamma ray. Other materials that can be used include NaI, CsI, and BGO.

ii) Sampling Calorimeter For hadrons a total absorption counter made totally of active material is impractical because the size becomes too big. For economic reasons, a sandwich of iron or lead absorbers and active detectors is frequently used. As detectors plastic scintillators or liquid argon proportional chambers are used. The calorimeter is usually placed in the magnetic field, but it is better to place the photomultipliers outside the field. Figure 12.34a shows the basic structure of a calorimeter. Showers are sampled when they pass through the active detector. Ideally a separate counter should be used for each sampled signal, but in order to avoid having a large space occupied by light guides and photomultipliers, which may create a dead angle or space, light outputs are often transmitted to a wavelength-shifter and led to a photomultiplier through light guides. Figure 12.34b shows a configuration in which UV light is converted to blue light by the wavelength-shifter in the scintillator, which in turn hits a second wavelength-shifter after passing through the air gap and is converted to green light. Light with longer wavelength can travel further and thus can be guided along a long path to photomultipliers that are placed behind the absorber/radiator or where there is little effect of the magnetic field. Recently, optical fibers such as in Fig. 12.21 have also been used. As showers can produce a large number of electrons, ionization chambers can be used as detectors that can extract more accurate information. Here absorbers play a double role, being electrodes, too (Figure 12.34c). A practical design of a liquid argon calorimeter (that used in the ATLAS/LHC detector) is shown in Fig. 12.35.

As the difference between the electromagnetic and hadronic calorimeters lies in the use of radiation length versus interaction length, different materials and sampling rates should be used. Usually lead is used for electromagnetic calorimeters and iron for hadronic calorimeters. However, the structure can be the same. If $20 X_0$ of lead is used for the electromagnetic calorimeter, this amounts to 0.65 interaction

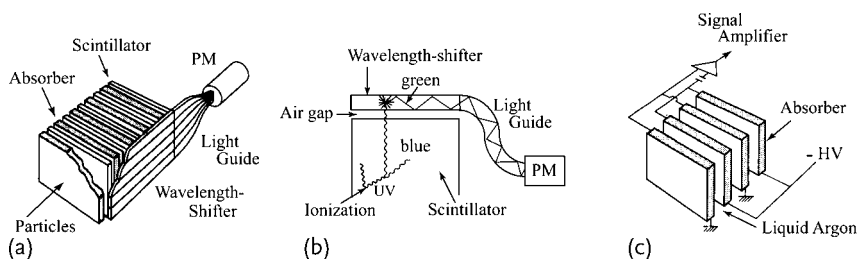


Figure 12.34 Structure of sampling calorimeters. (a) A calorimeter compactified to be used in a collider detector. To avoid a dead angle and to make the solid angle as large as possible, light is transported through a wavelength-shifter and light guide to distant photomulti-

pliers. (b) Conceptual and enlarged layout of the light transmission mechanism of (a). (c) Liquid argon is used instead of scintillators as photon detectors. Absorbers are immersed in liquid argon in a cryogenic vessel.

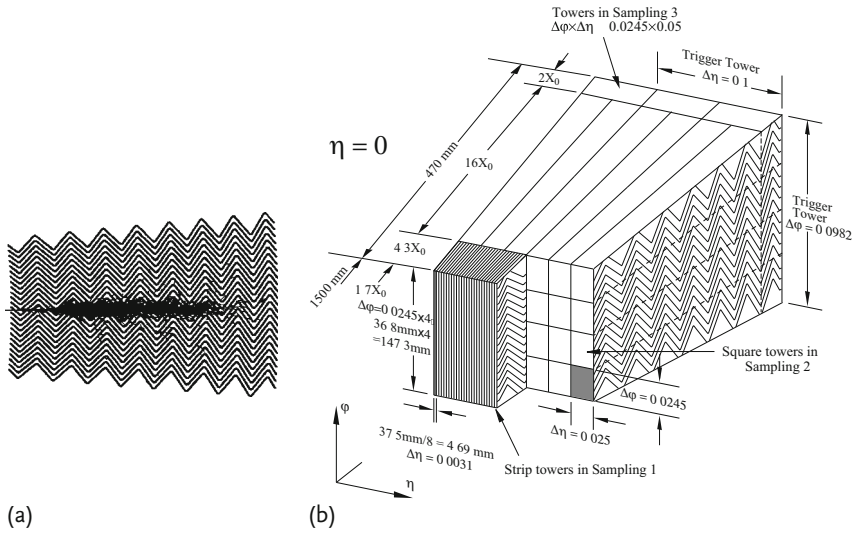


Figure 12.35 A module of the liquid argon calorimeter used in the ATLAS/LHC experiment. (a) Accordion structure of the sampling layers with a shower simulation superimposed. (b) Actual design of a module with solid angle coverage of $\Delta\eta \times \Delta\phi$ (radian) = 0.025×0.0245 . For

the definition of η , see Eq. (12.100). The thickness of a lead layer is 1–2 mm and a readout electrode made of thin copper and Kapton (polyimide) sheets is placed in the middle of the lead layers. The total thickness is 22 radiation lengths. It has attained an energy resolution of $\sigma/E = 10\%/\sqrt{E} + 0.3\%$ [32, 209].

length. For this reason, electromagnetic calorimeters are usually placed in front of hadronic calorimeters.

iii) **Energy Resolution of a Calorimeter** The calorimeter uses a shower event, which is a statistical phenomenon. The detected total length is proportional to the energy. If most of the shower can be contained in the calorimeter, the energy resolution σ is expected to be $\sim 1/\sqrt{E}$ considering the statistical (Poisson) nature of the shower. In the sampling calorimeter, the statistics become worse proportional to the absorber thickness d . The combined effects on the resolution are expected to be

$$\frac{\sigma}{\sqrt{E}} = k \sqrt{\frac{d}{E}} \quad (12.99)$$

As the above argument shows, the energy resolution of the calorimeter is worse at low energy ($\lesssim 10$ GeV). On the other hand, momentum resolution using the magnetic field is worse in proportion to the inverse of the momentum. Namely, in the low-energy region momentum measurement using a magnetic field is accurate, and at high energies the calorimeter is a better method. Modern collider detectors usually use both to cover all the energy regions.

Particle Identification (PID) The identification of particle species can be achieved by knowing its mass. It can be determined if two variables out of v , $E = m\gamma$, $p =$

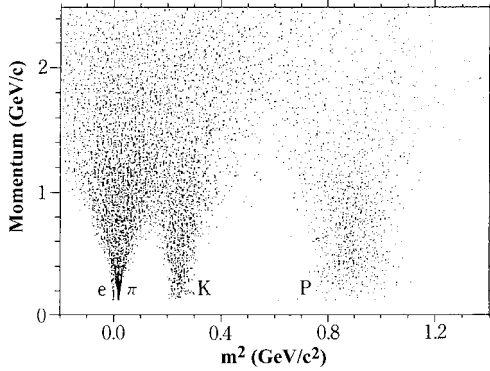


Figure 12.36 Particle identification using TOF method ($\Delta t = 200$ ps, $L = 175$ cm).

$m\beta\gamma$ are known. We have already shown that at fixed momentum the Cherenkov counter can select velocities and can identify particles.

i) Time-of-Flight (TOF) Method Particle identification is done by measuring the velocity of particles by the most direct method, namely by measuring the time of flight between two counters. This method usually uses two scintillation counters, but any counters capable of fast timing can be used. Figure 12.36 shows data taken by TOF counters in VENUS/TRISTAN, which has $\Delta t = 200$ ps and a flight distance of 175 cm. The figure shows that particles of $p \lesssim 1$ GeV/c can be identified with the above setup.

ii) E/p Method The momentum of a particle is measured with a track-measuring device in the magnetic field and the energy is measured by electromagnetic calorimeters. Electrons give $E/p \simeq 1$, but hadrons (pions) give $E/p \ll 1$, because they hardly deposit any energy in the electromagnetic calorimeter as λ (nuclear interaction length) $\gg X_0$ (radiation length). Usually the method is used to separate electrons from hadronic showers, but taking the shower profile further into consideration, a rejection factor e/π of 10^3 – 10^4 has been obtained without degrading the detection efficiency ($\epsilon \gtrsim 90\%$) [134, 389].

iii) Method of dE/dx As dE/dx is sensitive only to β , knowing the particle momentum enables PID to be achieved. For this purpose dE/dx has to be measured accurately. The most advantageous detector for this method is the time projection chamber (TPC). It can measure p , dE/dx and space position of all particles simultaneously and thus can reconstruct the topology of an entire reaction that has occurred in the chamber. It is a variation of the drift chamber and its structure is depicted schematically in Fig. 12.37. The chamber is a large-diameter cylindrical tube with a central partition that works as a cathode and two endcaps that work as anodes. It is placed in a solenoidal magnet and is used as a collider detector. The beam goes through the central axis of the TPC and the interaction point is at the center of the chamber.

A TPC that was used for TOPAZ/TRISTAN experiments has an inner radius of 2.6 m, length of 3 m and the applied voltage is -43.25 kV [223]. All the secondary

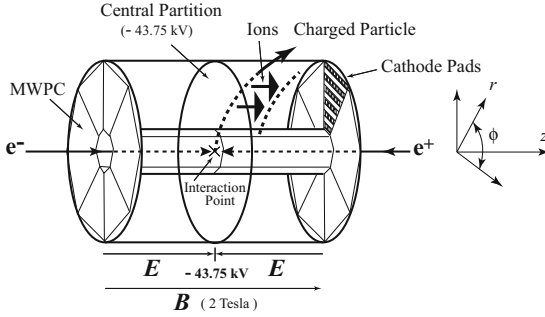


Figure 12.37 Illustration of the TPC used in TOPAZ/TRISTAN [223]. Inner radius 2.6 m, length 3 m. Ions from particle tracks are attracted away from the center of the TPC in the z direction. At the end wall anode wires of eight trapezoidal MWPCs read the r coordinate of the ions. At the back of the anodes,

cathode pads obtain the ϕ coordinate by reading the induced voltage. The arrival time gives the z coordinate and the number of ions gives dE/dx information. By reconstructing tracks and measuring their curvature, the momentum is determined and with dE/dx the particle species can be identified.

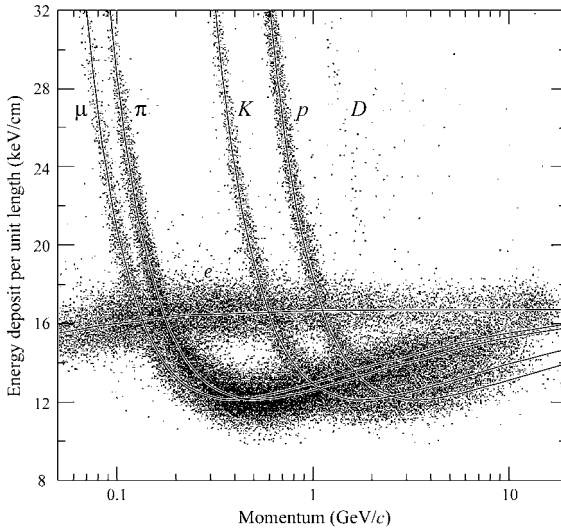


Figure 12.38 $dE/d\xi$ curves as a function of momentum. As $dE/d\xi$ is a function of $\beta\gamma$ only, curves made by different particle species at fixed momentum $p = m\beta\gamma$ separate. Figure from [311].

particles go through the chamber except at small solid angles along the beam, therefore it essentially covers 4π solid angle. A magnetic field is applied parallel to the cylinder axis, which is used to measure particle momenta. It is also designed to prevent transverse diffusion of ions so that they keep their r - ϕ position until they reach the anodes. Ions generated along tracks of the secondary particles drift parallel to the beam axis at constant velocity toward the endcap. The endcap con-

sists of eight trapezoidal MWPCs, which can read the r position of the ions. Pads at the back side of the MWPCs read the induced voltage capacitively and obtain the coordinate ϕ perpendicular to r . From the drift time of ions the z coordinate is known. The signal strength at the anode is proportional to the ionization loss of the particles, and is sampled as many times as the number of wires, giving good statistical accuracy for the measurement of dE/dx . From the curvature of the tracks in the magnetic field, the momentum is determined and together with dE/dx information this allows particle species to be identified. Figure 12.38 is a $dE/d\xi$ curve obtained by the TPC of the PEP4 experiment [311].

As described above, the TPC is capable of reconstructing three-dimensional coordinates of all the tracks, momenta and dE/dx . The information obtained is comparable to that of bubble chambers and yet TPC can function at high speed. It is a very powerful detector suitable for electron–positron collider experiments. However, it is not without its limitations in that it takes 30–50 μs for ions to reach the endcaps, and has difficulties in applications that require high-repetition data taking, such as in hadron collider experiments.

12.6

Collider Detectors

a) Basic Concept The mainstream modern detector is a general-purpose collider detector, which is a combination of detectors with various functions. Many of them share a common basic structure as given in Fig. 12.39, where the basic elements and the particle's behavior in them are illustrated. A typical example, the ALEPH detector used in LEP/CERN [10], is shown in Fig. 12.40. It is used to measure Z^0 properties accurately and, together with other detectors at LEP, has been instrumental in establishing the validity of the Standard Model.

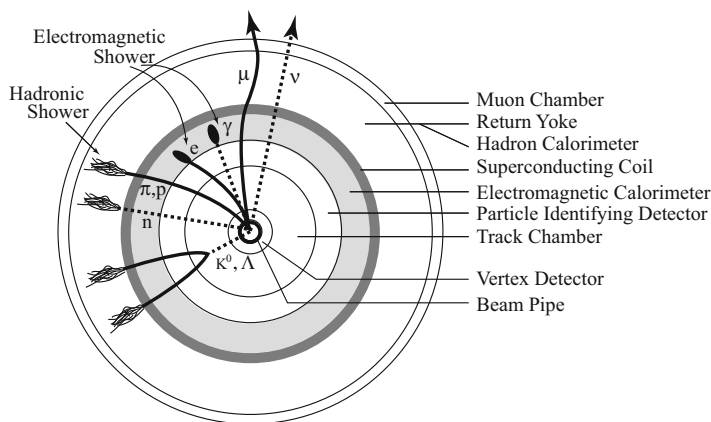


Figure 12.39 Schematic diagram of a typical collider detector shown in cross section perpendicular to the beam axis. Dotted lines indicate neutral and hence invisible particles.

We list a few characteristics common to all the general-purpose collider detectors.

(i) General purpose: They are made to respond to almost all ($\gtrsim 90\%$) kinds of conceivable reactions. Particle species that can be identified include e (electrons and positrons), γ and π^0 , μ , charged hadrons (mostly pions) and neutral hadrons (neutron, K^0 , Λ , etc.).

(ii) Necessary conditions for a general-purpose detector are:

1. capability of identifying a particle's species, measuring the energy-momentum no matter in which direction a particle flies,
2. uniformity of the sensitivity and accuracy over the complete solid angle,
3. hermeticity: ideally the detector should have no dead angle, but practically the beam pipe makes a dead zone in the forward and backward regions. But at least the momentum and energy components of all the particles transverse to the beam direction should be able to be measured. This is important in order to measure the missing energy, which gives important information about the production of neutrinos or some of the suggested new particles.

As can be seen from Fig. 12.40, the detector is separated into two parts, a barrel region ($45^\circ \lesssim \theta \lesssim 135^\circ$) and an endcap region (a few degrees $\lesssim \theta \lesssim 45^\circ$). They have radial symmetry in the cross-sectional view perpendicular to the beam direction (Fig. 12.41) and also forward-backward symmetry with respect to the interaction point.¹⁶⁾ The barrel and endcap parts have different geometrical structure

16) This is not the case for asymmetric colliders like those at B-factories and HERA (e-p collider at DESY).

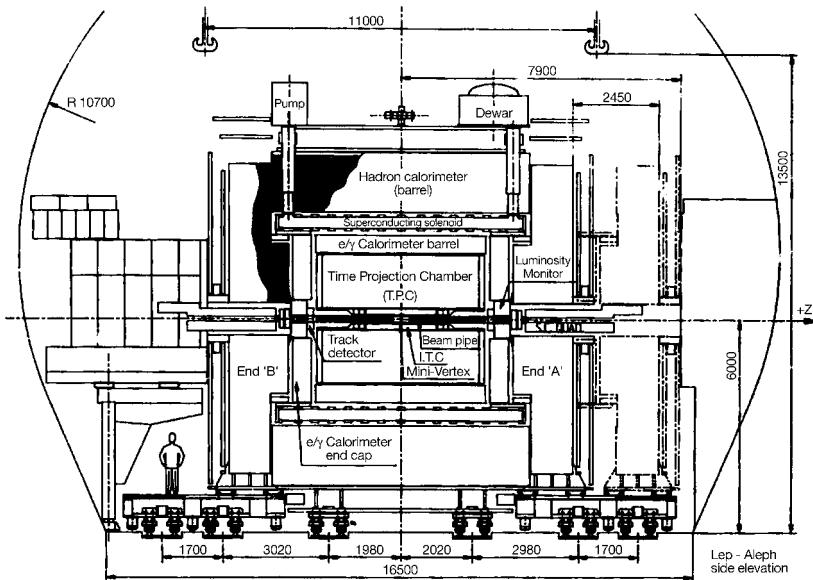


Figure 12.40 Side view of the ALEPH detector at LEP [10].

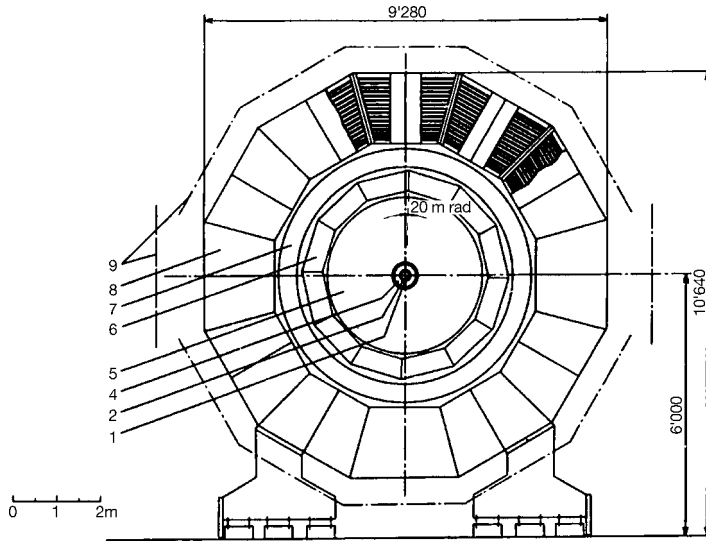


Figure 12.41 Cross-sectional view of the ALEPH detector at LEP [10]. The ALEPH detector consists of (from the center outwards) 1. beam pipe, 2. silicon semiconductor vertex

detector, 4. inner track chamber, 5. TPC, 6. electromagnetic calorimeter, 7. superconducting coil, 8. hadron calorimeter, and 9. muon chamber.

and necessitate different approaches in designing them, but functionally it is desirable to have uniformity to avoid experimental errors.

We will describe a few of the important functions of the components.

b) Measurement of Interaction and Decay Points of Short-Lived Particles Reactions produced by beam particles colliding with residual gas in the beam are unwanted events. To remove them it is necessary to determine the interaction point by tracing back several tracks. Track detectors inserted in the magnetic field can reproduce them in principle, but if one wants to measure decay points of short-lived particles (charm and beauty particles), which have lifetimes $10^{-12} \text{ s} \lesssim \tau \lesssim 10^{-13} \text{ s}$ and fly a distance of only $\ell = c\tau\gamma$, i.e. tens or hundreds of micrometers, a specialized vertex (crossing points of tracks) detector is needed. SSD detectors are used, and resolution of $\sim 10 \mu\text{m}$ in the $r\phi$ direction, $\sim 20 \mu\text{m}$ in the z direction were realized with the ALEPH detector. Figure 12.42 shows an example of decay-point measurements and Fig. 12.43 a lifetime curve of a short-lived particle.

c) Momentum Analysis Momenta of charged particles can be determined with a combination of magnet and track chamber. A superconducting solenoidal coil is used to generate a strong magnetic field of 1–1.5 T (Tesla). The direction of the magnetic field is usually chosen along the beam direction and thus only the transverse component of the momentum is measured but with knowledge of the polar angle θ of the particle the total momentum can be reconstructed. The tracking chamber uses drift chambers of one type or another. By stretching wires along the beam direc-

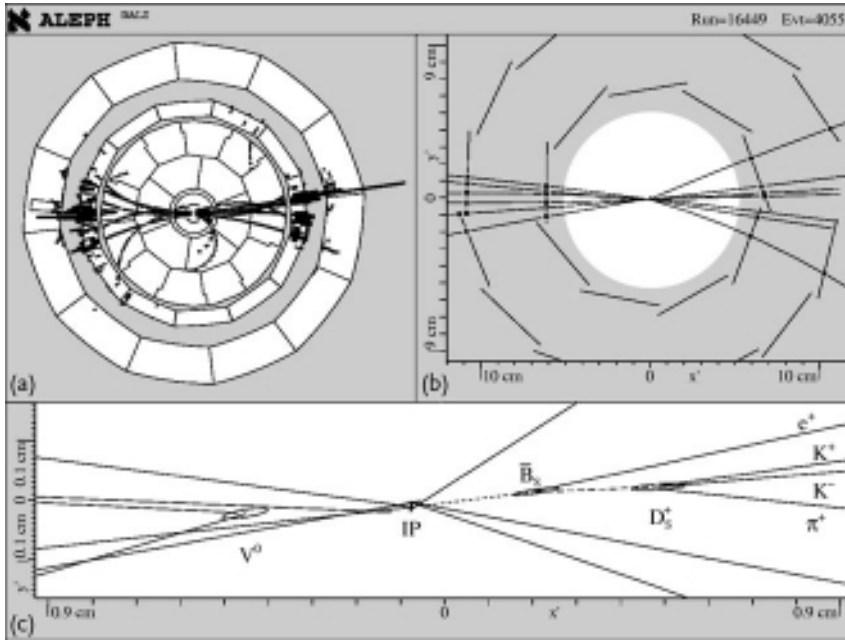


Figure 12.42 Display of a two-jet event recorded by the ALEPH detector, in which a B_{0S} candidate has been reconstructed in the following mode: $B_{0S} \rightarrow D_S^+ + \bar{e} + \nu_e$, $D_S^+ \rightarrow \varphi\pi^+$, $\varphi \rightarrow K^+ \dots$ (a) Overall

view in the plane transverse to the beams. (b) Detailed view of the vertex detector, showing the hits in the two layers of silicon. (c) Zoom close to the interaction point (IP), showing the reconstructed tracks and vertices [106].

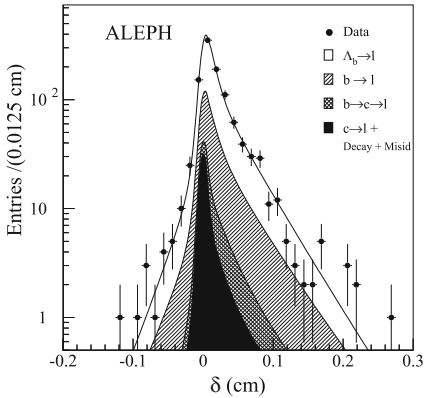


Figure 12.43 Lepton impact parameter distribution of $A\ell^-$ candidates. The solid curve is the probability function at the fitted value of the lifetime while the shaded areas show, one on the top of the other, the background contributions [11].

tion a good spatial accuracy can be obtained for r - ϕ coordinates. The z coordinates are usually obtained by layers of slanted wires stretched at a small angle relative to the main layers and their accuracy is not as good as that of r or ϕ . ALEPH used a TPC, and with $B = 1.5$ T, $\Delta s = 100 \mu\text{m}$ obtained $\delta p_T/p_T^2 = 0.0009 (\text{GeV}/c)^{-1}$. $\Delta s = 100 \mu\text{m}$ was obtained for particles having transverse track length $L = 1.4$ m. Because of the need to minimize multiple scattering, which restricts the accuracy of momentum and also of the long track length, the tracking detector is usually placed in the central, inner part of the whole detector. The track chamber can also identify unstable neutral particles such as K^0 and Λ by detecting their decay vertices.

d) Energy Vectors Calorimeters are used to measure energies of charged as well as neutral particles and are usually placed outside the tracking chambers. They are generally modularized to obtain space point information. One module is usually designed geometrically to have a fixed interval of rapidity $\Delta\eta$ and azimuthal angle $\Delta\phi$ so that counting rate per module becomes uniform. The QCD hadron production rate is known to be roughly constant in rapidity η or pseudorapidity η' , which is related to polar angle θ by

$$\eta' = \frac{1}{2} \ln \frac{E + p_{\parallel}}{E - p_{\parallel}} \simeq \frac{1}{2} \ln \frac{1 + \cos \theta}{1 - \cos \theta} = -\ln \tan \frac{\theta}{2} \quad (12.100)$$

In ALEPH, each module extends over a solid angle of $3 \times 3 \text{ cm}^2$ or $0.8^\circ \times 0.8^\circ$ and is distributed lattice wise. A position resolution of 1.5 mm in determining the shower generation point was obtained. By connecting this point with the interaction point the direction and energy (i.e. energy vector) of a particle can be measured. Calorimeters are physically and functionally divided into two parts, the electromagnetic calorimeter placed inside and the hadronic calorimeter placed outside the superconducting coil. The former is composed of a lead/scintillation counter sandwich, which has a total of 23 radiation lengths, whereas the latter is a sandwich of iron plates with 5–10 cm thickness and streamer tubes (a gas tube that works in streamer mode). The total absorption length is $\lambda = 7.6$ or ~ 1.2 m of iron. The iron absorber plays a double role as the return yoke of the magnetic flux created by the superconducting coil. As described, a particle's momentum is measured by the track chamber, its energy by the calorimeter and, using the E/p method, electrons (γ 's too if there are no tracks) and hadrons are separated. In ALEPH, the contamination of hadrons in the electron signal was less than 10^{-3} . The energy resolution was $\Delta E/E = 1.7 + 19/\sqrt{E} [\text{GeV}] \%$ for the electromagnetic calorimeter and $\sim 60\%/\sqrt{E} [\text{GeV}]$ for the hadronic calorimeter.

e) Particle Identification Depending on the physics aim of the detector, a particle identification device such as a transition radiation counter to separate e from π or a RICH counter to identify hadron species is usually inserted between the tracking chamber and the calorimeter. ALEPH's TPC serves a double role as a tracking chamber and a particle identification device.

f) Muon Identification The muon can be identified as a noninteracting long track. Therefore a track detector outside the magnet yoke, sometimes with additional absorbers totalling $\gtrsim 10\lambda$, serves as the muon identifier. The only other particle that can penetrate this much absorber is the neutrino or yet undiscovered weakly interacting particles. Backgrounds for muons directly produced in the reaction are decay muons from pions and false signals of pions, a small albeit nonnegligible component of pions that survive after the absorbers.

g) Missing Energy In the center of mass system of the colliding particles, the total sum of momenta should vanish. Missing energy is defined as the vector sum of all the energy vectors of charged as well as neutral particles but with the sign inverted. As each hadronic particle is not identified, they are assumed to be pions or zero-mass particles. Then the energy vector is the same as the momentum vector. When it is known that only one neutrino is missing the missing energy vector is that of the neutrino. A reaction to produce $W(u + \bar{d} \rightarrow W^+ \rightarrow e(\mu)^+ + \nu)$ was measured this way and the mass of W was determined.

New particles that are predicted by theoretical models, for example supersymmetry (SUSY) particles, also give a missing energy accompanied by (an) energetic lepton(s) or mono-jet. The accurate measurement of the missing energy is an important element of a general-purpose detector. This is why the hermeticity of the detector is important. Still, energies that escape to the beam pipe are sizeable and only the transverse component of the energy vectors can be determined.

h) Luminosity Measurement In addition to the main detector, small luminosity counters made of a tracking chamber and a calorimeter are placed in the forward and backward regions and measure well-known reactions such as small-angle Bhabha (e^-e^+ elastic) scattering for electron colliders. As the cross section is well known and calculable, knowing the counting rate gives back the intensity of incident beam. It is necessary to measure an absolute cross section of a reaction.

i) Summary A collider detector has all the properties described above, and can serve as a general-purpose detector. If the total energy $\sqrt{s}$ is known, reactions that can be induced can be more or less be predicted and similarity in designs of different groups is conspicuous. For instance, at the LEP/CERN collider, $\sqrt{s} = m_{Z^0} = 91$ GeV and four detectors ALEPH, DELPHI, L3 and OPAL were constructed. The major goals to measure the mass, decay width and modes of Z^0 accurately were common to all and all the detectors were capable of doing it. Therefore, from the standpoint of testing the validity of the Standard Model, differences between the detector performances can be considered small. The difference, however, appears in the details of the detector component, depending on whether one wants to make precision measurements of leptonic modes or to carry out detailed tests of QCD or whatever. As a result of this, sensitivity to or the potential for in-depth investigation of a new phenomenon may be different. How much effort is put into apparently irrelevant functions of detectors strongly reflects the group's philosophy and insight.

12.7

Statistics and Errors

12.7.1

Basics of Statistics

Particle reactions are statistical phenomena. Detectors are devices that convert a particle's passage into signals that a human can recognize and their performance is also dictated by statistics. We describe here basic concepts of statistical phenomena to understand the quality of measurements and the errors associated with them. Frequently used formula that are stated in a standard textbook of statistics and probability are given here without proof. Proofs are given only when they are specially connected to particle measurements.

There are no measurements without errors. They come in two kinds, systematic and statistical. Systematics are associated with the setup of reference points of measurement and make measured values either overestimated or underestimated. When the source of the error is understood, it could be removed in principle. But some part of it remains always unexplained or uncontrollable. For instance, we say the peak of Mt. Chomolungma is 8848 m above sea level. The reference point is at sea level $h = 0$, but the value of h could be either $+h_0$ or $-h_0$ and we never know its exact value. However, by careful arrangement, it is possible to set an upper limit to it, i.e. $(|h| < \varepsilon)$ and ε is referred to as the systematic error in the measurement of mountain heights. The systematic error has its origin in the setup of each experiment (or an experimenter's skill) and cannot be discussed in unbiased way. Therefore, we only discuss statistical errors. Their origin inherently lies in the measured object itself and/or in the measuring devices. The measured value fluctuates irregularly (at random) around a true value. Repeated measurement can decrease the size of statistical errors, but smaller values than the systematic error are fruitless, so usually measurements are ceased when the statistical error becomes comparable to the systematic error.

The behavior of errors or their distribution depends on the circumstances. When the probability function of events is unknown, a normal distribution, to be described later, is assumed. With such a blind assumption, usually no serious mistakes are induced in most circumstances because of

1. the law of large numbers,
2. central limit theorem, and
3. general properties of the normal distribution.

Law of Large Numbers Given a random variable x with a finite expected value (μ), if its values are repeatedly sampled (sample value x_i), as the number n of these observations increases, their mean will tend to approach and stay close to the expected

value. Namely,

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{x_i}{n} \rightarrow \mu \quad (12.101)$$

Central Limit Theorem The central limit theorem states that as the sample size n increases the distribution of the sample average of these random variables approaches the normal distribution with a mean μ and variance (to be described soon) σ^2/n irrespective of the shape of the original distribution.

However, one has to be careful because very often one tends to put too much confidence in the omnipotentiality of the normal distribution.

We first define some terminology and a few definitions.

Average Value or Mean If sample events described by a variable x are realized independently according to a distribution function $f(x)$, the average μ is defined by

$$\mu = \int dx x f(x) \equiv E(x) \quad (12.102)$$

with normalization

$$\int dx f(x) = 1 \quad (12.103)$$

Generally $E(g(x))$ is defined by

$$E(g(x)) = \int dx g(x) f(x) \quad (12.104)$$

When the variable x is not continuous but discrete, replace $f(x_i)$ by f_i and take the sum instead of integral ($\int dx \rightarrow \sum_i$). In the following we take x as a representative of both continuous and discrete variables unless otherwise stated.

Variance This is a measure of dispersion of the variable x around the mean μ and is defined by

$$\sigma^2 \equiv E((x - \mu)^2) = \int dx (x - \mu)^2 f(x) \quad (12.105)$$

$\sigma = \sqrt{\sigma^2}$ is called the standard deviation. It is easy to verify that

$$E((x - \mu)^2) = E(x^2) - E(x)^2 = E(x^2) - \mu^2 \quad (12.106)$$

Correlation When there are two statistical variables x, y with means μ_x, μ_y and standard deviations σ_x, σ_y , their covariance and correlation r are defined by

$$\text{cov}(x, y) = E((x - \mu_x)(y - \mu_y)) \quad (12.107a)$$

$$r = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} \quad (12.107b)$$

r takes values $-1 \leq r \leq 1$. When there exists a linear relation such as $y = ax + b$, $|r| = 1$, and when x and y are completely independent, $r = 0$. “Independent” means the distribution function can be separated like $f(x, y) = g(x)h(y)$. Figure 12.44 shows some examples of correlation.



Figure 12.44 Examples of correlation.

Normal Distribution The normal distribution is also called the Gaussian distribution. It is the most frequently occurring function in all statistical phenomena. Generally, measured values and detector fluctuation (errors) can be considered as obeying the normal distribution. The functional form with mean μ and standard deviation σ is given by

$$G(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right] \quad (12.108)$$

Figure 12.45 shows its functional form. The area in 1σ range ($\mu - \sigma \leq x \leq \mu + \sigma$) is 68.3%, 2σ 95.5%, and 3σ 99.7%.¹⁷⁾ In the figure, 90% of the total area is shown as a white area with boundary at $|x - \mu| = 1.64\sigma$. It is worth remembering that the probability of the variable x deviating more than one standard deviation is as large as 30%.

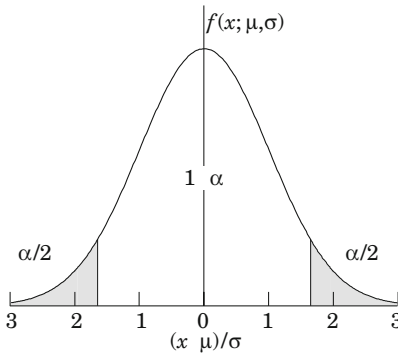


Figure 12.45 Gaussian distribution ($\alpha = 0.1$). 90 % of the area is within the region $(x - \mu)/\sigma = \pm 1.64$.

Another measure of dispersion is the full width at half maximum (FWHM), which is often quoted instead of the standard deviation. This is the x range where the probability exceeds half the maximum value. FWHM is denoted as Γ and for the normal distribution is given by

$$\Gamma = 2\sigma\sqrt{2\ln 2} = 2.35\sigma \quad (12.109)$$

Overlap of the Gaussian Distribution Let there be two independent variables x_1, x_2 , both following Gaussian distributions with mean 0 and variance σ_1^2, σ_2^2 , and mea-

¹⁷⁾ Conversely, the values of n ($|x - \mu| < n\sigma$) that give 90%, 95% and 99% area are 1.64, 1.96 and 2.58.

sured as the combination

$$x = ax_1 + bx_2 \quad (12.110)$$

Then their overlap will obey the distribution

$$G(x) = \int dx_1 dx_2 \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{x_1^2}{2\sigma_1^2}\right] \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left[-\frac{x_2^2}{2\sigma_2^2}\right] \quad (12.111)$$

Its Laplace transform is given by

$$\int_0^\infty dx e^{\lambda x} G(x) = \exp\left[\frac{1}{2}\lambda^2(a^2\sigma_1^2 + b^2\sigma_2^2)\right] \quad (12.112)$$

This is the Laplace transform of the Gaussian with the standard deviation

$$\sigma = \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2} \quad (12.113)$$

Namely, the error of an overlap of two independent statistical phenomena is given by the square root of the weighted sum of the error squared. If the difference instead of the sum is taken in Eq. (12.110), the error is still given by the same formula.

Poisson Distribution Uniformly distributed but independent events obey the Poisson distribution. Reactions and decays in particle physics, the number of ions per unit length created by passing charged particles, electrons emitted from a heated metal surface or the photocathode of photomultipliers, all of which are intimately connected with detector operations, obey the Poisson distribution. If we let the average occurrence of an event per unit time be μ , then the probability that it occurs r times in time t is given by

$$P(\mu; r) = \frac{(\mu t)^r}{r!} e^{-\mu t} \quad (12.114)$$

This can be explained as follows. Let us assume that the event occurred M times in a time interval T . Both T and M are taken to be large enough. Then the probability that r events fall in a small time interval t ($t \ll T$) is given by $p^r(1-p)^{M-r}$, where $p = t/T$. As there are ${}_M C_r$ ways to sample r out of M events, the probability that r events occur in the time interval t is given by the binomial distribution

$$\begin{aligned} P(M; r) &= {}_M C_r p^r (1-p)^{M-r} \\ &= \frac{M(M-1)\cdots(M-r+1)}{r!} p^r (1-p)^{M-r} \\ &= \frac{1 \cdot \{1-1/M\} \cdots \{1-(r-1)/M\}}{r!} (Mp)^r \left(1 - \frac{Mp}{M}\right)^{M-r} \end{aligned} \quad (12.115)$$

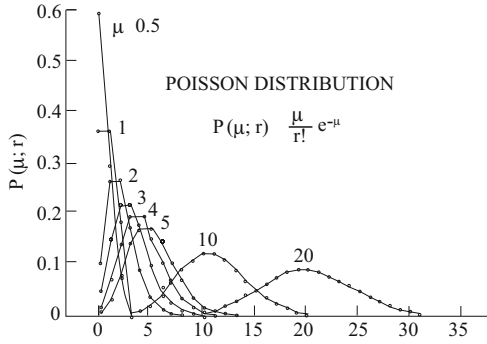


Figure 12.46 Poisson distribution: $(\mu^r / r!) e^{-\mu}$.

Taking the limits $M \rightarrow \infty$, $T \rightarrow \infty$, keeping $Mp = Mt/T \rightarrow \mu t$ fixed, and using $\lim_{M \rightarrow \infty} (1 - x/M)^M = e^{-x}$, we obtain the Poisson distribution Eq. (12.114). The mean occurrence $\bar{r}$ of events in unit time interval and the variance of the Poisson distribution are given by

$$\bar{r} = \sum_{r=0}^{\infty} r P(r) = \mu, \quad \sigma^2 = \sum_{r=0}^{\infty} (r - \mu)^2 P(r) = \mu \quad (12.116)$$

Figure 12.46 shows the Poisson distributions for several values of μ . Note the distribution is not symmetric around μ . If we equate $\mu = n$, when n is large the Poisson becomes the Gaussian with mean n and variance n . We might as well say it is a good approximation already at $n = 10$ (see Fig. 12.46). This is an example of the central limit theorem.

Problem 12.3

When n is large, show that the Poisson distribution with mean n , variance n approaches the Gaussian distribution:

$$\lim_{n \rightarrow \infty} \left(f(r) = \frac{n^r}{r!} e^{-n} \right) \rightarrow \frac{1}{\sqrt{2\pi n}} \exp \left[-\frac{(r - n)^2}{2n} \right] \quad (12.117)$$

Hint: Use Sterling's formula

$$n! = \sqrt{2\pi n} n^{n+1/2} e^{-n} \quad (12.118)$$

and expand $\ln f(r)$ in a Taylor series.

Example of Application: Deciding the Upper Limit of the Probability Value when no Events Are Observed When one tries to determine the validity of a certain theoretical formula, for instance the lifetime of double beta decay, and ends up with no observation, the upper limit of the rate or the lower limit of the lifetime can be determined as follows. Let the decay rate be λ . The probability that no events are

observed during the observation time T is given by the Poisson distribution as

$$P(\lambda; 0) = e^{-\lambda T} \quad (12.119)$$

The equation can be interpreted as the probability of observing no events as a function of λ . Therefore, the probability $\lambda \leq \lambda_u$ is given by

$$P(\lambda \leq \lambda_u) = T \int_0^{\lambda_u} d\lambda e^{-\lambda T} = 1 - e^{-\lambda_u T} \equiv \text{CL} \quad (12.120)$$

T is the normalization constant when λ is taken as the variable. We call the above equality the confidence level (CL) of λ lying in the interval $0 \leq \lambda \leq \lambda_u$,

$$\therefore \lambda_u = -\frac{1}{T} \ln(1 - \text{CL}) \quad (12.121)$$

For CL = 90%, $\lambda_u = 2.3/T$. Namely, when no events are observed, the upper limit with 90% confidence level is given by the rate with 2.3 as the number of events.

Simultaneous Observation of Two Independent Phenomena Suppose one is measuring the lifetime of a particle. The counts one obtains include noise of the detector, which is also statistical but of different and independent origin:

$$n = n_1(\text{particle}) + n_2(\text{noise}) \quad (12.122)$$

The probability of n events being observed is given by

$$\begin{aligned} P(\mu_1, \mu_2; n) &= \sum_{n_1 + n_2 = n} P(\mu_1; n_1) P(\mu_2; n_2) = \sum \frac{(\mu_1)^{n_1}}{n_1!} e^{-\mu_1} \frac{(\mu_2)^{n_2}}{n_2!} e^{-\mu_2} \\ &= \frac{(\mu_1 + \mu_2)^n}{n!} e^{-(\mu_1 + \mu_2)} = P(\mu_1 + \mu_2; n) \end{aligned} \quad (12.123)$$

12.7.2

Maximum Likelihood and Goodness of Fit

a) Maximum Likelihood Method The real value μ and its variance of a measured object can never be known, but a good guess can be obtained by measurements. Suppose there are n independent samples x_i ($i = 1 \dots n$), which we call a statistical variable obeying a distribution function $f(x; a)$. a is a parameter to be determined. The maximum likelihood function is defined by

$$L(x_1, \dots, x_n; a) = f(x_1; a) f(x_2; a) \dots f(x_n; a) \quad (12.124)$$

It can also be considered as the probability that gives the observed values $(x_1, x_2, \dots, x_n)$. If the function L is analytic, the value of a can be obtained by solving

$$\begin{aligned} \frac{dL}{da} &= 0 \\ \text{or } \frac{dL^*}{da} &\equiv \frac{d \ln L}{da} = 0 \end{aligned} \quad (12.125)$$

The two equations are equivalent in obtaining a and either one is used depending on the situation. When there is more than one parameter, simultaneous equations are obtained. When the solution of Eq. (12.125) is written as $\hat{a}$, it is not the true a but an estimated value. An estimated variance is given by

$$\hat{\sigma}^2 = \int dx_1 \dots dx_n (a - \hat{a})^2 L(x_1, \dots, x_n; a) \quad (12.126)$$

This formula, however, is not necessarily calculable unless the analytical form of $f(x; a)$ is known. When the probability function is given by the Gaussian, the maximum likelihood function is given by

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right] \quad (12.127)$$

If we write the estimate of μ as $\hat{\mu}$ and that of σ^2 as $\hat{\sigma}^2$, L is maximized by

$$\frac{\partial L}{\partial \mu} = L \sum \frac{x_i - \mu}{\sigma^2} = 0 \quad \Rightarrow \quad \hat{\mu} = \frac{1}{n} \sum x_i = \bar{x} \quad (12.128a)$$

$$\frac{\partial^2 L}{\partial \mu^2} = L \left[-\frac{n}{\sigma^2} + \sum \frac{1}{\sigma^2} \left(\frac{x_i - \mu}{\sigma^2} \right)^2 \right] = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \quad (12.128b)$$

As the real μ is unknown, it is replaced by $\hat{\mu}$. Then the degrees of freedom are reduced by 1 and it is known that

$$s^2 \equiv \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 = \frac{n}{n-1} \hat{\sigma}^2 \quad (12.129)$$

gives a better estimate of the variance. If n is sufficiently large, the difference is small. The arithmetic average $\bar{x}$ is an estimate of μ , but is not the real μ . This means another n samples will give a different $\bar{x}$. In this sense, $\bar{x}$ itself is a statistical variable. Consequently, the variance of $\bar{x}$ can be defined:

$$\begin{aligned} \sigma^2(\bar{x}) &= E((\bar{x} - \mu)^2) = \frac{1}{n^2} E\left(\left(\frac{1}{n} \sum_i x_i - \mu\right)^2\right) = \frac{1}{n^2} \sum E((x_i - \mu)^2) \\ &= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned} \quad (12.130)$$

Namely $\sigma(\bar{x}) = \sqrt{\sigma^2(\bar{x})} = \sigma/\sqrt{n}$ gives the error associated with the arithmetic mean $\bar{x}$. $\sigma(\bar{x})$ is referred to as the standard error of the mean. As is clear from the

above derivation, the equality $\sigma(\bar{x}) = \sigma/n$ holds independently of the functional form of the distribution.

The reader is reminded here that $\hat{\sigma}^2$ is an estimate of σ^2 , which is the real variance of the variable x , and not an estimate of the variance of $\bar{x}$. σ decides the fluctuation of the variable x_i distribution and does not change no matter how many samples are taken. $\hat{\sigma}^2$ approaches σ^2 as n increases. Likewise $\bar{x}$ approaches true μ as n increases, and its standard deviation is given by $\sigma/\sqrt{n}$, which approaches 0 as n increases.

Now, note that at a point $\partial L^*/\partial \mu = 0$,

$$-\frac{\partial^2 L^*}{\partial \mu^2} = \frac{n}{\sigma^2} = \frac{1}{\sigma^2(\bar{x})} \quad (12.131)$$

Equation (12.126) is a general expression, but when it cannot be calculated analytically one applies the central limit theorem and assumes the same formula as for the Gaussian distribution is valid. Namely,

$$\frac{1}{\sigma^2(\hat{a})} = -\left[\frac{\partial^2 L^*}{\partial a^2} \right]_{a=\hat{a}} \quad (12.132)$$

Note, however, this is generally an approximation and valid only when the statistics is large. When more than one parameter is involved, an error matrix U can be defined as the inverse of

$$(U^{-1})_{ij} = -\frac{\partial^2 L^*}{\partial a_i \partial a_j} \quad (12.133)$$

Then it can be shown that the diagonal elements σ_i^2 are equal to the variance of $\hat{a}_i$ and the nondiagonal elements are covariances. Namely

$$U = \begin{bmatrix} \sigma_1^2 & \text{cov}(1, 2) & \text{cov}(1, 3) & \cdots & \cdots \\ & \sigma_2^2 & \text{cov}(2, 3) & \cdots & \cdots \\ & & \sigma_3^2 & \cdots & \cdots \\ & & & \cdots & \cdots \end{bmatrix} \quad (12.134)$$

We note that the estimate determined by the maximum likelihood method is invariant under transformation. That is, if $b = g(a)$, the best estimate of b is given by $\hat{b} = g(\hat{a})$.

b) Propagation of Errors One often wants to calculate another quantity using the obtained sample variables. Here we are interested in the error of $u = f(x, y)$ given (x, y) with standard deviations σ_x, σ_y . The variance σ_u^2 of u is given by

$$\sigma_u^2 = E((u - \bar{u})^2) \quad (12.135)$$

In the first approximation, $\bar{u} = f(\bar{x}, \bar{y})$. How the variable u moves around $\bar{u}$ can be obtained by the Taylor expansion

$$u - \bar{u} \simeq (x - \bar{x}) \left. \frac{\partial f}{\partial x} \right|_{\bar{x}} + (y - \bar{y}) \left. \frac{\partial f}{\partial y} \right|_{\bar{y}} \quad (12.136)$$

Inserting the square of Eq. (12.136) in Eq. (12.135), we obtain

$$E((u - \bar{u})^2) = E \left[(x - \bar{x})^2 \left(\frac{\partial f}{\partial x} \right)^2 + (y - \bar{y})^2 \left(\frac{\partial f}{\partial y} \right)^2 \right] + 2(x - \bar{x})(y - \bar{y}) \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \quad (12.137)$$

Using Eqs. (12.105) and (12.107), we obtain

$$\sigma_u^2 = \sigma_x^2 \left(\frac{\partial f}{\partial x} \right)^2 + \sigma_y^2 \left(\frac{\partial f}{\partial y} \right)^2 + 2\text{cov}(x, y) \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \quad (12.138)$$

Namely, the variance of u is the sum of the squares of its argument errors multiplied by their derivative. The third term vanishes if the variables x and y are independent, but if there is correlation between the two, the propagation error is modified greatly from the sum of squares. In the following we ignore correlations.

Example 12.1

$$u = x \pm y$$

$$\sigma_u^2 \simeq \sigma_x^2 + \sigma_y^2 \quad (12.139)$$

which reproduces Eq. (12.113).

Consider an experiment with a neutron target. This is achieved by subtracting proton target data from deuteron target data. Let n_1, n_2 be the number of events from each experiment measured in the time intervals t_1, t_2 . One wants to decide the best balance in dividing the time into t_1 and t_2 with the total time $T = t_1 + t_2$ fixed. The counting rates are $r_1 = n_1/t_1, r_2 = n_2/t_2$ and $r_1 - r_2$ is the desired number. As the distribution is Poissonian, the variance σ^2 of n is $\sigma^2 = n$ and the error $\delta n = \sqrt{n}$. Using Eq. (12.139),

$$\delta r^2 = \delta r_1^2 + \delta r_2^2 = \frac{\delta n_1^2}{t_1^2} + \frac{\delta n_2^2}{t_2^2} = \frac{n_1}{t_1^2} + \frac{n_2}{t_2^2} = \frac{r_1}{t_1} + \frac{r_2}{T - t_1} \quad (12.140)$$

To minimize δr , we differentiate Eq. (12.140) with respect to t_1 , put it equal to 0 and obtain

$$\frac{t_1}{t_2} = \sqrt{\frac{r_1}{r_2}} \quad (12.141)$$

Example 12.2

$$u = x \cdot y \quad \text{or} \quad u = \frac{x}{y}$$

$$\frac{\sigma_u^2}{u^2} = \frac{\sigma_x^2}{x^2} + \frac{\sigma_y^2}{y^2} \quad (12.142)$$

Example 12.3: Asymmetry of Angular Distribution

Asymmetry of a distribution is a frequently measured quantity. Let R be the number of counts in the interval $0 \leq \cos \theta \leq 1$ and L in $-1 \leq \cos \theta \leq 0$. We want to estimate the error of its asymmetry

$$A = \frac{R - L}{R + L} \quad (12.143)$$

$$\frac{\partial A}{\partial R} = \frac{2L}{(R + L)^2}, \quad \frac{\partial A}{\partial L} = \frac{-2R}{(R + L)^2} \quad (12.144)$$

We put $N = R + L$:

$$\sigma_A^2 = \left(\frac{2L}{N^2} \right)^2 \sigma_R^2 + \left(\frac{2R}{N^2} \right)^2 \sigma_L^2 \quad (12.145)$$

As both L and R obey the Poisson distribution $\sigma_R^2 = R$, $\sigma_L^2 = L$

$$\sigma_A^2 = \frac{4L^2 R + 4R^2 L}{N^4} = 4 \frac{RL}{N^3} = \frac{1 - A^2}{N} \quad (12.146)$$

When A is small

$$\sigma_A \simeq \sqrt{\frac{1}{N}} \quad (12.147)$$

The error has to be smaller than the value of the asymmetry itself. Let us define a figure of merit (FOM) by

$$\text{FOM} \equiv \frac{A}{\sigma_A} \quad (12.148)$$

The number of events N to obtain FOM is

$$\text{FOM} = \frac{A}{\sigma_A} = \frac{A}{\sqrt{\frac{1 - A^2}{N}}} \rightarrow N = (\text{FOM})^2 \frac{(1 - A^2)}{A^2} \quad (12.149)$$

Let us further assume one wants at least $\text{FOM} \geq 3$. Then the required number of events is

$$N = \begin{cases} 30\,000 & A = 0.01 \\ 9 & A = 0.5 \end{cases} \quad (12.150)$$

The experimental time to achieve the desired FOM depends greatly on the value of the asymmetry itself.

12.7.3

Least Squares Method

Consider a function $y = f(x; a_1, a_2, \dots, a_m)$ having x as a variable. We want to determine the value of m a_i 's. Assume a measurement was carried out at n points x_i ($i = 1, \dots, n$) and measured values y_i and their errors σ_i have been obtained. The least squares method minimizes the following quantity:

$$S = \sum_{i=1}^n \left[\frac{y_i - f(x_i; a_1, \dots, a_m)}{\sigma_i} \right]^2 \quad (12.151)$$

In order to obtain a meaningful result the number n must be larger than the number of parameters to be determined. If the distribution is Gaussian with the mean $f(x_i; a_1, \dots, a_m)$, it is easy to prove the result agrees with that given by the maximum likelihood method. In this case S is called the chi square (χ^2) distribution. Consequently, it is often said that the least squares method minimizes the χ^2 distribution. If the y_i distribution is not Gaussian, the statement is not exact, but is not far from the truth because of the central limit theorem. Note we have assumed no errors in determining x_i in Eq. (12.151). If errors have to be assigned,

$$\sigma_y^2 \rightarrow \sigma_y^2 + \sigma_x^2 \left(\frac{\partial f}{\partial x} \right)^2 \quad (12.152)$$

has to be used instead of σ_y^2 . To obtain the best fit, S has to be a minimum and the conditions

$$\frac{\partial S}{\partial a_j} = 0 \quad (12.153)$$

give associated equations to be solved. When analytical solutions are not available, numerical calculations using computers have to be carried out. With similar arguments as before, the error matrix U of a_j is given by

$$U^{-1} = \frac{1}{2} \frac{\partial^2 S}{\partial a_i \partial a_j} \quad (12.154)$$

If S is nonlinear, computer calculation is the only way. In this case, the error matrix has to be solved numerically, too. A way to determine the variance is again guided by relations of the Gaussian distribution. When the distribution is one-dimensional Gaussian, we have

$$\sigma^2(\hat{a}) = \left(\frac{\partial^2 S}{\partial a^2} \right)_{a=\hat{a}}^{-1} \quad (12.155)$$

Taylor expanding S around the minimum, we obtain

$$S(a) \simeq S(\hat{a}) + \frac{1}{2}(a - \hat{a})^2 \frac{\partial^2 S}{\partial a^2} = S(\hat{a}) + \frac{1}{\sigma^2}(a - \hat{a})^2 \quad (12.156)$$

Therefore, we have the equality

$$S(\hat{a} + \sigma) = S(\hat{a}) + 1 \quad (12.157)$$

Generalizing the relation, we may define the standard error σ as the deviation of a from $\hat{a}$ to give S the value $S_{\min} + 1$. Equation (12.155) also shows how to obtain σ numerically. When there are two parameters to be determined, a closed contour where $S = S_{\min} + 1$ gives an error region within 1σ in the a_1 - a_2 plane.

χ^2 Distribution The chi square (χ^2) distribution is a theoretical probability distribution in inferential statistics, e.g., in statistical significance tests. The χ^2 distribution with n degrees of freedom using x ($x \geq 0$) as its variable is given by

$$K_n(x) = \frac{x^{n/2-1} e^{-x/2}}{2^{n/2} \Gamma(n/2)} \quad (12.158)$$

where $\Gamma(x)$ is the Gamma function. The function's form is given in Fig. 12.47.

Suppose the distribution function of a certain variable μ is given by $F(\mu, a_i)$, where a_i are m parameters one wants to determine by an experiment. Let y_i be measured points of μ obeying the normal distribution $G(\mu_i, \sigma^2)$. Then χ^2 defined by

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - \mu)^2}{\sigma^2} \quad (12.159)$$

obeys Eq. (12.158). The χ^2 distribution is referred to as the goodness of fit and is a measure of how well the determined parameters a_i reproduce the measured values. The measured n values $y_i - \mu$, ($i = 1, \dots, n$) are independent variables, but if the average $\bar{y}$ is substituted for μ , $y_i - \bar{y}$ are $n - 1$ independent parameters and $n - 1$ should be used as the degrees of freedom in applying the χ square test. When there are m parameters to be determined, the number of degrees of freedom becomes $n - 1 - m$. The confidence level CL is given by

$$\text{CL}(S) = \int_S^\infty dx K_n(x) \quad (12.160)$$

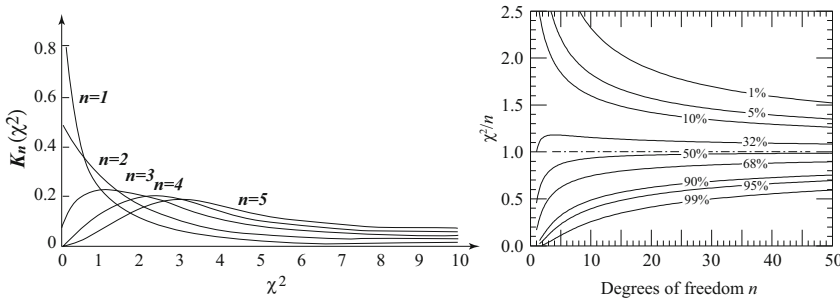


Figure 12.47 χ^2 distribution: $K_n(x) = (x^{n/2-1} e^{-x/2}) / (2^{n/2} \Gamma(n/2))$ and the reduced χ^2 is equal to χ^2/n for n degrees of freedom. The curves are given for given values of CL.

Usually, taking $CL = 5\%$ and when $\chi^2 \leq S$, the fit is approved as good. As a rough estimate, if μ is considered as a good representative, $(y_i - \mu)$ should take the approximate value σ^2 and so $\chi^2/(n - 1) \sim 1$ is expected. If the value of χ^2 is too far away, the original theoretical assumption has to be reconsidered.

Part Two A Way to the Standard Model

13

Spectroscopy

Spectroscopy originally meant the use of a prism to disperse visible light and analyze a complex wave form as a function of its wavelengths. The concept was expanded to include any measurement of observables as a function of wavelength or oscillation frequency. As the frequency is equivalent to energy in quantum mechanics, collecting and sorting of energy levels in any phenomenon is also called spectroscopy. It is an indispensable procedure in science to elucidate the underlying basic rules from observation of seemingly complex phenomena. The role of the periodic table in understanding atomic structure and radiation spectra in clarifying the nature of quantum mechanics come to mind immediately. In a similar vein, collecting the masses of hadrons, including both elementary particles and resonances, which were thought of as dynamic processes of interacting elementary particles, and determining their spin, parity and other quantum numbers were essential procedures in reaching the quark structure of the hadrons. Later on, we will deal with level structures of quark composites, but here we follow history to learn a heuristic way to reach the truth. The first step is to collect facts. The second step is to sort them out, find a regularity and see if they can be reproduced by a simple model. The third step is to elucidate the underlying law that governs the dynamics and reach the truth.¹⁾

Discoveries in the pre-accelerator era were made using isotopes and cosmic rays. That was the age of hunting particles, exciting but very much dependent on luck. With the advent of large accelerators – Cosmotron (3.3 GeV, 1953–1968), Bevatron (6.2 GeV, 1954–1971), PS (28 GeV, 1959–) and AGS (33 GeV, 1960–) – the age of harvesting particles had arrived. It became possible to plan and prepare experiments in an organized way, controlling the particle's conditions.

In this chapter, we describe how to extract information on hadrons (π , nucleon, K , ρ , ω , etc.) and resonances and determine their properties. This is the first step described above. We start from the basic structure of the interactions between pions and nucleons, which are representatives of the meson and baryon families of the hadron. Analysis of hadron interactions exemplified by π - N scattering is essential in order to understand their dynamics. By analyzing their behavior we can

1) This is Sakata's syllogism that prevailed in his school, whose pupils proposed IOO symmetry [212] (flavor $SU(3)$ of the Sakata model [339]), a predecessor of the quark model, lepton–baryon symmetry, the fourth quark [271] and the Kobayashi–Maskawa model [239].

extract resonances and determine their spin-parity as well as other features. Classification of resonances and construction of a model are the second step, which will be discussed in the next chapter. A brief description of low-energy nuclear forces and the high-energy behavior of hadronic cross sections is given in the last part of this chapter.

13.1

Pre-accelerator Age (1897–1947)

The history of particle physics as a research discipline may be said to have begun with Fermi's (1934) and Yukawa's (1935) proposal of the weak and strong (meson) theories. However, the first discovery of an elementary particle as we know it today dates back to as early as 1897, when J.J. Thomson discovered the electron, although the early history of the particle is more closely related to atomic structure and the quantum concept of particle-wave duality. The photon was recognized as a particle following Einstein's light quantum hypothesis in 1905 and subsequent demonstration of Compton scattering in 1925. The discovery of the neutron in 1935 showed that the nucleus consists of particles. It triggered Fermi's and Yukawa's applications of field theory to the weak and strong forces, incorporating particle creation and annihilation. In their theories, new particles (the neutrino and the meson) are claimed as indispensable players in the dynamical phenomena. Actually, the neutrino was proposed by Pauli in 1930 to reconcile apparent violations of energy and angular momentum conservation in beta decays. Fermi included it and applied the quantized field theory invented by Heisenberg and Pauli to beta decay. Yukawa's treatment in his first paper was semiclassical but his idea of a quantum of a force carrier was revolutionary and remains one of the prime axioms in modern particle theory.

It was the beginning of new era in which people no longer hesitated to claim new particles to explain newly discovered phenomena. Until then, the change of the concept of matter aroused by the development of quantum mechanics had been so violent that people were more willing to abandon fundamental laws than to introduce dubious new particles. For instance, Bohr, who represented mainstream thinking at the time, opted to discard energy conservation in beta decay rather than introduce unknown new particles. That he did not like Yukawa's idea either was indicated in his comment "You like new particles so much" when Yukawa presented the new theory to him on his visit to Japan in 1937.

Electron The first particle recognized as an elementary particle was the electron. It was known for a while as the cathode ray, which is emitted from a hot filament in a vacuum (low pressure to be exact) glass tube when high voltage is applied between an anode plate and the cathode. In fact, Hertz, who proved the existence of the electromagnetic wave experimentally, knew it. He erroneously concluded that the cathode ray was not charged. It was only after vacuum technology improved that the

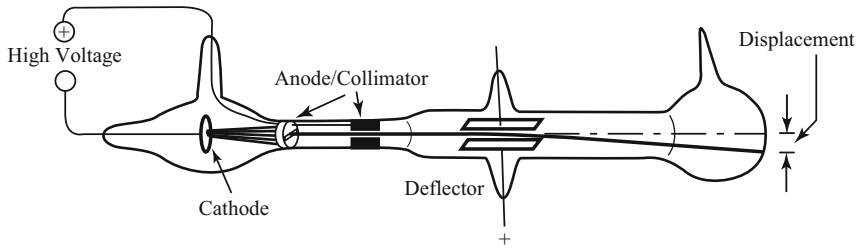


Figure 13.1 The apparatus with which J.J. Thomson discovered the electron. A high voltage is applied between the cathode and the anode plate. The emitted cathode ray goes through a hole in the anode, is collimated by

a slit, hits the glass wall and produces a fluorescent point. When a transverse electric field is applied across two plates in the middle (deflector), the cathode ray can be deflected [366].

deflection of the cathode ray was detected by J.J. Thomson in 1897 in the apparatus shown in Fig. 13.1 [366].

The cathode ray, after going through collimators, is deflected by a transverse electric field. From the direction of the deflection the charge can be determined. The beam can also be deflected by a magnetic field created by two coils applied perpendicularly to both the beam and the electric field. By adjusting the field strength to compensate the deflection by the electric field, the velocity of the particles can be measured. The value of e/m can be determined by the deflection angle created by the electric field. Its value was a thousand times smaller than that of the hydrogen ion (the proton) measured in electrolysis. Thus Thomson concluded that the cathode ray consisted of negative particles with mass much smaller than any known matter. An accurate value of the electric charge was obtained by Millikan's oil drop experiment and the mass was determined.

Atomic Structure Thomson proposed an atomic model that consists of a positive charge cloud of size $\sim 10^{-8}$ cm in which electrons are embedded like the seeds of a watermelon. The other model to be contrasted with Thomson's was Nagaoka's saturnian model [284], in which electrons circulate around a positive nucleus. The models could be tested by making an α -particle scatter on matter. The α -particle was discovered by Becquerel, and by 1911 it was known to have charge $+2e$ and a mass four times that of hydrogen. Since the α -particle is much heavier than the electron, it would not be deflected by electrons. Deflection by the positive charge would also be weak in the Thomson model because of its extended distribution. Rutherford noticed that these characteristics of the Thomson model did not agree with observations. He further conjectured that if the Thomson model were right, large-angle scattering is possible only as an accumulation of small scattering angles and the distribution would be Gaussian. On the other hand, if the charge is concentrated in a small region and the target mass is large, when the α -particle hits the atom head on it will be repelled back. If the collision is not head-on but within a small transverse distance (impact parameter), the deflection would still be large. Figure 13.2 shows trajectories (magnified 200 billion times) that an incoming par-

ticle makes for different impact parameters under the influence of the repulsive Coulomb force.

Following Rutherford's suggestion, Geiger and Marsden carried out an experiment using the detector shown in Fig. 13.3 and obtained the angular distribution shown in Fig. 13.2, which agreed with Rutherford's formula given in Eq. (6.46). The formula was calculated classically assuming the scattering pattern in Fig. 13.2 and

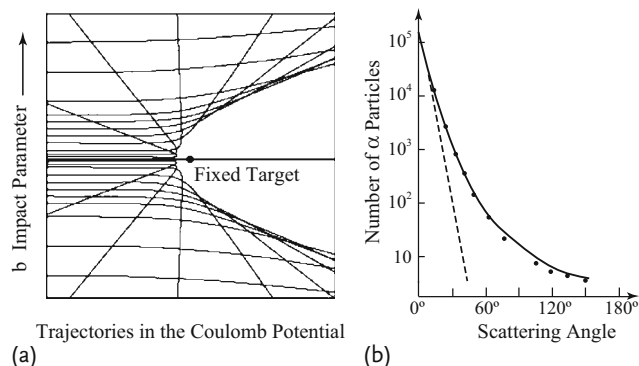


Figure 13.2 (a) Trajectories of incoming particles with different impact parameters scattered by the Coulomb force of the fixed scattering center. Consider it as a 2×10^{11} times magnified picture of an α -particle scattered by an atom. When the impact parameter is small,

the α -particle is deflected through a large angle. (b) Data fit well with the Rutherford scattering formula. The dashed line indicates the predicted scattering distribution for the Thomson model [158].

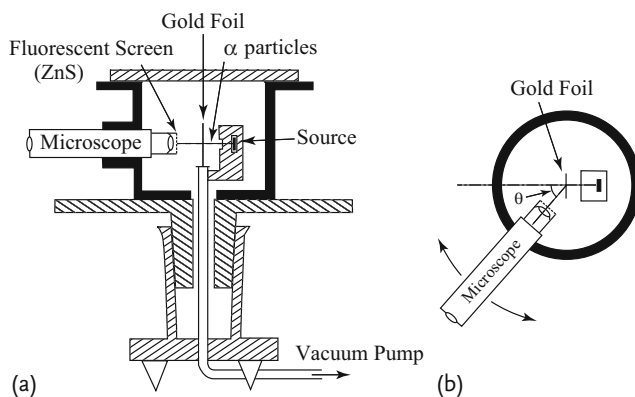


Figure 13.3 Side (a) and plan (b) view of Geiger and Marsden's apparatus: an α source, scattering gold foil, and a microscope whose front end is coated with fluorescent material (ZnS) are contained in a metal cylinder. The source is concentrated radon and α -particles are incident on the foil at a right angle through a slit. The source and the foil

are fixed to the axis, but the cylinder together with the microscope can be rotated around the axis so that the scattering angle of the α -particles can be varied. The measurement was carried out by flickers of the fluorescence being observed by the human eye and counted [158].

an uniform impact parameter distribution. Rutherford was fortunate because it is a rare example of the classical treatment producing the same result as quantum mechanics.

The expected curve according to the Thomson model is depicted as a dashed line. As we discuss later in Chap. 17 or more simply according to the Heisenberg uncertainty principle, a measure of the minimum distance that the projectile can penetrate into the target is given by the de Broglie wavelength of its momentum transfer $Q = 2p \sin(\theta/2) \lesssim p$. Since the data agreed with a formula that assumes a point target, the size of the nucleus is at most the inverse of the α -particle momentum. Assuming kinetic energy of 4 MeV, which gives $p \sim 170 \text{ MeV}/c$, one can estimate the nuclear size to be at most $r_0 \sim \hbar c / 170 \text{ MeV}/c \sim 10^{-13} \text{ cm}$, which is 10^5 times smaller than an atom. This experiment unambiguously established the fact that the atom has a structure with positive charge concentrated in a region $r < 10^{-13} \text{ cm}$ and a surrounding electron cloud extending to 10^{-8} cm .

Neutron In 1932, Irène and Frédéric Joliot-Curie noticed that neutral particles with strong penetrating power are emitted when α -particles from polonium collide with a beryllium target. Unfortunately, they mistook them for energetic gamma rays. Chadwick, knowing Rutherford's idea of the possible existence of a neutral heavy particle, a possible composite of the proton and the electron, proved in 1932 that the mass of the neutral particle of the Joliot-Curies was as large as that of the proton. His apparatus is shown in Fig. 13.4. The mass measurement was done as follows. In Newtonian mechanics, if a ball having mass m collides with another ball with the same mass at rest, the incident ball stops and the target moves on with the same velocity as that of the incident ball. If the target ball has mass $14m$, the velocity that the target obtains is $2/15$ that of the incident ball. These numbers can be obtained by a simple calculation using the law of conservation of momentum. Placing a paraffin foil which contains hydrogens or nitrogen ($m_N = 14m_H$) gas chamber, the maximum velocity of recoiled hydrogens or nitrogen are those ob-

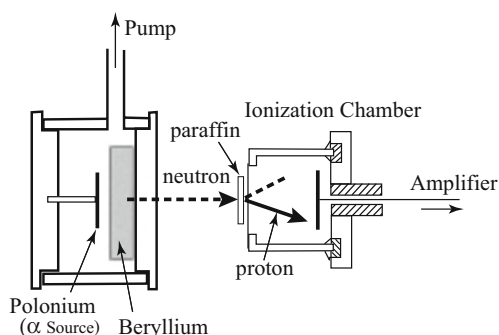


Figure 13.4 Chadwick's apparatus for detecting the neutron. The neutron is knocked out from Be by the α -particle emitted from polonium. Various materials are placed in front

of the ionization chamber on the right, which measures the ionization loss of charged particles scattered by the neutron [94].

tained from head-on collisions. The velocity can be determined from the ionization loss of the particles, which is converted to electric current, amplified and automatically recorded. It was the most advanced instrument at the time. The ratio of the velocity as determined by ionization loss between the paraffin and the nitrogen gas foil was $2/15$, confirming that the mass of the neutron was equal to that of the proton.

Positron Anderson's discovery of the positron (1933) has already been described in Chap. 4 in explaining the concept of the antiparticle. We only mention that the discovery was made by using a cloud chamber, which was invented by Wilson in 1911. It played an important role also in discovering the muon and pion, as well as strange particles later. It is a sealed box containing supercooled, supersaturated water or alcohol vapor. When a charged particle goes through the chamber, it ionizes the matter along its path. The ions act as condensation nuclei, around which a mist will form, creating a track just like an airplane contrail. The velocity and the momentum can be measured by the particle's ionization losses and the curvature of the track in the magnetic field.

Muon and Pion The discovery by Anderson and Neddermeyer in 1937 [293] of a particle with mass $\sim 200 m_e$ in cosmic rays aroused a considerable amount of excitement in view of Yukawa's prediction of the existence of a nuclear force carrier. However, the particle had a long life, interacting weakly with matter, which cast doubt on its role as the strong force carrier. It took a long and confusing time before the issue was finally settled by Powell's demonstration of sequential decay ($\pi \rightarrow \mu \rightarrow e$) on emulsion photographs (Fig. 13.5), thus establishing the pion as the strong force carrier predicted by Yukawa. As to the muon, however, although very instrumental in clarifying the properties of the weak interaction, its very existence, i.e. its role, remains a mystery to this day.

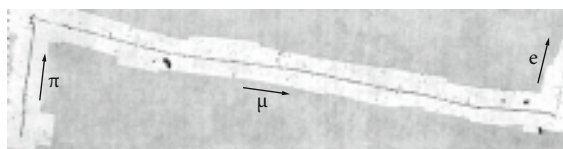


Figure 13.5 Emulsion photograph of $\pi - \mu - e$ decay. A pion decays into a muon ($\pi \rightarrow \mu + \nu_\mu$) and the muon decays into an electron ($\mu \rightarrow e + \nu_e + \nu_\mu$) [80].

Muon Neutrino The existence of the second neutrino was inferred from the muon decay in $\mu \rightarrow e + \nu_e + \nu_\mu$, because the energy spectrum of the produced electron was not monochromatic but showed a continuous spectrum. That ν_μ might be different from ν_e , which appears in beta decay, was suspected from the absence of $\mu \rightarrow e + \gamma$, which will be treated in more detail in the discussion of the weak interaction in Chap. 15.

Strange Particles Strange particles were also discovered in cloud chambers exposed to cosmic rays, but their detailed properties were investigated using accelerators. The pursuit of their properties and role in particle physics led to the discovery of the quark model, which will be explained in the next chapter.

13.2

Pions

Pions, nucleons and their interactions are the basis and also the best studied subjects of all hadrons. This is because the pion is the lightest hadron and an intense pion beam can be easily made by bombarding protons on matter. We start by determining the spin-parity of the pion.

Spin of π^+ To determine the spin of π^+ , the principle of detailed balance was applied to the reaction



Denoting the spin of π , D , p as s , s_D and s_p , respectively, and the momentum of π , p in their system as k , p , Eq. (9.146) is written as

$$\frac{d\sigma(p p \rightarrow \pi D)}{\sigma(\pi D \rightarrow p p)} = 2 \frac{k^2}{p^2} \frac{(2s+1)(2s_D+1)}{(2s_p+1)^2} = \frac{3}{2} \frac{k^2}{p^2} (2s+1) \quad (13.2)$$

The factor 2 in front of the equation is a correction factor for two identical particles in the final pp system. We used the fact that $s_D = 1$, $s_p = 1/2$ to derive the last equation. There is a factor 3 difference in the cross section when the spin is 1 compared to $s = 0$ and it is easy to distinguish them experimentally.

Knowing $\sigma(pp \rightarrow \pi D) = 4.5 \pm 0.8$ mb at $T_p = 340$ MeV or $T_\pi = 21.4$ MeV [91] gives the predictions $\sigma(\pi D \rightarrow pp) = 3.0 \pm 1.0$ mb if $s = 0$ and 1.0 ± 0.3 mb if $s = 1$. The measurement showed that $\sigma = 3.1 \pm 0.3$ mb at $T_\pi = 25$ MeV [100, 126]. The spin of the pion was determined to be 0.

Parity of π^- Let us consider a process in which a slow pion stops in deuterium (D), is absorbed by it and then dissociates into two neutrons:



The stopping π is trapped by the Coulomb force to fall in an orbit of the deuterium, drops down to the S-state and reacts. Therefore the orbital angular momentum is 0 before the reaction. Assuming the orbital angular momentum in the final state as L and using the parity conservation law [Eq. (9.73)]:

$$P_\pi = P_D(P_n)^2(-1)^L \quad (13.4)$$

Deuterium is a bound state of the proton p and the neutron n and is known to have spin 1 mostly in the S ($\ell = 0$) state with a small mixture of D ($\ell = 2$) state

$P_D = P_p P_n (-1)^\ell$. As we discussed before, if $P_p P_n = 1$ is assumed, the above equation leads to

$$P_\pi = (-1)^L \quad (13.5)$$

To determine L , we use the angular momentum conservation. As $s_n = 1/2$, the total spin of the two neutron system is either 0 or 1. Taking into account $s_\pi = 0$, the initial value of the angular momentum is $J_i = s_D = 1$. The final value is a vector sum of s and L . Then

$$\begin{aligned} s = 0 & \rightarrow J_f = J_i = 1 & \rightarrow L = 1 \\ s = 1 & \rightarrow J_f = L - 1, L, L + 1 & \rightarrow L = 2, 1, 0 \end{aligned} \quad (13.6)$$

We need one more condition to determine the parity. We note that the final state is a two-fermion state and has to be antisymmetric under permutation. Since the wave function is a product of both the spin and spatial wave functions, if the spin wave function is symmetric then the spatial wave function must be antisymmetric and vice versa. As the $s = 0$ (1) state is antisymmetric (symmetric), L has to be even (odd) correspondingly:

$$\begin{aligned} s = 0 & \rightarrow L \neq 1 \\ s = 1 & \rightarrow L = \text{odd} \end{aligned} \quad (13.7)$$

The only allowed value is $L = 1$, therefore

$$P_\pi = -1 \quad (13.8)$$

This means if the reaction (13.3) exists, the pion's parity has to be negative. If the parity is positive the reaction does not happen. The reaction has been observed and the parity of the pion (π^- to be exact) is determined to be negative.

Spin and Parity of π^0 The neutral pion decays into two photons with nearly 100% probability. This fact alone excludes the pion's spin from being 1. In the π^0 rest frame, the two γ 's are emitted back to back. Taking the z -axis along the photon momentum, and writing RL when the photon moving in the z direction is right handed ($s_z = 1$) and the other moving in the $-z$ direction is left handed ($s_z = -1$), a total of four states (RL, LR, RR, LL) are conceivable (Fig. 13.6). As $L_z \sim (\mathbf{x} \times \mathbf{p}) \cdot \mathbf{p} = 0$, J_z includes only spin components. RL, LR states, which have $J_z = \pm 2$, cannot be realized unless $J \geq 2$. J_z of RR, LL states is 0, which does not rule out having $J = 1$. But if $J = 1$, the wave function should behave as the eigenfunction of $(J, J_z) = (1, 0)$, namely $Y_{10}(\theta, \phi) \sim \cos \theta$, which gives -1 at $\theta = \pi$. But 180° rotation of RR or LL is equivalent to particle exchange of 1 and 2, which has to give $+1$ by their bosonic nature. This is a contradiction, hence $J = 1$ is not allowed.

Next we assume the spin of π^0 is zero and derive its parity. The parity transformation changes the direction of the momentum but not that of the spin. Therefore R and L are exchanged, which, in turn, means RR and LL are exchanged. Then the parity eigenstate must be of the form

$$|P\rangle = \frac{RR \pm LL}{\sqrt{2}} \leftrightarrow P = \pm 1 \quad (13.9)$$

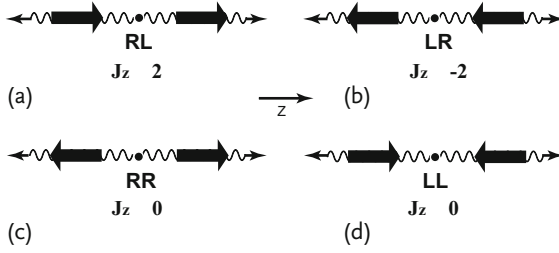


Figure 13.6 Polarization states of two photons in π^0 CM. The wavy lines denote the photon momentum and the bold arrows denote their spin orientation. (a) RL , (b) LR have $J_z = \pm 2$ while (c) RR , (d) LL have $J_z = 0$.

Writing the photon polarization vectors as e_x, e_y , we can express the right-handed and left-handed states as

$$|R\rangle = -\frac{e_x + ie_y}{\sqrt{2}}, \quad |L\rangle = \frac{e_x - ie_y}{\sqrt{2}} \quad (13.10)$$

Then

$$P = \pm 1 \leftrightarrow RR \pm LL \sim \begin{cases} e_x e_x - e_y e_y \\ e_x e_y + e_y e_x \end{cases} \quad (13.11)$$

This means that if $P = +1$, the polarization of the two photons is parallel, while the polarization of the negative parity states is orthogonal. Since the direction of the polarization is that of the electric field, electron-positron pairs produced in $\gamma \rightarrow e^+ + e^-$ tend to be emitted in the direction of the electric field. Among the π^0 decay reactions, there is a mode called Dalitz decay

$$\pi^0 \rightarrow (e^+ + e^-) + (e^+ + e^-) \quad (13.12)$$

which can happen via inner bremsstrahlung. Therefore an experiment to measure the angular distribution of the two electron pairs can reveal the parity of π^0 . A theoretical calculation gives $1 \pm A \cos \theta$, where θ is the angle between the two planes that the pairs make and $A \simeq 0.48$. Measurement gives $P_\pi = -1$.

Tensor Method It is instructive to rederive the previous result using another method, the tensor method. It is complementary to the partial wave analysis and in many cases simpler and intuitive. We try to construct the S-matrix, at least part of it, using materials available in the problem. We note first that the S-matrix is a scalar because of rotational invariance. In the decay $\pi^0 \rightarrow 2\gamma$, if π^0 is has spin 0, the wave function has only one degree of freedom and is a scalar. But if π^0 has spin 1, it has three degrees of freedom, hence the S-matrix has to be a scalar product of the pion polarization vector $\boldsymbol{\pi}$ with another vector. The same argument applies to any incoming or outgoing particles. Therefore the amplitude must be linear in $\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2$ and also $\boldsymbol{\pi}$ if the pion is a vector. The materials to construct the S-matrix are

Table 13.1 Tensor analysis of $\pi^0 \rightarrow \gamma\gamma$ decay amplitude.

$J^P(\pi^0)$	Amplitude	Polarization direction
0^+	$(\epsilon_1 \cdot \epsilon_2) f(k^2)$	$\epsilon_1 \parallel \epsilon_2$
0^-	$(\mathbf{k} \cdot \epsilon_1 \times \epsilon_2) f(k^2)$	$\epsilon_1 \perp \epsilon_2$
1^-	$(\boldsymbol{\pi} \cdot \mathbf{k})(\epsilon_1 \cdot \epsilon_2) g(k^2)$	$\epsilon_1 \parallel \epsilon_2$
1^+	$(\boldsymbol{\pi} \cdot \epsilon_1 \times \epsilon_2) g(k^2)$	$\epsilon_1 \perp \epsilon_2$

Note: $\mathbf{k} \times (\epsilon_1 \times \epsilon_2) = \epsilon_1(\mathbf{k} \cdot \epsilon_2) - \epsilon_2(\mathbf{k} \cdot \epsilon_1) = 0$

the momenta $(\mathbf{k}_1, \mathbf{k}_2)$ and polarization vectors (ϵ_1, ϵ_2) of the photon. The problem simplifies if we choose the working frame properly, namely, the pion rest frame and the Coulomb gauge, $\mathbf{k}_1 = -\mathbf{k}_2 = \mathbf{k}$ and $\epsilon_1 \cdot \mathbf{k} = \epsilon_2 \cdot \mathbf{k} = 0$. The amplitudes that can be made of these materials are listed in Table 13.1, where $f(k^2)$, $g(k^2)$ represent the remainder of the amplitude.

Under the parity transformation, $\mathbf{k} \rightarrow -\mathbf{k}$ and the pion wave function also changes its sign depending on its intrinsic parity. Since the final state is composed of two identical bosons, the amplitude has to be symmetric under exchange of the two particles, i.e. $\mathbf{k} \leftrightarrow -\mathbf{k}$ and $\epsilon_1 \leftrightarrow \epsilon_2$. The $1^\pm$ amplitudes do not satisfy this requirement and are forbidden. The remaining choice is between 0^+ and 0^- , which must be determined experimentally.

While the possibility of $s \geq 2$ is not excluded, it is unlikely considering the fact that the three members of the pion (π^+ , π^0 , π^-) have almost the same mass. They constitute an isospin triplet (to be discussed later), and are considered to share the same properties. If $s \geq 2$, the angular distribution of $\pi^0 \rightarrow 2\gamma$ would be anisotropic, which is not observed experimentally. In high-energy reactions, a great number of pions are produced and their multiplicity is expected to be proportional to their degrees of freedom ($2s + 1$). But it is the same for all $\pi^\pm, \pi^0$, so we may assume they have the same spin and parity. Today, pions are considered as composites of u, d quarks and $\bar{u}, \bar{d}$ antiquarks, $\pi^+ \sim -u\bar{d}$, $\pi^0 \sim u\bar{u} - d\bar{d}$, $\pi^- \sim d\bar{u}$. Their total spin is $s = 0$, the relative orbital angular momentum $\ell = 0$ and they constitute $J^P = 0^-$ mesons.

Problem 13.1

Prove the following statements assuming parity invariance.

1. The decay $0^+ \rightarrow 0 + 0$ is allowed, but $0^- \rightarrow 0 + 0$ is not, where the final state is composed of two identical particles. Experimentally, $\eta(548)$ is known to have spin 0 and

$$\frac{\Gamma(\eta \rightarrow \pi^+ \pi^-)}{\Gamma_{\text{total}}} < 1.3 \times 10^{-5}, \quad \frac{\Gamma(\eta \rightarrow \pi^0 \pi^0)}{\Gamma_{\text{total}}} < 4.3 \times 10^{-4} \quad 90\% \text{ CL} \quad (13.13)$$

which tells us that η has $J^P = 0^-$.

2. $0^+ \rightarrow 0^- + 0^- + 0^-$ is forbidden.
3. $0^- \rightarrow 0^+ + \gamma$, $0^- + \gamma$ are forbidden.
4. $0^- \rightarrow 0^- + 0^- + 1^-$ is allowed.
5. $1^- \rightarrow 0^- + 1^-$ is allowed.
6. Let slow antiprotons stop in hydrogen. Prove that the process $\bar{p} + p \rightarrow 2\pi^0$ is forbidden.

G Parity As was shown in Section 9.3.4, the transformation of a particle doublet to an antiparticle doublet is obtained by the successive operation of C and $e^{-i\pi I_2}$, rotation about the axis I_2 by angle π . It is referred to as the G operation:

$$\begin{bmatrix} p \\ n \end{bmatrix} \rightarrow \begin{bmatrix} -\bar{n} \\ \bar{p} \end{bmatrix} = C \begin{bmatrix} -n \\ p \end{bmatrix} = C \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p \\ n \end{bmatrix} \equiv G \begin{bmatrix} p \\ n \end{bmatrix} \quad (13.14)$$

$$G = C e^{-i\pi I_2}$$

The fact that the G -operated state transforms in the same way under isospin rotation means that the G and isospin operations are commutable. Let us operate with G on the pion fields. The C operation interchanges $\pi = (\pi_1 + i\pi_2)/\sqrt{2}$ and $\pi^\dagger = (\pi_1 - i\pi_2)/\sqrt{2}$, which means $\pi_2 \rightarrow -\pi_2$. The charge symmetry operation is a rotation around the y -axis and changes $\pi_1, \pi_3 \rightarrow -\pi_1, -\pi_3$. The net effect of the G transformation is to change the sign of all the pions:

$$G\pi = -\pi \quad (13.15)$$

This means a system of n pions has G parity $(-1)^n$. Unlike C parity, G parity can be defined for a charged system and is more convenient. However, since the G operation includes a rotation in isospin space, the applicability is limited to the strong interaction. In the electromagnetic interaction, G parity is not conserved. Conversely it can be used to differentiate one from the other. Especially useful is an application to the neutral $I_3 = 0$ system.

□ Theorem 13.1

The G parity of a system consisting of a particle q and an antiparticle $\bar{q}$ is given by $(-1)^{S+L+1}$.

Proof: As the isospin mathematics is the same as that of the angular momentum, the eigenvalue of $e^{-i\pi I_2}$ has the same form as that of $e^{i\pi L_y}$. As the system is neutral and has $I_3 = 0$, the eigenfunction is $Y_{I0} \sim P_I$ (the Legendre polynomial of I th order). As $P_I(-x) = (-1)^I P_I(x)$, and rotation by π around the y -axis changes $(\theta \rightarrow \pi - \theta, \phi \rightarrow \pi - \pi)$ or $\cos \theta \rightarrow -\cos \theta$,

$$e^{i\pi I_2} Y_{I0} = (-1)^I \quad (13.16)$$

On the other hand C operation on a neutral system yields $C = (-1)^{L+S}$ [see Eq. (9.177)]. Consequently, G -parity of a neutral system is $G = (-1)^{L+S+I}$. $\square$

Examples of the G Parity Selection Rule

(1) π production in $\bar{p}n$ system: The system has $I = 1$. For slow $\bar{p}$ stopping in deuterium, the reaction takes place in the S ($L = 0$) state. Then an even (odd) number of pions are produced in the decay from $^3S_1$ ($^1S_0$) states.

(2) $\rho \rightarrow 2\pi$:

$$\rho^\pm(1^-) \rightarrow \pi^\pm \pi^0 \quad \therefore \quad G_\rho = + \rightarrow \rho \not\rightarrow 3\pi \quad (13.17a)$$

$$\omega(1^-) \rightarrow \pi^+ \pi^- \pi^0 \quad \therefore \quad G_\omega = - \rightarrow \omega \rightarrow 2\pi \quad (13.17b)$$

Problem 13.2

Main decay modes of the η meson ($m_\eta = 548$ MeV, $\Gamma_\eta = 1.3$ MeV) are

$$\begin{aligned} \eta &\rightarrow \pi^+ + \pi^- + \pi^0 & 22.73 \pm 0.34\% \\ &\rightarrow \pi^0 + \pi^0 + \pi^0 & 32.56 \pm 0.23\% \\ &\rightarrow 2\gamma & 39.31 \pm 0.20\% \\ &\rightarrow \pi^+ + \pi^- + \gamma & 4.60 \pm 0.16\% \end{aligned} \quad (13.18)$$

All the decays are via electromagnetic interaction. Why? Considering C, P, J selection rules, show that J^{PC} of η is 0^{-+} if $J < 2$.

13.3

πN Interaction

13.3.1

Isospin Conservation

Charge Independence of πN Coupling As the pion is a spin 0 scalar in real space, the simplest Lorentz invariant interaction can be expressed as

$$\mathcal{H}_{\text{INT}} \sim ig\bar{\psi}\gamma^5\psi\pi^2 \quad (13.19)$$

In isospin space, an obvious invariant that can be made from the two vectors $\bar{\psi}\boldsymbol{\tau}\psi$ and $\boldsymbol{\pi}$ is $\bar{\psi}\boldsymbol{\tau} \cdot \boldsymbol{\pi}\psi$. The interaction Hamiltonian is now

$$\begin{aligned} \mathcal{H}_{\text{INT}} &= ig\bar{\psi}\gamma^5\boldsymbol{\tau} \cdot \boldsymbol{\pi}\psi \\ &= ig[(\bar{n}p + \bar{p}n)\pi_1 + i(\bar{n}p - \bar{p}n)\pi_2 + (\bar{p}p - \bar{n}n)\pi_3] \\ &= ig[\sqrt{2}(\bar{p}n\pi^+ + \bar{n}p\pi^-) + (\bar{p}p - \bar{n}n)\pi^0] \end{aligned} \quad (13.20a)$$

2) $(f/m_\pi)\bar{\psi}\gamma^\mu\gamma^5\psi\partial_\mu\pi$ is equally valid, but in the low-energy limit it is equivalent to the pseudoscalar interaction of Eq. (13.19) with $f = g(m_\pi/2m_p)$ [refer to the low-energy approximation of the Dirac bilinears in Eq. (4.2)].

$$\pi^- = \frac{\pi_1 + i\pi_2}{\sqrt{2}}, \quad \pi^+ = \frac{\pi_1 - i\pi_2}{\sqrt{2}}, \quad \pi^0 = \pi_3 \quad (13.20b)$$

where in the second and third lines we have omitted the space-time variables as they play no role in the following discussion. The last line of Eq. (13.20a) means that the strengths of the interactions

$$p \leftrightarrow n + \pi^+ : n \leftrightarrow p + \pi^- : p \leftrightarrow p + \pi^0 : n \leftrightarrow n + \pi^0 \quad (13.21)$$

are proportional to

$$\sqrt{2} : \sqrt{2} : 1 : -1 \quad (13.22)$$

The charge independence condition Eq. (9.193) $V_{pp} = V_{nn} = V_{pn}$ is derived from the above relation. The argument goes like this. The strength of the potential between nucleons is expressed in Fig. 13.7.

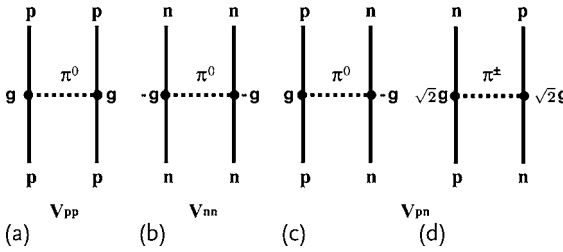


Figure 13.7 The strengths of (a) V_{pp} , (b) V_{nn} and (c,d) V_{pn} are the same.

Fig. 13.7a	Exchange of π^0	$V_{pp} \propto g^2[(1) \times (1)] = g^2$
Fig. 13.7b	Exchange of π^0	$V_{nn} \propto g^2[(-1) \times (-1)] = g^2$
Fig. 13.7c	Exchange of π^0	$V_{pn} \propto g^2[(1) \times (-1)] = -g^2$
Fig. 13.7d	Exchange of $\pi^\pm$	$V_{pn} \propto g^2[(\sqrt{2}) \times (\sqrt{2})] = 2g^2$

Since pp and nn are both $I = 1$ states and symmetric with exchange of the particles 1 and 2, the $I = 1$ part of the potential becomes

$$V_{pn}(c) + V_{pn}(d) = g^2 = V_{pp} = V_{nn} \quad (13.23)$$

which shows the charge independence of the nuclear force.

Isospin Exchange Force In Section 6.8 we saw that when two distinguishable fermions scatter by exchanging a scalar meson the scattering matrix is given by Eq. (6.157)

$$T_{fi} = (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \times \bar{u}(p_3) \Gamma u(p_1) \frac{g^2}{q^2 - \mu^2} \bar{u}(p_4) \Gamma u(p_2) \Big|_{q=p_3-p_1} \quad (13.24)$$

where $\Gamma = 1$ ($i\gamma^5$) for a (pseudo)scalar meson exchange. In the nonrelativistic limit, the scalar meson propagator becomes $-g^2/(q^2 + \mu^2)$, which is the Fourier transform of an attractive Yukawa potential. We now extend the discussion to include the isospin degrees of freedom. The interaction including the isospin is given in Eq. (13.20):

$$\mathcal{H}_{\text{INT}} = g \bar{\psi} \Gamma \boldsymbol{\tau} \cdot \boldsymbol{\pi} \psi = g \sum_i \bar{\psi} \Gamma \tau_i \psi \pi_i \quad (13.25)$$

where ψ is now a 2×1 matrix in the isospin space. Then Eq. (13.24) is modified to

$$T_{fi} \sim \bar{u}(p_3) \Gamma \tau_{1,i} u(p_1) \frac{g^2 \delta_{ij}}{q^2 - \mu^2} \bar{u}(p_4) \Gamma \tau_{2,j} u(p_2) \Big|_{q=p_3-p_1} \quad (13.26)$$

where we have omitted the δ -function. We see the only modification due to inclusion of the isospin is

$$\frac{g^2}{q^2 - \mu^2} \rightarrow \frac{g^2}{q^2 - \mu^2} (\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2) \quad (13.27)$$

where we have omitted the isospin wave function and expressed the scattering matrix as a 2×2 matrix in isospin space. The subscripts 1, 2 were attached to denote that the τ matrix acts only on particle 1 or 2.

Now, a system of two identical fermions must be odd under permutation of particles 1 and 2. Since we have included the isospin degree of freedom, regarding the nucleon (p, n) as identical but in different states, the permutation must be done on all three kinds of wave functions:

$$(1 \leftrightarrow 2) \sim (\text{isospin}) \times (\text{spin}) \times (\text{spatial wave function}) \quad (13.28)$$

Since the two nucleons both have $I = 1/2$, the total isospin is either 1 or 0 and symmetric or antisymmetric accordingly. The isospin wave functions are given by

$$I = 1 = \begin{cases} |pp\rangle & I_3 = +1 \\ \frac{1}{\sqrt{2}}(|pn\rangle + |np\rangle) & I_3 = 0 \\ |nn\rangle & I_3 = -1 \end{cases} \quad (13.29a)$$

$$I = I_3 = 0 \quad \frac{1}{\sqrt{2}}(|pn\rangle - |np\rangle) \quad (13.29b)$$

The value of $\langle f | \boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2 | i \rangle$ is

$$\begin{aligned} \frac{\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2}{2} &= \left(\frac{\boldsymbol{\tau}_1}{2} + \frac{\boldsymbol{\tau}_2}{2} \right)^2 - \left(\frac{\boldsymbol{\tau}_1}{2} \right)^2 - \left(\frac{\boldsymbol{\tau}_2}{2} \right)^2 = I(I+1) - 2 \times \frac{1}{2} \left(\frac{1}{2} + 1 \right) \\ &= \begin{cases} \frac{1}{2} & I = 1 \\ -\frac{3}{2} & I = 0 \end{cases} \end{aligned} \quad (13.30)$$

The deuteron, a bound state of (pn), has $I = 0$, $S = 1$ and only the even orbital angular momentum $L = 0, 2, \dots$ states are allowed, which is indeed the case. It has a dominant S ($L = 0$) state with a small mixture of D ($L = 2$) state.

For $I = 0$, $\tau_1 \cdot \tau_2 < 0$ and the sign of the potential is reversed. Namely, the scalar meson exchange gives a repulsive force and cannot make a bound state if the isospin degree of freedom is taken into account. However, if the exchanged meson is a pseudoscalar, Eq. (13.26) gives an attractive potential and can explain the existence of a nuclear bound state. Since it takes some extra algebra to show that the pseudoscalar exchange indeed gives an attractive force, explicit calculations will be deferred to Section 13.6.

The appearance of the isospin exchange force ($\sim \tau_1 \cdot \tau_2$) is due to the exchange of isospin carried by the pion. Whenever the force carrier has a quantum number that is described by $SU(N)$, $N \geq 2$, a similar exchange force term appears. The color degree of freedom carried by the gluon induces a color exchange force in $SU(3)$, and it played a crucial role in disentangling the hadron mass spectrum and in proving the color charge to be the source of the strong force, which will be discussed in detail in Section 14.6.

33 Resonance ($\Delta(1232)$) In the early days before the discovery of quarks, hadrons, including pions and nucleons, were considered elementary and their dynamics was extensively studied through their reactions. Figure 13.8 shows one example of such investigations. One immediately notes a rich structure in their behavior and many resonances have been identified. Today, they are recognized as superpositions of excited energy levels of quark composites. But discovery does not happen in a day. Given this kind of data, it is necessary to isolate levels, identify their properties and classify them to make a Mendeleev periodic table of hadrons. Namely, finding regularities among the many levels is the first step in reaching a deeper understanding of the structure. This is the process called hadron spectroscopy. Before launching a systematic investigation, we pick out the most conspicuous resonance at $M = 1232$, which used to be called the 33 resonance, and apply the algebra of isospin symmetry to show how it works. There are eight combinations in πN scattering, but we pick out three easily accessible reactions in the following:

$$\pi^+ + p \rightarrow \pi^+ + p \quad (13.31a)$$

$$\pi^- + p \rightarrow \pi^- + p \quad (13.31b)$$

$$\pi^- + p \rightarrow \pi^0 + n \quad (13.31c)$$

Expressing only the isospin indices, we write the scattering amplitude as

$$\langle f | T | i \rangle = \sum_{I, I_3} \sum_{I', I'_3} \langle f | I I_3 \rangle \langle I I_3 | T | I' I'_3 \rangle \langle I' I'_3 | i \rangle \quad (13.32)$$

If charge independence is valid, the total isospin I and I_3 are conserved. The difference of the I_3 is that of the electric charge, and so I_3 conservation follows from electric charge conservation, too. Just like rotational invariance in real space, the scattering amplitude can be decomposed into partial amplitudes of definite isospin.

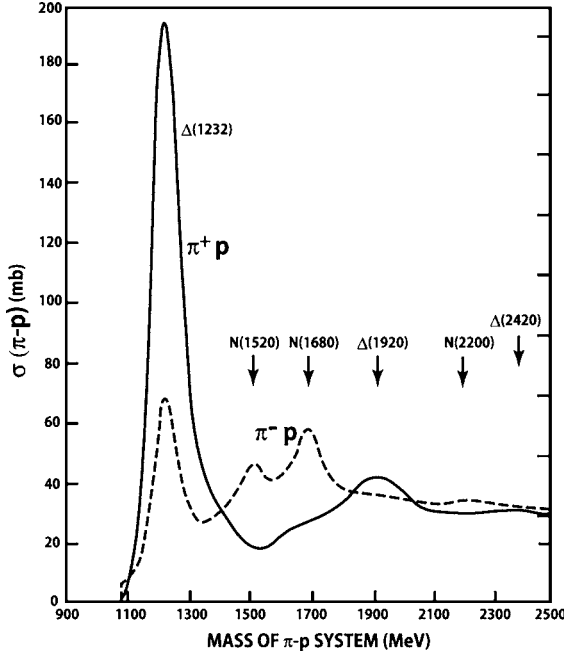


Figure 13.8 Resonance structure of low-energy $\pi^\pm p$ scattering data. The names are N for $l = 1/2$ and Δ for $l = 3/2$ resonances. The number in parenthesis is its mass value. S, P, D, F, G, and H stand for the $L = 0, 1, 2, 3, 4$ and 5 states. D_{15} means $L = 2, l = 1/2, J = 5/2$. Some resonances are overlaps of

two or more states. $N(1680)$ may be an overlap of $N(1675) D_{15}$ and $N(1680) F_{15}$. Similarly, $\Delta(1920)$ may be an overlap of $\Delta(1920) P_{33}$ and $\Delta(1930) D_{35}$. $N(2200)$ may also be an overlap of $(N(2190) G_{17}$ and $N(2220) H_{19})$ [156].

The Wigner–Eckart theorem tells us that

$$\langle I I_3 | T | I' I'_3 \rangle = \delta_{II'} \delta_{I_3 I'_3} \langle I | T | I \rangle \equiv \delta_{II'} \delta_{I_3 I'_3} T(I) \quad (13.33)$$

The transformation matrices $\langle I I_3 | i \rangle$ can be obtained from the Clebsch–Gordan coefficient table (see Appendix E). Let us first rewrite the $|\pi N\rangle$ states in terms of $|I I_3\rangle$:

$$\left| \frac{3}{2}, \frac{3}{2} \right\rangle = |\pi^+ p\rangle \quad (13.34a)$$

$$\left| \frac{3}{2}, \frac{1}{2} \right\rangle = \sqrt{\frac{1}{3}} |\pi^+ n\rangle + \sqrt{\frac{2}{3}} |\pi^0 p\rangle \quad (13.34b)$$

$$\left| \frac{3}{2}, -\frac{1}{2} \right\rangle = \sqrt{\frac{2}{3}} |\pi^0 n\rangle + \sqrt{\frac{1}{3}} |\pi^- p\rangle \quad (13.34c)$$

$$\left| \frac{3}{2}, -\frac{3}{2} \right\rangle = |\pi^- n\rangle \quad (13.34d)$$

$$\left| \frac{1}{2}, \frac{1}{2} \right\rangle = \sqrt{\frac{2}{3}} |\pi^+ n\rangle - \sqrt{\frac{1}{3}} |\pi^0 p\rangle \quad (13.35a)$$

$$\left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \sqrt{\frac{1}{3}} |\pi^0 n\rangle - \sqrt{\frac{2}{3}} |\pi^- p\rangle \quad (13.35b)$$

Solving the above equations in terms of $|I I_3\rangle$, we obtain³⁾

$$|\pi^+ p\rangle = \left| \frac{3}{2} \frac{3}{2} \right\rangle \quad (13.36a)$$

$$|\pi^- p\rangle = \sqrt{\frac{1}{3}} \left| \frac{3}{2}, -\frac{1}{2} \right\rangle - \sqrt{\frac{2}{3}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \quad (13.36b)$$

$$|\pi^0 n\rangle = \sqrt{\frac{2}{3}} \left| \frac{3}{2}, -\frac{1}{2} \right\rangle + \sqrt{\frac{1}{3}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \quad (13.36c)$$

Using the Wigner–Eckart theorem

$$\langle \pi^+ p | T | \pi^+ p \rangle = \left\langle \frac{3}{2} \frac{3}{2} \right| T \left| \frac{3}{2} \frac{3}{2} \right\rangle = T \left(\frac{3}{2} \right) \quad (13.37a)$$

$$\begin{aligned} \langle \pi^- p | T | \pi^- p \rangle &= \left[\sqrt{\frac{1}{3}} \left\langle \frac{3}{2}, -\frac{1}{2} \right| - \sqrt{\frac{2}{3}} \left\langle \frac{1}{2}, -\frac{1}{2} \right| \right] T \\ &\quad \times \left[\sqrt{\frac{1}{3}} \left| \frac{3}{2}, -\frac{1}{2} \right\rangle - \sqrt{\frac{2}{3}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \right] \\ &= \frac{1}{3} \left[T \left(\frac{3}{2} \right) + 2T \left(\frac{3}{2} \right) \right] \end{aligned} \quad (13.37b)$$

$$\begin{aligned} \langle \pi^0 n | T | \pi^- p \rangle &= \left[\sqrt{\frac{2}{3}} \left\langle \frac{3}{2}, -\frac{1}{2} \right| + \sqrt{\frac{1}{3}} \left\langle \frac{1}{2}, -\frac{1}{2} \right| \right] T \\ &\quad \times \left[\sqrt{\frac{1}{3}} \left| \frac{3}{2}, -\frac{1}{2} \right\rangle - \sqrt{\frac{2}{3}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \right] \\ &= \frac{\sqrt{2}}{3} \left[T \left(\frac{3}{2} \right) - T \left(\frac{3}{2} \right) \right] \end{aligned} \quad (13.37c)$$

The three amplitudes (or all eight amplitudes) are written in terms of two independent amplitudes $T(3/2)$ and $T(1/2)$. Combining the three equations in Eq. (13.37),

$$-\langle \pi^+ p | T | \pi^+ p \rangle + \langle \pi^- p | T | \pi^- p \rangle + \sqrt{2} \langle \pi^0 n | T | \pi^- p \rangle = 0 \quad (13.38)$$

The three amplitudes make a triangle in the complex plane. If we express the corresponding cross sections by $d\sigma^+$, $d\sigma^-$, $d\sigma^0$, Eq. (13.38) can be rewritten as

$$|\sqrt{d\sigma^-} - \sqrt{2d\sigma^0}| < \sqrt{d\sigma^+} < \sqrt{d\sigma^-} + \sqrt{2d\sigma^0} \quad (13.39a)$$

3) The Clebsch–Gordan coefficients make an orthogonal matrix. Therefore, reading the table vertically instead of horizontally, the inverse matrix elements can be obtained.



$$|\sqrt{2d\sigma^0} - \sqrt{d\sigma^+}| < \sqrt{d\sigma^-} < \sqrt{2d\sigma^0} + \sqrt{d\sigma^+} \quad (13.39b)$$

$$|\sqrt{d\sigma^+} - \sqrt{d\sigma^-}| < \sqrt{2d\sigma^0} < \sqrt{d\sigma^+} + \sqrt{d\sigma^-} \quad (13.39c)$$

As the cross sections are proportional to the square of the amplitude, they can be expressed as

$$d\sigma^+ = k \left| T \left(\frac{3}{2} \right) \right|^2 \quad (13.40a)$$

$$d\sigma^- = k \left| \frac{1}{3} \left\{ T \left(\frac{3}{2} \right) + 2T \left(\frac{1}{2} \right) \right\} \right|^2 \quad (13.40b)$$

$$d\sigma^0 = k \left| \frac{\sqrt{2}}{3} \left\{ T \left(\frac{3}{2} \right) - T \left(\frac{1}{2} \right) \right\} \right|^2 \quad (13.40c)$$

If $|T(3/2)| \gg |T(1/2)|$

$$d\sigma^+ : d\sigma^- : d\sigma^0 = 9 : 1 : 2 \quad (13.41)$$

If $|T(3/2)| \ll |T(1/2)|$

$$d\sigma^+ : d\sigma^- : d\sigma^0 = 0 : 2 : 1 \quad (13.42)$$

The above relations use only isospin conservation; no other restrictions apply and consequently they are valid for any value of dynamic variables such as the scattering angle or the energy as long as the comparison is made with the same condition. The data of a conspicuous resonance $\Delta(1232)$ at the total energy $\sqrt{s} = 1100 \sim 1300$ MeV or the kinetic energy of the pion $100 \lesssim T_\pi \lesssim 300$ are shown in Fig. 13.9. The measured total cross sections at the peak are

$$210 \text{ mb} : 24 \text{ mb} : 48 \text{ mb} \quad ^4) \quad (13.43)$$

which shows that $\Delta(1232)$ has $I = 3/2$. On the other hand, $\Delta(1232)$ is observed in $\pi^+ p$, $\pi^+ n$, $\pi^- p$, and $\pi^- n$ channels having electric charge $(++, +, 0, -)$, making a quartet, and $I = 3/2$ can be inferred without the above demonstration. But the fact that the cross sections were just as predicted is strong evidence that the hypothesis of charge independence is valid in πN reactions or more generally in the strong interaction.

Other Tests Other evidence is given by the absence of

$$d + d \rightarrow {}^4\text{He} + \pi^0 \quad (13.44)$$

The deuteron and helium are isosinglets ($I = 0$), and the pion an isotriplet ($I = 1$), so the above reaction, if it exists, violates isospin conservation.

4) $1 \text{ mb} = 10^{-27} \text{ cm}^2$.

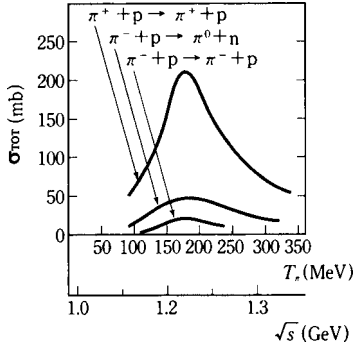


Figure 13.9 Three scattering cross sections in the $\Delta(1232)$ region.

Problem 13.3

Helium ${}^3\text{He}$ and tritium ${}^3\text{H}$ make an isospin doublet. Show that

$$\frac{d\sigma(pd \rightarrow {}^3\text{H}\pi^+)}{d\sigma(pd \rightarrow {}^3\text{He}\pi^0)} = 2 \quad (13.45)$$

Problem 13.4

Show that

$$\frac{d\sigma(pp \rightarrow d\pi^+)}{d\sigma(np \rightarrow d\pi^0)} = 2 \quad (13.46)$$

Problem 13.5

(K^+, K^0) , $(\Sigma^+, \Sigma^0, \Sigma^-)$ make $I = 1/2$ and 1 multiplets. Writing the amplitudes for

$$\pi^+ + p \rightarrow K^+ + \Sigma^+, \quad \pi^- + p \rightarrow K^0 + \Sigma^0, \quad \pi^- + p \rightarrow K^+ + \Sigma^- \quad (13.47)$$

as a^+, a^0, a^- , show that

$$\sqrt{2}a^0 + a^- - a^+ = 0 \quad (13.48)$$

Problem 13.6

(K^+, K^0) , (Ξ^0, Ξ^-) make $I = 1/2$ doublets. Consider the reactions

$$K^- n \rightarrow K^0 \Xi^-, \quad K^- p \rightarrow K^+ \Xi^-, \quad K^- p \rightarrow K^0 \Xi^0 \quad (13.49)$$

and show the relation among the amplitudes

$$f(K^0 \Xi^-) = f(K^+ \Xi^-) + f(K^0 \Xi^0) \quad (13.50)$$

Problem 13.7

Λ^0 is an isosinglet particle. Show that

$$d\sigma(\pi^+ n \rightarrow K^+ \Lambda^0) = d\sigma(\pi^- p \rightarrow K^0 \Lambda^0) \quad (13.51)$$

Fermi–Yang Model If we make an isotriplet from two isodoublets (p, n) and $(-\bar{n}, \bar{p})$ we have

$$|1, 1\rangle = -p\bar{n}, \quad |1, 0\rangle = \frac{1}{\sqrt{2}}(p\bar{p} - n\bar{n}), \quad |1, -1\rangle = n\bar{p} \quad (13.52)$$

These states have baryon number 0, isospin $I = 1$ and the same quantum number as the pions (π^+, π^0, π^-). Therefore the idea that the pions are bound states of a nucleon and an antinucleon was proposed by Fermi and Yang [139]. This was the predecessor of the Sakata model, which extended the $SU(2)$ Fermi–Yang model to $SU(3)$, taking (p, n, Λ) as the three basic constituents of all the hadrons, which in turn developed into the successful quark model.

13.3.2

Partial Wave Analysis

A scattering process can be decomposed into a sum of contributions with definite angular momentum J , which is called partial wave expansion. Even if a resonance is observed at a certain value of J , there are contributions from other partial waves, which may not form resonances. Generally they are slowly varying backgrounds as a function of the energy. If the energy is high and many inelastic channels consisting of more than one π and a nucleon are open, the contribution of the resonance may be only a small fraction and may not appear as a sharp peak. To find resonances in such a case, a partial wave analysis taking into account other contributions is necessary. The starting point is the partial wave expansion formula derived in Chap. 9. We reproduce them here:

$$\frac{d\sigma}{d\Omega} = |f_{\lambda_3 \lambda_4, \lambda_1 \lambda_2}|^2 \quad (13.53a)$$

$$f_{\lambda_3\lambda_4,\lambda_1\lambda_2} = \sqrt{\frac{p_f}{p_i}} \frac{\mathcal{M}_{fi}}{8\pi\sqrt{s}} \quad (13.53b)$$

$$= \frac{1}{2ip_i} \sum_J (2J+1) \langle \lambda_3\lambda_4 | [S^J(E) - 1] | \lambda_1\lambda_2 \rangle d_{\lambda\mu}^J(\theta) e^{i(\lambda-\mu)\phi} \quad (13.53c)$$

$$P |J\lambda\lambda_1\lambda_2\rangle = \eta_1\eta_2(-1)^{J-s_1-s_2} |J\lambda, -\lambda_1, -\lambda_2\rangle \quad (13.55a)$$

$$\langle -\lambda_3, -\lambda_4 | S^J | -\lambda_1, -\lambda_2 \rangle = \eta_S \langle \lambda_3\lambda_4 | S^J | \lambda_1\lambda_2 \rangle \quad (13.55b)$$

$$\eta_S = \eta_1\eta_2\eta_3\eta_4(-1)^{s_3+s_4-s_1-s_2} \quad (13.55c)$$

$$\langle -\lambda_3, -\lambda_4 | f(\theta, \phi) | -\lambda_1, -\lambda_2 \rangle = \eta_S \langle \lambda_3\lambda_4 | f(\theta, \pi - \phi) | \lambda_1\lambda_2 \rangle \quad (13.55d)$$

where all the kinematical variables are to be evaluated in the center of mass frame. P is the parity operator and λ_i, η_i are helicity and intrinsic parity of the particle i respectively. As the nucleon and the pion have spin 1/2 and 0 respectively, we set $\lambda_1 = 0, \lambda_2 = \pm 1/2$. In the following we write $\pm$ instead of $\pm 1/2$ for simplicity. According to Eq. (13.55b), we have

$$S_{++}^J = S_{--}^J, \quad S_{+-}^J = S_{-+}^J \quad (13.56)$$

Denoting the orbital angular momentum as L , we have $L = J \pm 1/2$. Considering that π has intrinsic parity -1 , the parity transformation formula (13.55a) becomes

$$P |JM; \pm\rangle = -(-1)^{J-1/2} |JM; \mp\rangle \quad (13.57)$$

where J, M denote the angular momentum and its component along z-axis. Then, eigenstates of the parity operator can be constructed as

$$\begin{aligned} |J-\rangle &= \frac{|JM; +\rangle + |JM; -\rangle}{\sqrt{2}} \\ |J+\rangle &= \frac{|JM; +\rangle - |JM; -\rangle}{\sqrt{2}} \end{aligned} \quad (13.58)$$

and their parity is given by

$$P |J_{\pm}\rangle = \pm(-1)^{J-1/2} |J_{\pm}\rangle = -(-1)^{J\pm 1/2} |J_{\pm}\rangle \quad (13.59)$$

Equation (13.59) means $|J_{\pm}\rangle$ corresponds to $L = J \pm 1/2$. According to Eq. (13.56), there are only two independent amplitudes for a fixed J . Two L 's with the same J

5) The relation with the scattering matrix is given by Eq. (6.35):

$$S_{fi} = \delta_{fi} - iT_{fi} = \delta_{fi} - i(2\pi)^4 \delta^4\left(\sum_i p_i - \sum_f p_f\right) \mathcal{M}_{fi} \quad (13.54)$$

state differ by 1 and hence correspond to states of different parity. If parity is conserved, they do not mix. Therefore, partial matrices made of states $J_{\mp}$ are diagonalized. From the unitarity of the S -matrix, we have

$$S^{J-} = e^{2i\delta_{J-}}, \quad S^{J+} = e^{2i\delta_{J+}} \quad (13.60)$$

where δ_J is referred to as the phase shift. When inelastic channels are open, the unitarity condition changes. If we are considering only elastic scattering channel, the effect appears as attenuation of the flux. It can be taken into account by considering a complex phase shift that has a positive imaginary part. We define

$$e^{2i\delta_J} = e^{2i(\text{Re}\delta_J + i\text{Im}\delta_J)} = \eta_J e^{2i\text{Re}\delta_J} \quad (13.61)$$

Then, using Eqs. (13.58), we have

$$\begin{aligned} \frac{\langle JM; + | S - 1 | JM; + \rangle}{2ip} &\equiv \frac{S_{++}^J - 1}{2ip} = \frac{e^{2i\delta_-} + e^{2i\delta_+} - 2}{4ip} \\ &\equiv \frac{f_{J-} + f_{J+}}{2} \end{aligned} \quad (13.62a)$$

$$\frac{\langle JM; + | S - 1 | JM; - \rangle}{2ip} \equiv \frac{S_{+-}^J - 1}{2ip} = \frac{f_{J-} - f_{J+}}{2} \quad (13.62b)$$

$$f_{J\pm} = \frac{1}{p} e^{i\delta_{J\pm}} \sin \delta_{J\pm} \quad (13.62c)$$

which defines the partial scattering amplitudes. Inserting Eqs. (13.62) in Eq. (13.53c) and rewriting $d_{\lambda\mu}^j$ using formula in Eq. (E.22), we obtain

$$f_{++} = \sum_J (f_{J-} + f_{J+}) \cos \frac{\theta}{2} \left[P'_{J+1/2}(\theta) - P'_{J-1/2}(\theta) \right] \quad (13.63a)$$

$$f_{+-} = \sum_J (f_{J-} - f_{J+}) \sin \frac{\theta}{2} \left[P'_{J+1/2}(\theta) + P'_{J-1/2}(\theta) \right] e^{-i\phi} \quad (13.63b)$$

$$P'_L = \frac{dP_L(\theta)}{d \cos \theta} \quad (13.63c)$$

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{3x^2 - 1}{2},$$

$$P_3(x) = \frac{5x^3 - 3x}{2}, \dots \quad (13.63d)$$

where $P_L(\theta)$'s are Legendre polynomials. The other two amplitudes can be obtained from Eq. (13.55d). Without loss of generality we can set $\phi = 0$ and we have

$$f_{--} = f_{++}, \quad f_{-+} = -f_{+-} \quad (13.64)$$

In ordinary experiments, helicity states are not observed. In this case, the average of the initial and sum of the final helicity states are taken to give

$$\frac{d\sigma}{d\Omega} = \frac{1}{2} \sum_{\lambda\mu} |f_{\mu\lambda}|^2 = |f_{++}|^2 + |f_{+-}|^2 \quad (13.65)$$

Polarization of the Nucleon If the target nuclei are unpolarized, nucleons in the final states of πN scattering are polarized perpendicular to the reaction plane. The reasoning is as follows. For fixed z -axis, there are two spin degrees of freedom in both the initial and final states. Consequently the scattering amplitude can be expressed as a 2×2 matrix. Any 2×2 matrix can be expanded in terms of a unit matrix and Pauli's matrices σ . The rotational invariance of the interaction restricts them to be scalars. Since the only available physical quantities in the center of mass frame are initial momentum $\mathbf{p}_i$ and final momentum $\mathbf{p}_f$, the scalars that can be made are limited to

$$1, \quad \sigma \cdot \mathbf{p}_i, \quad \sigma \cdot \mathbf{p}_f, \quad \sigma \cdot \mathbf{p}_i \times \mathbf{p}_f \quad (13.66)$$

$\sigma \cdot \mathbf{p}_i$ and $\sigma \cdot \mathbf{p}_f$ change sign under the parity operation and hence do not contribute in parity-conserving reactions. Only the terms 1 and $\sigma \cdot \mathbf{p}_i \times \mathbf{p}_f$ are allowed and the former does not change the polarization of the nucleon. From the latter, we know that the polarization generated by the scattering is perpendicular to the scattering plane.

If we set $\phi = 0$, the scattering occurs in the x - z plane and the polarization is in the y direction. If we adopt the z -axis along the final momentum direction, the eigenstates of σ_y can be constructed from the helicity eigenstates in the following way:

$$\begin{aligned} |\uparrow\rangle &= \frac{1}{\sqrt{2}}(|+\rangle + i|-\rangle) \\ |\downarrow\rangle &= \frac{1}{\sqrt{2}}(|+\rangle - i|-\rangle) \end{aligned} \quad (13.67)$$

The probability of the spin being upward ($s_y = 1/2$) is given by

$$\begin{aligned} \frac{d\sigma}{d\Omega}(\uparrow) &= \frac{1}{2} \sum_{\lambda} |\langle \uparrow | f | \lambda \rangle|^2 = \frac{1}{4} \sum_{\lambda} |f_{+\lambda} - i f_{-\lambda}|^2 \\ &= \text{Im}(f_{+-}^* f_{++}) + \frac{1}{2}(|f_{++}|^2 + |f_{+-}|^2) \end{aligned} \quad (13.68)$$

As the probability of the spin being downward can be calculated similarly, the degree of polarization is given by

$$P \equiv \frac{d\sigma(\uparrow) - d\sigma(\downarrow)}{d\sigma(\uparrow) + d\sigma(\downarrow)} = \frac{2\text{Im}(f_{+-}^* f_{++})}{|f_{++}|^2 + |f_{+-}|^2} \quad (13.69)$$

The polarization gives information independent from that given by the scattering cross section. For instance, as is described in the following, the angular distribution of scattering cross sections having the same J but different L is the same and cannot be distinguished for different L or equivalently different parity states. Then a polarization measurement can resolve the degeneracy.

Differential Cross Section When only a specific J contributes, the unpolarized cross section is given by

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \left(J + \frac{1}{2}\right)^2 |f_{J\pm}|^2 \left[(d_{1/2,1/2}^J)^2 + (d_{1/2,-1/2}^J)^2\right] \\ &= 2\pi(2J + 1) |f_{J\pm}|^2 I_J(\theta)\end{aligned}\quad (13.70)$$

where we used Eq. (E.18) and defined $I_J(\theta)$ as a normalized angular distribution such that integration over 4π solid angle gives 1. When $J = 1/2$, f_{J-} corresponds to $L = 0$ and f_{J+} to $L = 1$. If only a single partial wave contributes, we have, using Eqs. (13.63),

$$\begin{aligned}S_{1/2}: \quad f_{++} &= f_{1/2-} \cos(\theta/2), \quad f_{+-} = f_{1/2-} \sin(\theta/2) \\ P_{1/2}: \quad f_{++} &= f_{1/2+} \cos(\theta/2), \quad f_{+-} = -f_{1/2+} \sin(\theta/2)\end{aligned}\quad (13.71)$$

Using the notation commonly used in spectroscopy,⁶⁾ the differential cross sections are expressed as

$$S_{1/2}: \quad \frac{d\sigma}{d\Omega} = |f_{++}|^2 + |f_{+-}|^2 = |f_{1/2-}|^2 \quad (13.72a)$$

$$P_{1/2}: \quad \frac{d\sigma}{d\Omega} = |f_{1/2+}|^2 \quad (13.72b)$$

Similarly using Eq. (E.22)

$$d_{\pm 1/2,1/2}^{3/2} = \frac{(3 \cos \theta \pm 1)}{4} \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \quad (13.72c)$$

$$P_{3/2}: \quad \frac{d\sigma}{d\Omega} = 8\pi |f_{3/2-}|^2 \frac{(1 + 3 \cos^2 \theta)}{8\pi} \quad (13.72d)$$

$$D_{3/2}: \quad \frac{d\sigma}{d\Omega} = 8\pi |f_{3/2+}|^2 \frac{(1 + 3 \cos^2 \theta)}{8\pi} \quad (13.72e)$$

From the above formula and also from Eq. (13.70), we see that the angular distributions are solely functions of J and are independent of parity or L for the same J .

13.3.3

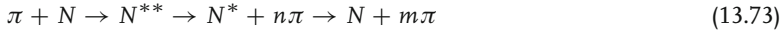
Resonance Extraction

Reactions between pions (π) and nucleons (denoted as N , not to be confused with n to denote neutrons) are among the best investigated phenomena. This is because the π meson is light and an intense pion beam can easily be made by bombarding any material with protons with energy higher than 300 MeV. Resonances, which appear as large peaks of the scattering cross sections, were of immediate interest. With the advent of (then) high-energy intense accelerators, together with the invention of a powerful detector, the bubble chamber, (see Fig. 14.2 for more detailed description) many hadronic resonances were discovered and investigated in the

6) $S, P, D, F, \dots$ denotes $L = 0, 1, 2, 3, \dots$

7) As the antinucleon has the baryon number -1 , extra nucleon-antinucleon pairs are allowed in the final state, but we do not consider them staying in relatively low-energy regions.

1960s. Although the resonances as well as the pion and the nucleon (collectively referred to as hadrons) are actually composites of quarks, we will stick mostly to the terminology used in spectroscopy in this chapter. The composite nature of the hadrons will be treated in the next chapter. Kaons and hyperons Y (baryons that carry strangeness quantum number) were also investigated in this era. Apart from selection rules arising from the strangeness, kinematical structures of KN or πY are similar to those of πN . Therefore, we mostly concentrate on πN phenomena. Resonances in the πN system denoted as N^* , N^{**} , $\dots$ are dynamical enhancements of many particles in scattering states. Like those of atoms and nuclei, higher resonances quickly descend to lower energy levels, eventually settling in the ground state as a nucleon and mesons.⁷⁾ Namely



Because of baryon number conservation, N^{**} and N^* have the baryon number 1. It is a number assigned to the nucleon, hyperon and their excited states. In contrast, mesons have baryon number 0. A general pattern of πN reactions is depicted in Fig. 13.10. Hadronic resonances typically have widths (Γ) of 10–100 MeV, which

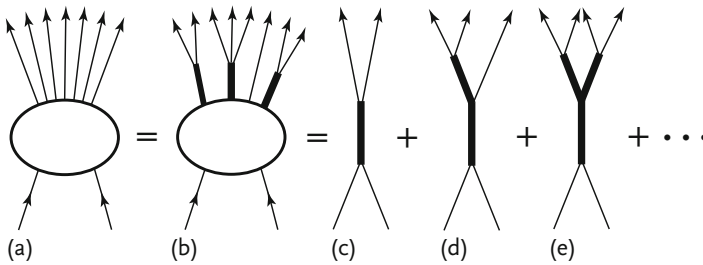


Figure 13.10 A general pattern of πN reactions. Thin lines denote π and N and thick lines denote resonances. Resonances have short decay times, typically on the order of

the strong time scale $\sim 10^{-23}$ s. π Mesons and other weakly decaying particles (K , Λ , Σ , etc.) which have lifetimes of 10^{-10} – 10^{-8} s are considered as stable in strong interactions.

means their lifetimes are $\tau \sim \hbar/\Gamma \sim 10^{-22}$ – 10^{-23} s.⁸⁾ At flight velocity near the speed of light their flight distance before decay is at most $\sim 10^{-15}$ m, comparable to hadron size and far below observable limits. What is observed is debris of decayed resonances. However, energy-momentum conservation helps us to reconstruct their mass. We define the invariant mass M of a many-particle system by

$$M^2 = (p_1 + p_2 + \dots)^2 = (E_1 + E_2 + \dots)^2 - (\mathbf{p}_1 + \mathbf{p}_2 + \dots)^2 \quad (13.74)$$

where p_i 's are four-component momenta of particles. Therefore, one method to find the resonances is to pick up $m\pi$ or $m\pi + N$ out of an n -particle final state in the multi-pion production reaction, plot the number of appearances as a function of their invariant mass and see if a peak or peaks appear. In the special case

8) A particle flying at near light speed can have longer life time because of Lorentz time dilatation. Here we are considering only the order of magnitude.

where resonances can be formed by an incoming π (particle 1) and a target nucleon (particle 2) (Fig. 13.10c), the invariant mass is given by

$$\begin{aligned} M^2 &= (p_1 + p_2)^2 \equiv s = (E_1 + m)^2 - \mathbf{p}_2^2 \\ &= (m_1 + m_2)^2 + 2m_2 T_1 \\ T_1 &= E_1 - m_1 \end{aligned} \quad (13.75)$$

Here we assumed the target is at rest and used $\mathbf{p}_2 = 0$. A plot of scattering cross section as a function of the invariant mass immediately indicates the existence of a resonance as a peak or bump (see Fig. 13.8) and its mass can be determined by Eq. (13.75). To sort out many resonances, their quantum numbers, including spin, parity and isospin, have to be determined.

Resonance Behavior Let us look at the behavior of a resonance when it is a dominant contribution to the scattering amplitude. Since the resonance is an unstable intermediate state, we may treat it as a particle that has energy with negative imaginary part $\Gamma/2$ [see Eq. (2.14)]. When a resonance intermediate state $|R\rangle$ dominates, perturbation expansion of the scattering amplitude in quantum mechanics gives

$$T_{fi} = \sum_n \frac{\langle f|H_I|n\rangle\langle n|H_I|i\rangle}{E_0 - E_n} \sim \frac{\langle f|H_I|R\rangle\langle R|H_I|i\rangle}{E_0 - (E_R - i\frac{\Gamma}{2})} \propto \frac{\sqrt{\Gamma_f}\sqrt{\Gamma_i}}{E_0 - (E_R - i\frac{\Gamma}{2})} \quad (13.76)$$

Here $\Gamma_f \sim |\langle f|H_I|R\rangle|^2$, $\Gamma_i \sim |\langle i|H_I|R\rangle|^2$ are partial decay widths to the states $|f\rangle, |i\rangle$ and the total decay width is given by

$$\Gamma = \sum_k \Gamma_k \propto \sum_k |\langle k|H_I|R\rangle|^2 \quad (13.77)$$

The elastic scattering matrix for a single partial wave can be expressed as

$$S = e^{2i\delta} = \frac{\cot \delta + i}{\cot \delta - i} \quad (13.78)$$

Let us expand $\cot \delta$ in Taylor series in the neighborhood of E_R , which gives $\delta = \pi/2$,

$$\cot \delta = \cot \delta_{E_R} + (E - E_R) \left. \frac{d \cot \delta}{dE} \right|_{E_R} + \dots \quad (13.79)$$

At $E = E_R$ the real part of the denominator vanishes, we can put

$$\cot \delta_{E_R} = 0, \quad - \left. \frac{d \cot \delta}{dE} \right|_{E_R} = - \left. \frac{d \delta}{dE} \right|_{E_R} \equiv - \frac{2}{\Gamma} \quad (13.80a)$$

The scattering matrix can be approximated as

$$S \sim \frac{E - E_R - i(\Gamma/2)}{E - E_R + i(\Gamma/2)} \quad (13.80b)$$

which results in the scattering amplitude

$$f(E) = \frac{e^{i\delta} \sin \delta}{p} = -\frac{1}{p} \frac{(\Gamma/2)}{E - E_R + i(\Gamma/2)} \quad (13.81)$$

This has the desired form to reproduce a resonant shape of the scattering cross section. In other words, the resonance occurs where the phase shift δ_J of a partial wave becomes $\pi/2 \pm n\pi$. As can be seen from the definition of Γ [Eq. (13.80a)], if the phase shift as a function of energy increases rapidly in the vicinity of $\pi/2$, Γ becomes small, giving a sharp peak.

Unitary Limit For each partial wave, the range of value $S^J = e^{2i\delta_J}$ is constrained to $|e^{2i\delta_J}| \leq 1$ and we have the following unitary limit for the scattering amplitude:

$$|f_{J\pm}|^2 = \frac{\sin^2 \delta_{J\pm}}{p^2} \leq \frac{1}{p^2} \quad (13.82)$$

This is equivalent to saying that the probability should not exceed 1. The cross section of a partial wave for an unpolarized target in the vicinity of a resonance is given by

$$\frac{d\sigma}{d\Omega} = \frac{2\pi(2J+1)}{p^2} \frac{(\Gamma/2)^2}{(E - E_R)^2 + (\Gamma/2)^2} I_J(\theta) \quad (13.83)$$

For inelastic scattering ($i \neq f$) it becomes [see Eq. (13.76)]

$$\frac{d\sigma}{d\Omega} = \frac{2\pi(2J+1)}{p^2} \frac{(\Gamma_f/2)(\Gamma_i/2)}{(E - E_R)^2 + (\Gamma/2)^2} I_J(\theta)^9 \quad (13.84)$$

In particular, for $J = 3/2$

$$\frac{d\sigma}{d\Omega} = \frac{8\pi}{p^2} \frac{(\Gamma_f/2)(\Gamma_i/2)}{(E - E_R)^2 + (\Gamma/2)^2} \left(\frac{1 + 3 \cos^2 \theta}{8\pi} \right) \quad (13.86)$$

From Eq. (13.83), we see that the resonance occurs at $E = E_R$ and when no inelastic channels are open it can reach the unitary limit.

Spin and Isospin of Δ (1232) $\Delta(1232)$, denoted as Δ in this chapter, appeared in the πN scattering cross section as the first prominent peak in Fig. 13.8:

$$\pi + N \rightarrow \Delta \rightarrow \pi + N \quad (13.87)$$

Its isospin is $3/2$ since four charged states ($\Delta^{++}, \Delta^+, \Delta^0, \Delta^-$) were observed. In order to determine its angular momentum J , we compare the total cross section

9) Equations (13.83) and (13.84) apply to spin-averaged πN scattering. More generally, for incident particles with spin s_a and s_b , the spin-averaged resonance formula is given by

$$\frac{d\sigma}{d\Omega} = \frac{4\pi(2J+1)}{(2s_a+1)(2s_b+1)} \frac{1}{p^2} \frac{(\Gamma_f/2)(\Gamma_i/2)}{(E - E_R)^2 + (\Gamma/2)^2} I_J(\theta) \quad (13.85)$$

with the unitary limit. As it occurs around $T \simeq 180$ MeV, which is low enough to be able to ignore the inelastic channel contribution, we may approximate $\Gamma_i = \Gamma_f = \Gamma$. Then the cross section should take the value of unitary limit at the resonance peak.

Figure 13.11a [264] shows that the unitary limit is reached almost completely, which determines the spin of $\Delta(1232)$ to be $J = 3/2$. Then, the angular distribution should be that of $J = 3/2$ given by Eq. (13.86). The data [31] (Fig. 13.11b) at $T = 170$ MeV confirm this. As previously explained, the angular distribution at the resonance alone cannot distinguish the L value or parity of Δ ,

$$P_\Delta = P_\pi P_p (-1)^L = (-1)^{L+1} \quad (13.88)$$

but considering the interference effect with other partial waves, it can be determined. Classically, the orbital angular momentum corresponds to $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, where $\mathbf{r}$ is restricted by the nuclear force range ($|\mathbf{r}| < R$), which is known to be $\sim 10^{-13}$ cm. At the Δ region, the pion momentum is ~ 300 MeV/ c , and only $L = 0, 1$ are significant. Thus we are led to conjecture that Δ is $P_{3/2}$ with $L = 1$. To confirm this, let us consider how the angular distribution is affected if there is a small $L = 0$ background, which is expected to vary slowly. Using Eqs. (13.63), the scattering amplitude including both S ($L = 0$) and P ($L = 1$) wave contributions is expressed as

$$d\sigma \sim |S|^2 + |P|^2(1 + 3\cos^2\theta) + 4\text{Re}(S^*P)\cos\theta \quad (13.89a)$$

On the other hand, if we assume S ($L = 0$) and D ($L = 2$), the expression becomes

$$d\sigma \sim |S|^2 + |D|^2(1 + 3\cos^2\theta) + 4\text{Re}(S^*D)(3\cos^2\theta - 1) \quad (13.89b)$$

From Eqs. (13.89), we see that the interference term gives $\cos\theta$ dependence and produces asymmetry for the assignment $\Delta = P$, but no asymmetry for D . In general, asymmetric terms arise from interferences between waves of opposite parity. The resonance amplitude is expressed as

$$f_J \sim \frac{1}{E - E_R + i(\Gamma/2)} = \frac{E - E_R - i(\Gamma/2)}{(E - E_R)^2 + (\Gamma/2)^2} \quad (13.90)$$

and its functional form is described in Fig. 13.12. Assuming the S wave does not produce a resonance in this region, its amplitude is a slowly varying function. Then the interference term with the real part of the resonance amplitude changes its sign below and above the resonance energy. Angular distributions at $T_\pi = 98, 310$ MeV shown in Fig. 13.11b behave exactly as expected for the existence of $\cos\theta$ interference, which vanishes at the resonance but produces opposite asymmetries below and above the resonance, confirming our previous conjecture that $L = 1$. In summary, the spin parity of Δ is determined to be $J^P = (3/2)^+$.

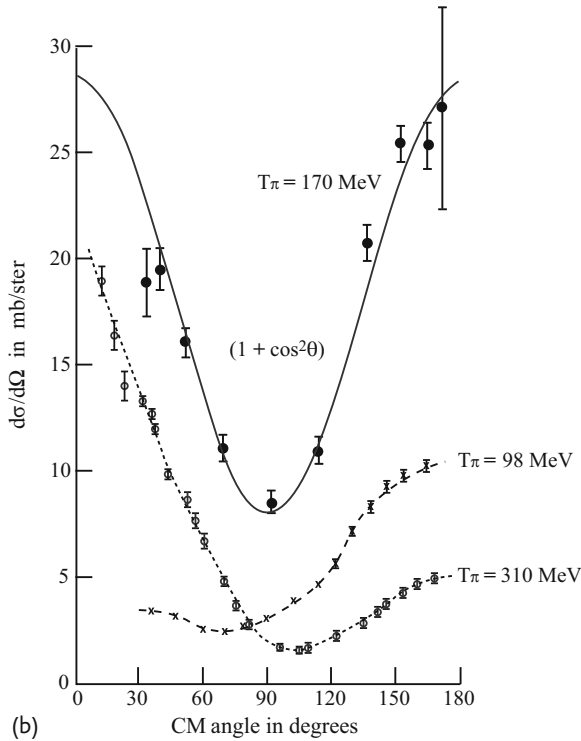
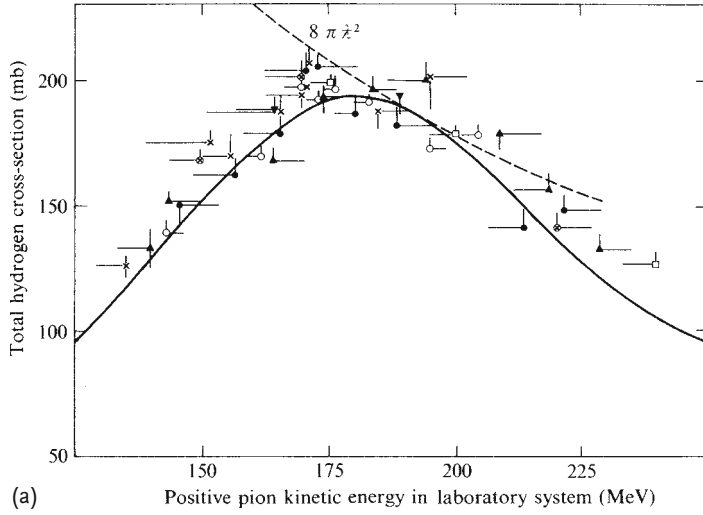


Figure 13.11 (a) The production cross section of $\Delta(1232)$ reaches its unitary limit $8\pi\lambda^2 = 8\pi/p^2$ [264]. (b) Angular distributions of the πN system in the neighborhood of $\Delta(1232)$. At the resonance peak

($T_\pi = 180 \text{ MeV}$, $(s = 1232 \text{ MeV})^2$), the angular distribution is $\sim 1 + 3 \cos^2 \theta$. Below and above the resonance, the angular distribution is asymmetric due to interference with S-wave amplitude (see text) [23, 31, 90].

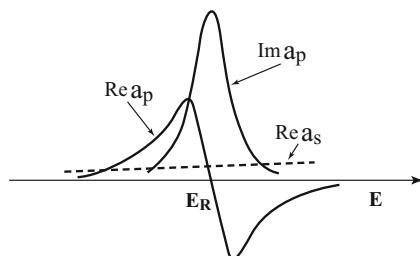


Figure 13.12 Qualitative behavior of a resonance amplitude a_p and background amplitude a_s .

13.3.4

Argand Diagram: Digging Resonances

Not all resonances are as conspicuous as $\Delta(1232)$. Some of them appear as peaks and bumps, but many are hidden behind conspicuous resonances or overwhelming backgrounds that come from inelastic channels. To disentangle many overlapping resonances and backgrounds, it is necessary to make a detailed partial wave analysis as a function of energy and angles. If it is assumed that the maximum orbital angular momentum is limited by $L_{\max} \simeq pR$, $R \simeq 1/m_\pi$, the scattering amplitude can be expressed as a finite number of partial waves. When the cross-section as well as polarization data are fitted as functions of energy and angle, every partial wave scattering amplitude (A_J) can be determined, at least in principle. Let us examine how these partial amplitudes behave. If inelastic channels are open that

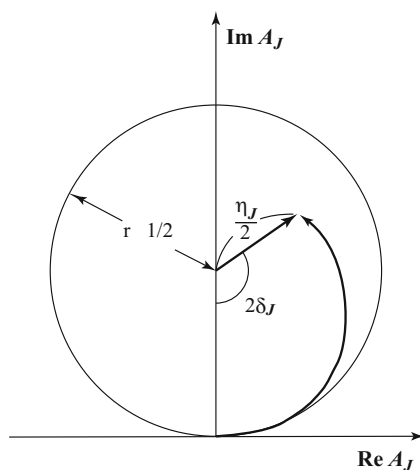


Figure 13.13 A complex partial-wave amplitude lies within a circle of radius $1/2$ (unitarity). For an elastic channel, it is located at the origin when the kinetic energy is zero. As the energy increases, it moves clockwise or coun-

terclockwise depending on whether the force is repulsive or attractive. When $\delta_J = \pi/2$, namely when the amplitude crosses the imaginary axis counterclockwise, it shows a resonance behavior.

damp the elastic waves, its effect can be taken into account by introducing elasticity as defined in Eq. (13.61). Then the partial-wave scattering amplitude can be written as

$$A_J \equiv p f_J = \frac{\eta_J e^{2i\delta_J} - 1}{2i} \quad (13.91)$$

On the $\text{Re}(A_J)$ – $\text{Im}(A_J)$ plane, A_J is confined to within a circle of unit radius with its center located at $x = 0$, $y = 1/2$. This is called the Argand diagram. This is another expression for the unitary limit. Starting at the origin when the kinetic energy $T = 0$, A_J proceeds on the circle clockwise when the force is repulsive and counterclockwise when attractive. As the rotation angle is given by 2δ and its distance from the center by $\eta_J/2$, the track departs from the circle when inelastic channels such as $2\pi + N$, $3\pi + N, \dots$ are open (see Fig. 13.13). When it crosses the imaginary axis counterclockwise, $\delta_J = \pi/2$, it becomes a resonance amplitude. When the speed at crossing is fast, the resonance width is narrow [see Eq. (13.80)]. In the vicinity of the resonance, the amplitude A_J takes the form of Eq. (13.81); we can express it as a function of two parameters x_σ , x_λ and ϵ defined by

$$A_{\sigma\lambda} = -\frac{\sqrt{x_\sigma x_\lambda}}{\epsilon + i} \quad (13.92)$$

$$x_\sigma = \frac{\Gamma_\sigma}{I}, \quad \epsilon = (E - E_R) \frac{2}{I}$$

When many channels mix, the behavior of the resonance (in the Argand diagrams) is very complex (see Problem 13.8). Generally, if the track is circling counterclockwise, a place where it runs parallel to the x -axis can be considered as a resonance. For πN and KN scattering, there are rich stocks of data and detailed analyses have been carried out [204]. Figure 13.14 shows part of two independent analyses using different formalisms. In general, they show good agreement. We see that many resonances that are not apparent in Fig. 13.8 were revealed by doing detailed partial wave analysis.

Problem 13.8

Show that resonances in the Argand diagram in Fig. 13.15 occur for the following cases:

- (a) Resonance amplitude only.
- (b, c) Resonance when it coexists with a background wave that has attractive force and the same spin-parity.
- (d) Resonance when it coexists with a background wave that has repulsive force and the same spin-parity.

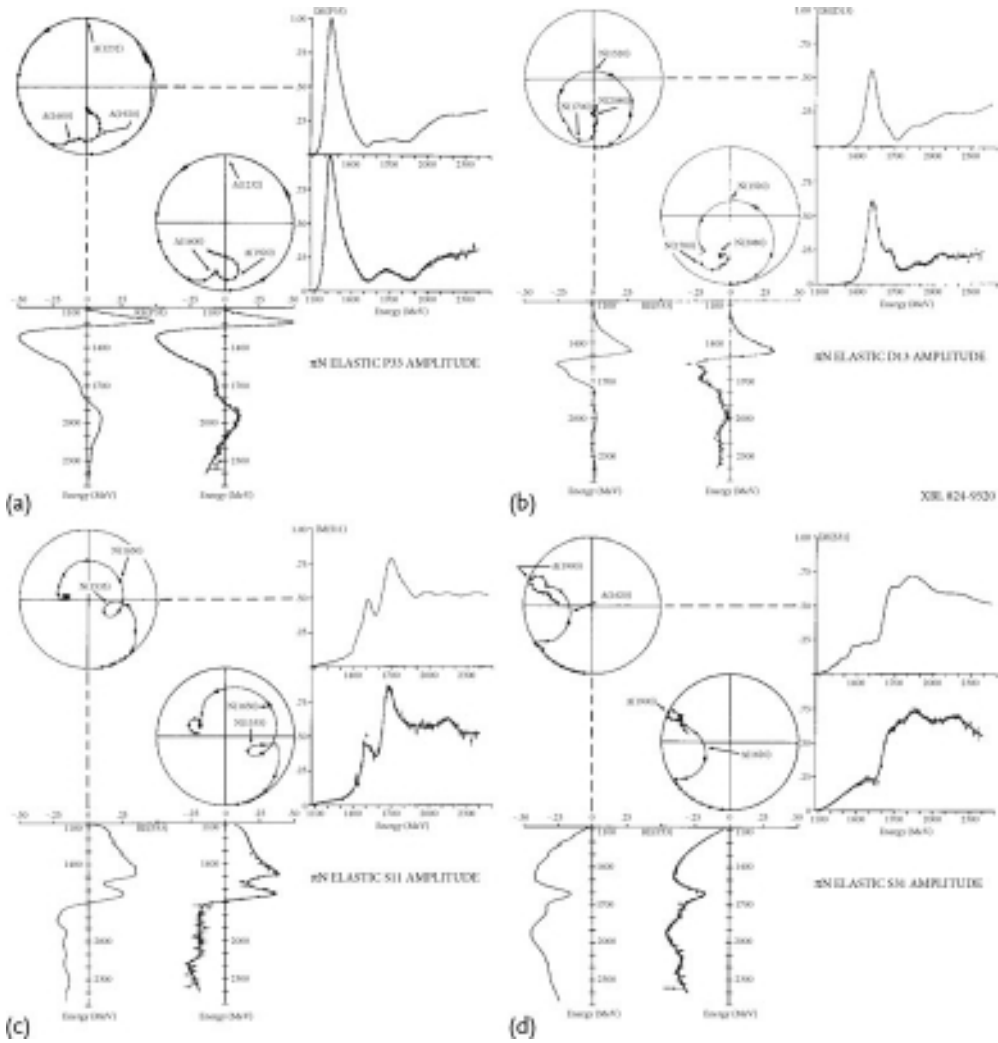


Figure 13.14 Examples of πN scattering partial-wave analysis. It is common to use N and Δ to denote $I = 1/2$ and $I = 3/2$ with S (strangeness) = 0 members. (a) P_{33} ($L = 1, I = J = 3/2$) contour includes $\Delta(1232)$, etc. (b) D_{13} ($L = 2, I =$

$1/2, J = 3/2$) contour includes $N(1520)$, etc. (c) S_{11} ($L = 0, I = 1/2, J = 1/2$) contour includes $N(1535)$, etc. (d) S_{31} ($L = 0, I = 3/2, J = 1/2$) contour includes $\Delta(1620)$, etc. [204].

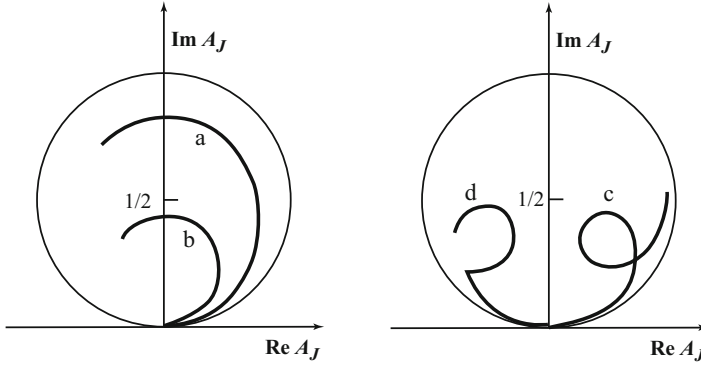


Figure 13.15 Depending on the elasticity η_J being smaller or larger than 0.5 or/and the existence of background amplitudes, there exist a variety of resonance behaviors.

13.4

ρ (770)

Mass and Width ρ was the first resonance discovered in the $\pi\pi$ channel. Three charged states are known

$$\rho^+ \rightarrow \pi^+ + \pi^0, \quad \rho^0 \rightarrow \pi^+ + \pi^-, \quad \rho^- \rightarrow \pi^- + \pi^0 \quad (13.93)$$

which means the isospin of ρ is $I = 1$. If real $\pi\pi$ scattering is possible, its spin and parity could be determined in the same way as $\Delta(1232)$. π is an unstable particle and cannot be used as a real target, but there is a way to make one effectively. Due to the Heisenberg uncertainty principle, a proton can emit and reabsorb pions as follows:

$$p \leftrightarrow n + \pi^+, \quad p \leftrightarrow p + \pi^0 \quad (13.94)$$

The transition is an effect of strong interaction and occurs frequently. Consequently, the proton can be considered as surrounded by a cloud of pions, which could serve as the target. If we can shoot pions at the cloud, we can make $\pi\pi$ scattering happen. A necessary condition is that the time for which a π is virtually free is long compared to the collision time. Referring to Fig. 13.16, this means the contribution of diagram (b) must dominate. Using (q, μ) , (p_p, M) , (p_n, M) to denote the four-momenta and masses of the target π , proton and neutron, respectively, the condition for the pion to be free is

$$q^2 = q_\alpha q^\alpha = \mu^2 \quad (13.95)$$

10) Note there is no $\rho^0 \rightarrow \pi^0 + \pi^0$ mode. This is forbidden by isospin invariance and boson statistics. The $I = 1$ state made of two $I = 1$ states is antisymmetric, and $2\pi^0$ can only be symmetric.

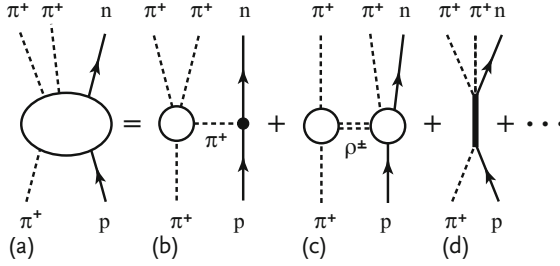


Figure 13.16 Diagrams contributing to $\pi N \rightarrow \pi\pi N$ reaction: π exchange (b) and other contributions (c,d).

In a frame where $|\mathbf{p}_p| = |\mathbf{p}_n| = p$, assuming the masses of the proton and the neutron are the same,

$$\begin{aligned} E_p &= E_n \\ q^2 &= (p_p - p_n)^2 = (E_p - E_n)^2 - (\mathbf{p}_p - \mathbf{p}_n)^2 = 2p^2(1 - \cos \theta) < 0 \end{aligned} \quad (13.96)$$

q^2 is negative, meaning it is not free but a virtual particle. The difference between the mass of the free and the virtual particle $\delta\mu = \sqrt{\mu^2 - q^2}$ indicates the degree of virtuality. From the uncertainty principle, the time duration for which the target pion is free may be measured by $\tau_{\text{free}} \sim 1/\delta\mu$. q^2 cannot be made positive. Then, to make τ_{free} large, i.e. $\delta\mu$ as small as possible, $|q^2|$ must be small, or from Eq. (13.96) we have to make $\cos \theta \rightarrow 1$. Considering the requirement that the target π should be well away from the nucleon, in other words, the space extension of the virtual π should be much larger than the size of the proton $R \sim 1/\mu$

$$|\mathbf{q}| = |\mathbf{p}_p - \mathbf{p}_n| \simeq \sqrt{2p^2(1 - \cos \theta)} \ll \mu \quad (13.97)$$

we see $\tau_{\text{free}} \sim 1/\delta\mu \sim 1/\mu$. The collision time τ_{coll} is given by the time necessary for a particle to go through the target. Since the target looks thin because of the Lorentz contraction, it is given by $\tau_{\text{free}} \sim R/c\gamma \sim (1/\mu)(E_{\text{in}}/M)^{-1}$, where E_{in} is the energy of the incoming π . Therefore conditions for $\tau_{\text{coll}} \ll \tau_{\text{free}}$ are given by Eq. (13.97) plus

$$E_{\text{in}} \gg M \quad (13.98)$$

If they are satisfied we can consider the concept of the virtual target to be valid.¹¹⁾

11) The objection may be raised that because of the Regge exchange dominance in the high-energy forward scattering (see Section 13.7) the exchanged particle may not be limited to a single π . The reason justifying the dominance of diagram (b) in Fig. 13.16 comes from the experimental observation that

$$r \equiv \frac{\sigma(\pi^- p \rightarrow \pi^- \pi^+ n)}{\sigma(\pi^- p \rightarrow \pi^- \pi^0 n)} \simeq 2 \quad (13.99)$$

Assumption of isospin invariance gives

$$\frac{\sigma(\pi^- \pi^+ \rightarrow \pi^- \pi^+)}{\sigma(\pi^- \pi^0 \rightarrow \pi^- \pi^0)} = 1 \quad (13.100)$$

Since there is a factor $\sqrt{2}$ difference between π^+ and π^0 in the Clebsch-Gordan coefficients of πN coupling, the observed ratio confirms the isospin invariance, hence the dominance of the target pion.

Therefore, when performing the experiment

$$\pi^+ p \rightarrow \pi^+ + \pi^0 + p \quad (13.101)$$

which includes two π 's in the final state, if we set the incident π energy sufficiently high, and choose events with small recoil angle of the target proton, we are doing the pseudoexperiment

$$\pi^+ + \pi_{\text{tgt}} \rightarrow \pi^+ + \pi \quad (13.102)$$

where π_{tgt} is π 's emitted from the proton. Once the experimental conditions for $\pi\pi$ scattering are settled, the determination of the mass and width of a $\pi\pi$ resonance is straightforward. All we have to do is to plot the cross section as a function of incident energy and measure the position and width of a peak. The observed values

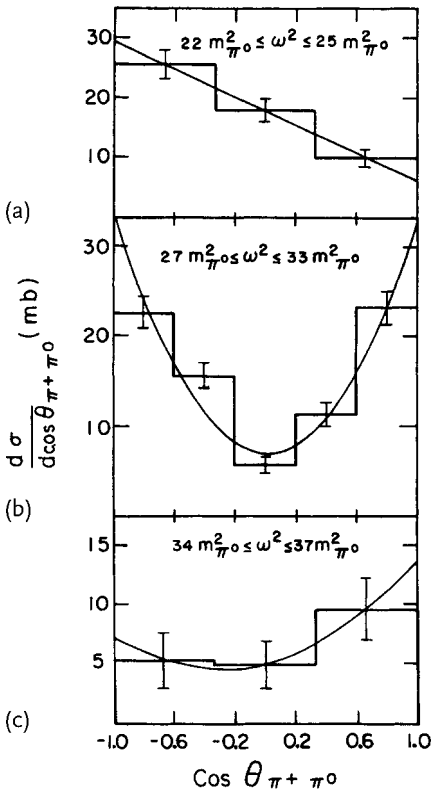


Figure 13.17 Angular distributions in the di-pion CM frame of the reaction $\pi^+ p \rightarrow \rho^+ p \rightarrow \pi^+ + \pi^0 + p$. (a) $640 < \sqrt{s} < 680$, (b) $700 < \sqrt{s} < 780$, (c) $790 < \sqrt{s} < 820$. At $\sqrt{s} = m_\rho \simeq 740$ MeV, the angular dis-

tribution fits with $\cos^2 \theta$. The cross section is not zero at $\theta = 90^\circ$ because of nonresonant contributions. Note the absence of the $\cos^2 \theta$ term below and above the resonance [88].

are

$$\begin{aligned} m(\rho) &= 775 \text{ MeV} \\ \Gamma &= 146 \text{ MeV} \end{aligned} \quad (13.103)$$

Spin In a similar manner to the case of $\Delta(1232)$, by choosing the z -axis along the incident π^+ direction, we can set $J_z = L_z + S_z = 0$, where we have used the fact that the spin of π is 0. If a resonance is made with angular momentum J , the amplitude which the two π 's make is proportional to $d_{00}^J = P_J$ in the center of mass (CM) frame of the two π 's. P_J is a Legendre polynomial and is given by

$$\begin{aligned} I(\theta) &\sim |P_0|^2 \sim \text{constant} & J &= 0 \\ &\sim |P_1|^2 \sim \cos^2 \theta & J &= 1 \\ &\sim |P_2|^2 \sim (3 \cos^2 \theta - 1)^2 & J &= 2 \end{aligned} \quad (13.104)$$

When we plot the cross section as a function of the invariant mass $m_{\pi\pi}^2$, the peak position gives m_ρ^2 . The angular distribution at the resonance is given in Fig. 13.17, which fits well with $\cos^2 \theta$ and proves $J^P = 1^-$.

13.5

Final State Interaction

13.5.1

Dalitz Plot

In the previous section, we described a method to extract $\pi\pi$ scattering from a reaction

$$\pi + N \rightarrow \pi + \pi + N \quad (13.105)$$

There, the target nucleon played an auxiliary role in providing a virtual π target. A more general method to establish the existence of resonances is to make an invariant mass distribution of an arbitrary number of particles extracted from multiparticle final states. Using the invariant mass, we can investigate resonances not only in the $\pi\pi$ channel but also in many π' s, πN channels and other combinations. In order to investigate the behavior of the matrix element, we need to know event distributions when it is a constant, namely the distribution of phase space. In general, when there are n particles, the scattering cross section is given by (see Section 6.5.4)

$$d\sigma, d\Gamma \propto |\langle f | T | i \rangle|^2 dLIPS \quad (13.106a)$$

$$dLIPS = \frac{1}{m!} (2\pi)^4 \delta^4 \left(P_i - \sum_f p_f \right) \prod_f^n \frac{d^3 p_f}{(2\pi)^3 2E_f} \quad (13.106b)$$

where P_i is the total four-momentum in the initial state and $m!$ is a correction factor for m identical particles.

Let us consider, as an example, a case with $n = 3$ like reaction Eq. (13.105). Integrating over three-momentum $\mathbf{p}_3$, the rest of the integral becomes proportional to

$$\frac{p_1^2 p_2^2 d p_1 d p_2 d \Omega_1 d \Omega_2}{E_1 E_2 E_3} \delta(E - E_1 - E_2 - E_3) \quad (13.107)$$

where $p_i = |\mathbf{p}_i|$. In the CM frame of the three particles, using $\mathbf{p}_3 = -(\mathbf{p}_1 + \mathbf{p}_2)$, we have

$$\begin{aligned} E_1 &= \sqrt{p_1^2 + m_1^2}, & E_2 &= \sqrt{p_2^2 + m_2^2} \\ E_3 &= \sqrt{p_1^2 + p_2^2 + m_3^2 + 2p_1 p_2 \cos \xi} \end{aligned} \quad (13.108)$$

where ξ is the angle between $\mathbf{p}_1$ and $\mathbf{p}_2$. If the initial state is unpolarized, the coordinate axis can be chosen in any direction, giving $\int d\Omega_1 = 4\pi$ and $d\Omega_2 = 2\pi d \cos \xi$. Using Eq. (13.108), variables can be changed from $(p_1, p_2, \cos \xi)$ to (E_1, E_2, E_3) . Calculating the Jacobian of the transformation, we obtain

$$d p_1 d p_2 d \cos \xi = \frac{E_1 E_2 E_3}{p_1^2 p_2^2} d E_1 d E_2 d E_3 \quad (13.109)$$

Because of the δ -function, the integration over E_3 is trivial, giving

$$d L I P S \sim \int \frac{d^3 p_1 d^3 p_2}{E_1 E_2 E_3} \delta(E - E_1 - E_2 - E_3) \propto d E_1 d E_2 \quad (13.110)$$

Namely, the phase space of the three-body final states is flat on the E_1 - E_2 plane. Therefore, the event rate dN is given by

$$dN \sim |\langle f | T | i \rangle|^2 d E_1 d E_2 \quad (13.111)$$

and the event number density on the E_1 - E_2 plane is directly proportional to the square of the matrix element, which is known as the Dalitz plot. The boundary of the distribution is given by $|\cos \xi| \leq 1$ and using Eq. (13.108) it leads to

$$-1 \leq \frac{E^2 - 2E(E_1 + E_2) + 2E_1 E_2 + m_1^2 + m_2^2 - m_3^2}{2\sqrt{E_1^2 - m_1^2}\sqrt{E_2^2 - m_2^2}} \leq 1 \quad (13.112)$$

Sometimes it is more convenient to use the invariant mass instead of the particle energy. The invariant mass (m_{23}^2) of particles 2 and 3 is given by

$$\begin{aligned} m_{23}^2 &= (p_2 + p_3)^2 = (E - E_1)^2 - \mathbf{p}_1^2 = E^2 - 2E E_1 + m_1^2 \\ \therefore d m_{23}^2 &= -2E d E_1 \end{aligned} \quad (13.113a)$$

Similarly

$$d m_{31}^2 = -2E d E_2 \quad (13.113b)$$

We see the distribution on the E_1 - E_2 plane is directly proportional to that on the m_{23}^2 - m_{31}^2 plane.

Figure 13.18 [38] shows an event distribution on the $m^2(\bar{K}^0\pi^-)$ - $m^2(n\pi^-)$ plane obtained by shooting a 3 GeV K beam on a neutron target (actually a deuteron target) and extracting $\bar{K}^0\pi^-n$ from the final state. One can clearly see concentrations of events corresponding to $K^*(892)$ and $\Delta(1232)$. Projection of events on each coordinate axis shows a sharp peak on top of a smooth slowly varying hill. The hill corresponds to the phase space volume and the sharp peak reflects the shape of the matrix element. It indicates that the matrix element $\mathcal{M}$ is proportional to

$$\mathcal{M} \propto \frac{1}{(E - E_R) + i\Gamma/2} \quad (13.114)$$

and that the resonance appears as dominant intermediate states. Similarly, the invariant mass distributions of $\pi^+\pi^0$, $\pi^+\pi^-$ and $\pi^+\pi^-\pi^0$ obtained from the reaction $\pi^+ + p \rightarrow 3\pi$ (or 4π) + p are plotted in Fig. 13.19 [12]. One immediately sees $\rho^+(770)$ in Fig. 13.19a, $\rho^0(770)$ in (b) and $\eta(548)$ and $\omega(782)$ in (c).

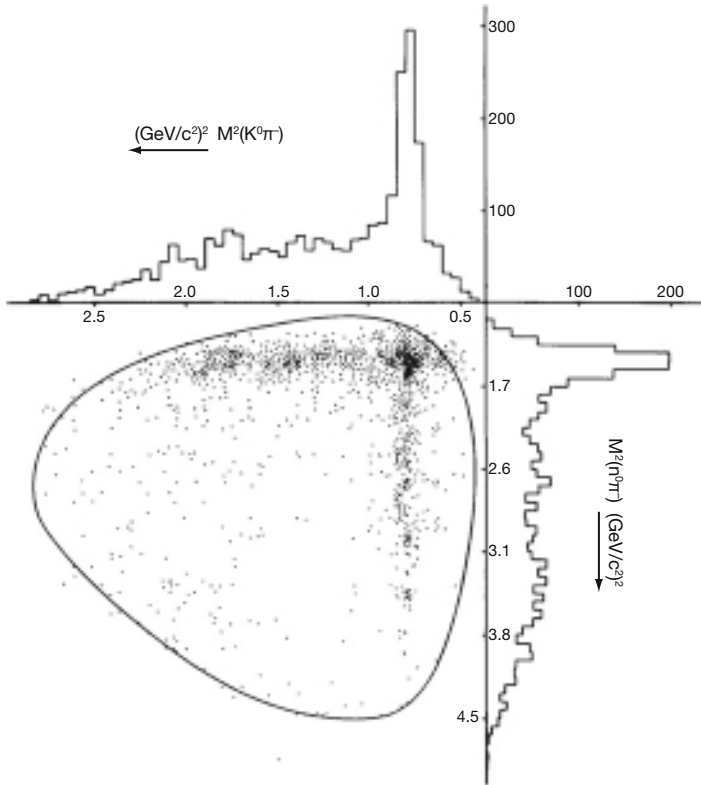


Figure 13.18 Dalitz plot of the reaction $K^- n \rightarrow K^0 \pi^- n$ and its projection on each coordinate axis. Two resonances $K^*(892)$ in $K\pi$ and $\Delta(1232)$ in the πN channel are clearly seen [38].

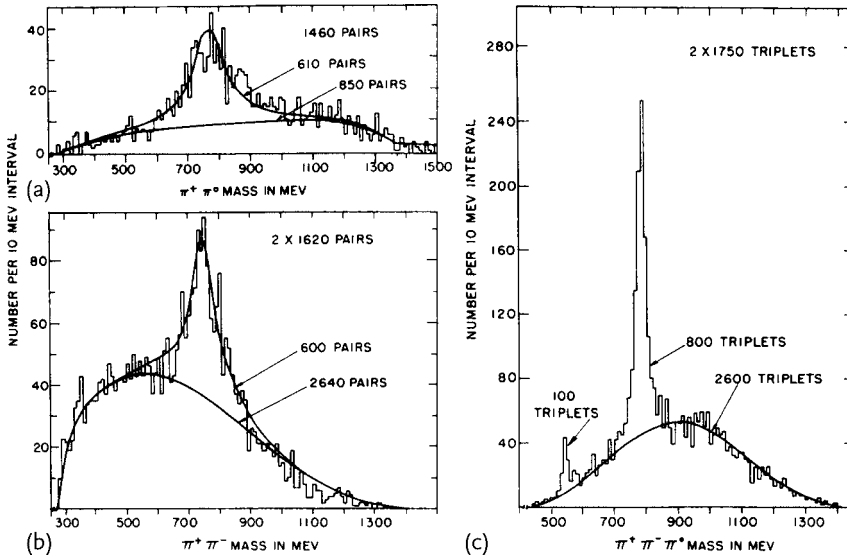


Figure 13.19 Distributions of invariant masses in the final state of $\pi^+ p$ reactions. (a) $\pi^+ + p \rightarrow \pi^+ + \pi^0 + p$ (b) $\pi^+ + p \rightarrow \pi^+ + \pi^+ + \pi^- + p$ (c) $\pi^+ + p \rightarrow \pi^+ + \pi^+ + \pi^- + \pi^0 + p$. The

abscissa is the invariant mass (a) $m(\pi^+ \pi^0)$, (b) $m(\pi^+ \pi^-)$, (c) $m(\pi^+ \pi^- \pi^0)$ in units of MeV. One can see ρ^+ in (a), ρ^0 in (b) and $\eta(550)$ and $\omega(783)$ in (c) [12].

Thus, by observing invariant mass distributions of multiparticle final states in πN and KN reactions, many resonances were discovered in the 1960s.

13.5.2

K Meson

In the K meson decay

$$K^+ \rightarrow \pi^+ + \pi^+ + \pi^- \quad (13.115)$$

the final-state particles have a good symmetry in that all the particles are π 's of more or less the same mass. In such a case it is better to treat the three variables E_1, E_2, E_3 equally. Instead of orthogonal coordinates, one may plot them as a point on the inner area of an equilateral triangle as in Fig. 13.20. If we use kinetic energy $T_i = E_i - m_\pi$ as variables,

$$T_1 + T_2 + T_3 = m_K - 3m_\pi = Q \quad (13.116)$$

Since the sum of three perpendicular lines from a point P to the bases is equal to the height CC' and is a constant, the point P can represent an event in $K \rightarrow 3\pi$. We identify the lengths of three lines as T_1, T_2, T_3 divided by Q so that the sum is normalized to 1. The boundary is constrained by $|\cos \xi| \leq 1$. For $T_i \ll m_\pi$,

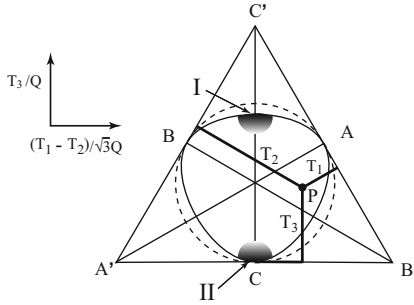


Figure 13.20 Symmetrical Dalitz plot for 3-particle decays. Point P corresponds to an event with $T_1 + T_2 + T_3 = Q$, as the sum of the three perpendicular lines T_1, T_2, T_3 to the bases is equal to the height CC' , which is a constant. The abscissa measured from C is

equal to $(T_1 - T_2)/\sqrt{3}$. For nonrelativistic particles, the boundary is a circle (dashed line), but it is deformed to a shape close to a triangle as they become relativistic. For the shaded regions I, II see the text.

it becomes a circle; this is almost the situation for K decay since the Q -value is 75 MeV, which has to be shared among the three π 's.

Problem 13.9

Show that the boundary in the triangular Dalitz plot of Fig. 13.20 is a circle or a triangle with apexes at A, B, C for $T_i \ll m_\pi$ or $T_i \gg m_\pi$.

At the boundary, two of the momenta are parallel or antiparallel. On the lines AA', BB', CC' , the energy of the two particles is the same. Setting $\pi_1 = \pi_2 = \pi^+$, the distribution should be symmetric for $1 \leftrightarrow 2$ exchange, and hence we use only the right half of the plane. If we set the orbital angular momentum in the CM frame of $\pi^+\pi^+$ as L_+ and relative orbital angular momentum of π^- to the two π^+ CM as L_- , the spin of K^+ (J) is in the range

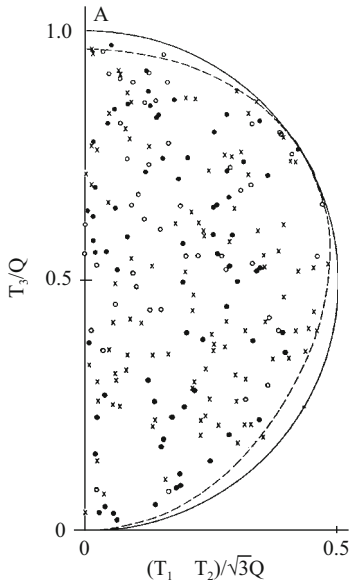
$$|L_+ - L_-| \leq J \leq L_+ + L_- \quad (13.117)$$

Since L_+ is the orbital angular momentum of two identical spinless particles, it must be even. We list possible values of J in Table 13.2. If $J \neq 0$ one of L_+, L_- is not zero. If $L_+ \geq 2$, and $\pi^+\pi^+$ are relatively at rest, the angular momentum barrier (see the following section) would diminish the value of the matrix element. This happens when the energy of π^- is maximum, namely in the region I in Fig. 13.20. Similarly if $L_- > 0$, the event rate in region II, where π^- is at rest, would also decrease. The data shown in Fig. 13.21 [305] seem very uniform over the entire region and we conclude the spin of K^+ is zero.

The three π final state has parity $P_\pi^3(-1)^{L_++L_-} = -1$ and was called the τ decay mode. There is another decay channel, $K^+ \rightarrow \pi^+\pi^0$, which was called the θ mode. The parity of θ is positive, hence τ and θ must be different particles if parity is conserved. However, τ and θ had the same mass and lifetime and were most likely the same particle. The two contradictory facts were called the τ - θ puzzle.

Table 13.2 Possible values of J^P for given (L_+, L_-) .

L_-	$L_+ = 0$	$L_+ = 2$
0	0^-	2^-
1	1^+	$1^+, 2^+, 3^+$
2	2^-	$0^-, 1^-, 2^-, 3^-, 4^-$

**Figure 13.21** Dalitz plot of $K^+ \rightarrow \pi^+ + \pi^+ + \pi^-$. The ordinate is T_3 and the abscissa is $(T_1 - T_2)/\sqrt{3}$ normalized to Q . The solid line describes a circle and the dashed line

indicates the actual boundary. From the uniformity of the distribution, the spin of the K meson was determined to be 0 [305].

zle. It led Lee and Yang to question the validity of parity invariance in the weak interaction. We now know that parity is 100% broken in the weak interaction and hence the parity of the K meson cannot be determined with reactions in which the weak force is active. The parity of K was eventually determined by kaon capture by helium, forming a hypernucleus in the reaction



In the hypernucleus one neutron is replaced with a Λ^0 particle having $m = 1116 \text{ MeV}$ and strangeness $S = -1$. The measurements showed that both He^4 and ${}_\Lambda \text{H}^4$ have $J = 0$. The K meson is expected to be captured in the S ($L = 0$) state and the parity of K is determined to be negative, assuming the parity of Λ relative to the

proton is positive. The indefiniteness originates from the fact that the strangeness is conserved in the strong interaction (see argument in Section 9.2.1).

13.5.3

Angular Momentum Barrier

When the range R of a force is finite, classically there should be no contributions from partial waves with $l > l_{\max} = pR$, where p is the particle momentum or wave number. In quantum mechanics, however, all the partial waves have components at $r < R$ and they contribute. This can be seen as follows. A plane wave can be expanded as

$$e^{ipx} = \sum_{l=0} i^l (2l+1) P_l(\cos \theta) j_l(pr) \quad (13.119)$$

This shows that a component of a partial wave in the plane wave is proportional to regular solutions of the spherical Bessel function $j_l(pr)$. Since $\lim_{pr \rightarrow 0} j_l(pr) \sim (pr)^l$, at low energies where $p \ll R^{-1}$, the l th component of the partial wave in the potential is proportional to $(pR)^l$. It can be rephrased as follows. If a particle has angular momentum l , a centrifugal potential $\propto l(l+1)/r^2$ is created, which acts as a repulsive force. When it is superimposed on the short-range attractive force, the potential looks like that in Fig. 13.22. An excited bound state level can be formed far below the peak of the barrier but with $E > 0$. It is an unstable state, because the wave can penetrate the barrier to leak outside of it (tunneling effect) and thus has finite probability of decaying. With the above reasoning, the decay matrix element becomes proportional to p^l , making the decay rate $\Gamma \propto p^{2l+1}$. The mechanism is effective for incoming waves (scattering or absorption) as well as outgoing waves, and makes the scattering amplitude proportional to p^{2l} . Then Eq. (13.62c) tells us that $\delta_l \sim p f_l \sim p^{2l+1}$. Thus, in the low-energy region ($p \ll R^{-1}$), components with high angular momentum are highly damped, which is referred to as the angular momentum barrier.

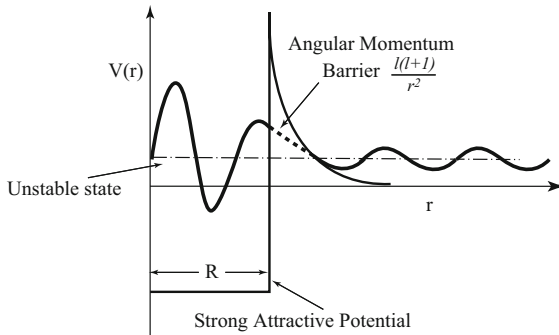


Figure 13.22 The quantum tunneling effect allows a particle to leak in or out through the angular momentum barrier with a damping factor $\sim (pR)^l$.

13.5.4

 ω Meson

The ω meson was discovered in the reaction [268]

$$\bar{p} + p \rightarrow \pi^+ + \pi^+ + \pi^0 + \pi^- + \pi^- \quad (13.120)$$

as a relatively narrow resonance. We also have seen it in Fig. 13.19c. It was seen only in neutral combinations and not in charged ones. It fixes the isospin $I = 0$. The ω meson has mass and width

$$m_\omega = 782 \text{ MeV}, \quad \Gamma_\omega = 8.5 \text{ MeV} \quad (13.121)$$

This implies the decay is strong and that the G parity is -1 . $I = 0$ implies $C = -1$. Since the isospin is a good quantum number in the strong interaction, it constrains the form of the wave function. As π is an isovector, the only way to construct an $I = 0$ combination from three pions is $I_1 \cdot I_2 \times I_3$, which is totally antisymmetric. Then Bose statistics fixes the spatial part of the wave function to be totally antisymmetric, too. This property can be used to determine the spin and parity of the ω meson in the Dalitz plot of the 3π final state. We consider constructing matrix elements using the tensor method as we did for π^0 in Section 13.2. Materials we can use are E_1, E_2, E_3 and $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$. We consider three cases.

$J^P = 0^-$: A totally antisymmetric scalar that can be made of three momenta is $\mathbf{p}_1 \cdot \mathbf{p}_2 \times \mathbf{p}_3$, which vanishes identically because $\mathbf{p}_1 = -(\mathbf{p}_2 + \mathbf{p}_3)$. Therefore, the distribution must be proportional to

$$f \propto (E_1 - E_2)(E_2 - E_3)(E_3 - E_1) \quad (13.122)$$

The event density on the Dalitz plot must vanish when two particles have the same energy, which are the lines AA' , BB' and CC' in Fig. 13.20, and is depicted as a three-dimensional plot in Fig. 13.23a. Note the difference from K and η decay (Problem 13.2). For them there are no constraints from the isospin. For ω to have a flat distribution, it has to be an isovector [398].

$J^P = 1^-$: Because 3π has negative parity, we need an axial vector here. Since $\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 = 0$, we have only one axial vector

$$f \propto \mathbf{q} = \mathbf{p}_1 \times \mathbf{p}_2 = \mathbf{p}_2 \times \mathbf{p}_3 = \mathbf{p}_3 \times \mathbf{p}_1 \quad (13.123)$$

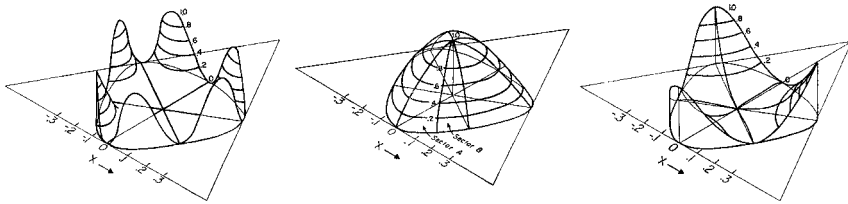


Figure 13.23 Event distributions in Dalitz plot for $I = 0$ and $J^P = 0^-, 1^-, 1^+$ [354, 398].

which is already totally antisymmetric. At the boundary of the Dalitz plot plane, two momenta are parallel or antiparallel, hence the density vanishes as depicted in Fig. 13.23b.

$J^P = 1^+$: A completely antisymmetric vector can be reconstructed as

$$f \propto \mathbf{p}_1(E_2 - E_3) + \mathbf{p}_2(E_3 - E_1) + \mathbf{p}_3(E_1 - E_2) \quad (13.124)$$

At the center $E_1 = E_2 = E_3$, hence $f = 0$. For $E_2 = E_3$, the first term vanishes and the second and third can vanish if $\mathbf{p}_1 = \text{maximum}$, because then $\mathbf{p}_2 = \mathbf{p}_3 = -\mathbf{p}_1/2$. The depletion region is shown in Fig. 13.23c. We can go on to consider higher spins but we stop here. The interested reader can refer to [369, 398].

The results can be compared with experimental data (Fig. 13.24). Since all the distributions considered here have sixfold symmetry, it is convenient to plot all the points in one sector. Figure 13.24b shows such a plot and is compared with theoretical predictions. It is clear ω has $J^P = 1^-$.

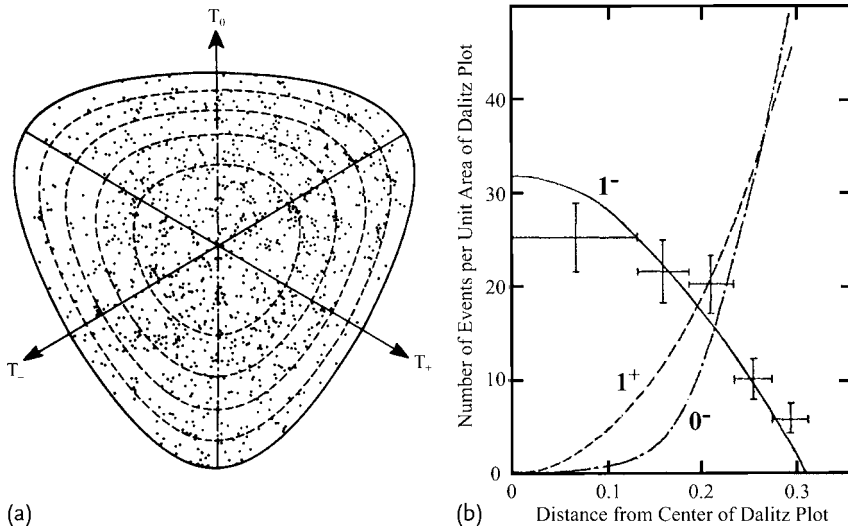


Figure 13.24 (a) Dalitz plot for the decay $\omega^0 \rightarrow \pi^+ + \pi^- + \pi^0$ [12, 13]. (b) Number of events in one sextant of the Dalitz plot. The curves correspond to the theoretical distribu-

tion for a vector meson (solid line), an axial vector (dashed line) and a pseudoscalar (dot-dashed line) [354].

13.6

Low-Energy Nuclear Force

13.6.1

Spin–Isospin Exchange Force

The scattering amplitude for NN interaction with exchange of a spin 0 isoscalar meson was given in Eq. (6.157):

$$T_{fi} = (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \bar{u}(p_3) \Gamma u(p_1) \frac{g^2}{q^2 - \mu^2} \bar{u}(p_4) \Gamma u(p_2) \quad (13.125)$$

where μ is the mass of the exchanged meson. Here we investigate the exchange of a π meson, which is a pseudoscalar ($\Gamma = i\gamma^5$) in space-time and a vector in isospin space. We have seen in Section 13.3.1 that the modification caused by inclusion of the isospin is only multiplication by an additional factor $\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2$, but the calculation of the pseudoscalar exchange effect needs some algebra. Using Eq. (4.63), $\bar{u}(p') i\gamma^5 u(p)$ becomes in the nonrelativistic limit

$$\begin{aligned} \bar{u}_s(p') i\gamma^5 \boldsymbol{\tau} u_r(p) &= i(\boldsymbol{\tau})_{fi} \xi_s^\dagger \left[\sqrt{(p' \cdot \boldsymbol{\sigma})(p \cdot \boldsymbol{\sigma})} - \sqrt{(p' \cdot \boldsymbol{\sigma})(p \cdot \boldsymbol{\sigma})} \right] \xi_r \\ &\simeq i(\boldsymbol{\tau})_{fi} \xi_s^\dagger \boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{p}') \xi_r = i(\boldsymbol{\tau})_{if} \xi_s^\dagger \boldsymbol{\sigma} \cdot \mathbf{q} \xi_r \end{aligned} \quad (13.126)$$

where $(\boldsymbol{\tau})_{fi}$ is a matrix element in isospin space with f, i denoting the proton or neutron state. The equation shows that γ^5 interaction is effectively a derivative coupling, or that the pion couples with the nucleon predominantly in the P wave as compared to S wave, which is generally dominant in low-energy reactions. This is due to the negative parity of the pion, because in the reaction $N \rightarrow N + \pi$ parity and angular momentum conservation forbids the pion to be produced in the S state. NN interaction is another matter. The scattering amplitude contains a propagator due to exchange of the meson. It has the form

$$T_{fi} \sim \frac{i}{(p_1 - p_3)^2 - \mu^2} = \frac{i}{m_1^2 + m_2^2 - 2(E_1 E_2 - p_1 p_2 \cos \theta) - \mu^2} \quad (13.127)$$

which contains the S wave as well as other waves.¹²⁾

From now on we drop the spin and isospin wave function for simplicity and leave the scattering amplitude in matrix form in the spin and isospin space. The scattering matrix for the pseudoscalar exchange is expressed as

$$T_{fi} = (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) g^2 \frac{(\boldsymbol{\sigma}_1 \cdot \mathbf{q})(\boldsymbol{\sigma}_2 \cdot \mathbf{q})}{q^2 - \mu^2} (\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2) \quad (13.129)$$

12) From Eq. (9.26), the J th partial wave amplitude can be extracted using

$$\frac{S_J(E) - 1}{2i} = \int_{-1}^{+1} d \cos \theta f(\theta) P_J(\theta) \quad (13.128)$$

where the subscripts 1, 2 were attached to distinguish the particles 1 and 2. To gain insight into the properties of the potential, we rewrite the matrix element as

$$\langle p_3 | i g \int d^3 x \bar{\psi}(x) \gamma^5 \boldsymbol{\tau} \psi(x) \boldsymbol{\pi}(x) | p_1, q \rangle \quad (13.130a)$$

$$\begin{aligned} &\simeq i g \int d^3 x e^{-i(p_1-p_3) \cdot x} \boldsymbol{\sigma} \cdot (\mathbf{p}_1 - \mathbf{p}_3) \langle 0 | \boldsymbol{\tau} \cdot \boldsymbol{\pi}(x) | q \rangle \\ &= -e^{-i(E_1-E_3)t_x} g \int d^3 x e^{-i(p_1-p_3) \cdot x} (\boldsymbol{\sigma} \cdot \nabla) \langle 0 | \boldsymbol{\tau} \cdot \boldsymbol{\pi}(x) | q \rangle \end{aligned} \quad (13.130b)$$

where we used a charge-independent interaction and $\simeq$ means a nonrelativistic approximation. We made a partial integration to obtain the last equality. The approximation is equivalent to replacing the Hamiltonian by

$$H_{\text{INT}} = i g \int d^3 x \bar{\psi}(x) \gamma^5 \boldsymbol{\tau} \psi(x) \boldsymbol{\pi}(x) \rightarrow -g \sum_k \psi^\dagger(x) \sigma_k \boldsymbol{\tau} \psi(x) \cdot \partial_k \boldsymbol{\pi}(x) \quad (13.131)$$

where $\psi^\dagger(x)$, $\psi(x)$ in the last equation are two-component fields. This is called the static approximation. Then the scattering matrix is given by

$$\begin{aligned} S_{fi} - \delta_{fi} &= -i T_{fi} = \langle p_3 p_4 | (-i)^2 \int d^4 x H_{\text{INT}}(x) \int d^4 y H_{\text{INT}}(y) | p_1 p_3 \rangle \\ &= -g^2 \int d^4 x d^4 y e^{-i(p_1-p_3) \cdot x - i(p_2-p_4) \cdot y} \\ &\quad \times (\boldsymbol{\sigma}_1 \cdot \nabla_x) (\boldsymbol{\sigma}_2 \cdot \nabla_y) \tau_{1i} \langle 0 | T[\pi_i(x) \pi_j(y)] | 0 \rangle \tau_{2j} \\ &= -g^2 \int d^4 x d^4 y e^{-i(p_1-p_3) \cdot x - i(p_2-p_4) \cdot y} \\ &\quad \times (\boldsymbol{\sigma}_1 \cdot \nabla_x) (\boldsymbol{\sigma}_2 \cdot \nabla_y) \tau_{1i} \left[\int \frac{d^4 q}{(2\pi)^4} e^{-iq \cdot (x-y)} \frac{i \delta_{ij}}{q^2 - \mu^2} \right] \tau_{2j} \end{aligned} \quad (13.132)$$

Separating the CM system $\mathbf{R} = (\mathbf{x} + \mathbf{y})/2$ from the relative motion $\mathbf{r} = \mathbf{x} - \mathbf{y}$ and integrating over $\mathbf{R}$ gives a δ -function for momentum conservation:

$$\begin{aligned} T_{fi} &= (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \int d^3 r e^{-i q r} \Big|_{r=x-y} \\ &\quad \times \left[(\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2) (\boldsymbol{\sigma}_1 \cdot \nabla_x) (\boldsymbol{\sigma}_2 \cdot \nabla_y) \left\{ \frac{1}{(2\pi)^3} \int d^3 q e^{i q \cdot (x-y)} \frac{-g^2}{q^2 + \mu^2} \right\} \right] \end{aligned} \quad (13.133)$$

We have used $E_1 \simeq E_3 \simeq m$ in the denominator of the pion propagator. Then the content of the curly brackets is the static Yukawa potential $[= -g^2 e^{-\mu r}/(4\pi r)]$. The

nonrelativistic potential in space coordinates is given as the content of the square brackets divided by $(2m)^2$ to correct for the relativistic normalization:

$$V(\mathbf{r}) = (\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2)(\boldsymbol{\sigma}_1 \cdot \nabla_x)(\boldsymbol{\sigma}_2 \cdot \nabla_y) \left(-\frac{f^2}{4\pi\mu^2} \frac{e^{-\mu r}}{r} \right) \Big|_{\mathbf{r}=\mathbf{x}-\mathbf{y}} \quad (13.134a)$$

$$f = \frac{\mu}{2m} g \quad (13.134b)$$

Performing the differentiation, we obtain

$$V(\mathbf{r}) = \frac{f^2}{4\pi} \mu (\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2) \left[\frac{1}{3} (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) + \left(\frac{1}{s^2} + \frac{1}{s} + \frac{1}{3} \right) S_{12} \right] Y(s) \quad (13.135a)$$

$$S_{12} = \frac{3(\boldsymbol{\sigma}_1 \cdot \mathbf{s})(\boldsymbol{\sigma}_2 \cdot \mathbf{s})}{s^2} - (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2), \quad Y(s) = \frac{e^{-s}}{s} \quad (13.135b)$$

$$s = \mu r, \quad s = |\mathbf{s}|, \quad \mathbf{r} = \mathbf{x} - \mathbf{y} \quad (13.135c)$$

$\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2$ is called a spin exchange force¹³⁾ and S_{12} a tensor force. S_{12} gives a repulsive force along the spin direction and attractive force perpendicular to it (or vice versa depending on the overall sign), and averages to zero if integrated over all angles. It vanishes for the S wave. therefore we will concentrate on the spin exchange force (the first term).

Problem 13.10

Derive Eq. (13.135) from Eq. (13.134).

We have seen in Section 13.3.1 that $\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2$ gives the isospin exchange force

$$\frac{\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2}{2} = \begin{cases} \frac{1}{2} & I = 1 \\ -\frac{3}{2} & I = 0 \end{cases} \quad (13.136)$$

The same formula holds for the spin functions and gives a positive sign for $S = 1$. The deuteron, a bound state of (pn), has $I = 0$, $S = 1$ and only the even orbital angular momentum $L = 0, 2, \dots$ states are allowed, which is indeed the case. It has a dominant S ($L = 0$) state with a small mixture of D ($L = 2$) state. For $I = 0$, $\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2 < 0$ and the sign of the potential is reversed. Namely, the scalar meson exchange gives a repulsive force and cannot make a bound state if the isospin degree of freedom is taken into account. But if the exchanged meson is a pseudoscalar, Eq. (13.135) gives an attractive potential for the $I = 0$, $S = 1$ state and can explain the deuteron bound state.

The appearance of the spin exchange force ($\sim \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2$) is due to the exchange of the angular momentum carried by a P-wave meson. Similarly, the appearance of the isospin exchange force ($\sim \boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2$) is due to the exchange of isospin carried by the pion. Whenever the force carrier has a quantum number that is described by $SU(N)$, $N \geq 2$ a similar exchange force term appears. The color degree of freedom carried by the gluon induces a color exchange force in $SU(3)$, which will be discussed in detail in the next chapter.

13) For the electromagnetic force, it is a magnetic dipole-dipole interaction.

13.6.2

Effective Range

A highlight of the Yukawa theory was the prediction of the π meson with $m_\pi \simeq 200 m_e$. It was based on the knowledge that the nuclear force is short range $\sim 10^{-13}$ cm. How do we know the range of a force? In the low-energy limit $\delta_l \propto k^{2l+1}$ (see Section 13.5.3), where k is the CM momentum. The S-wave phase shift δ_0 being $\sim k$, we have in the low-energy limit

$$k \cot \delta \simeq -\frac{1}{a} + \frac{1}{2} r_0 k^2 + O(k^3) \quad (13.137)$$

which is referred to as the effective range approximation (see Problem 13.12). Then the scattering amplitude, hence the cross section, can be calculated using Eqs. (13.78) and (13.53). The formula (13.137) is general in the sense that it is insensitive to the potential shape as long as it falls faster than $1/r$. “ a ” is referred to as the scattering length and is a measure of the target size (see Problem 13.11). It is positive (negative) for an attractive (repulsive) force. r_0 is referred to as the effective range and is a measure of the force range. Equation (13.137) means that by measuring low energy scattering data, a and r_0 can be obtained as well as the sign of a .

Problem 13.11

Show that in the limit $k \rightarrow 0$, $\sigma_{\text{TOT}} = 4\pi a^2$.

Problem 13.12

Consider S-wave scattering due to a spherical potential with finite range:

$$V(r) = 0 \quad (r > R), \quad V(r) \neq 0 \quad (r < R) \quad (13.138)$$

Setting $E = k^2/2m$ and the $L = 0$ radial wave function as $S(r)$, we define $u(r) = rS(r)|_{r < R}$, $v(r) = rS(r)|_{r > R}$ and denote the solution at $k = 0$ as $u_0(r)$, $v_0(r)$. Prove the following equations:

$$v_0(r) = 1 - \frac{r}{a} \quad (13.139)$$

$$k \cot \delta = -\frac{1}{a} + \frac{1}{2} r_0 k^2 + O(k^3) \quad (13.140)$$

$$r_0 = 2 \int_0^R dr [v_0^2 - u_0^2] \quad (13.141)$$

Using the potential Eq. (13.135), the properties of the deuterons and low-energy NN scattering behavior were reproduced well by setting

$$\frac{f^2}{4\pi} \simeq 0.08 \quad \text{or} \quad \frac{g^2}{4\pi} \simeq 14 \quad (13.142)$$

which was the original success of the Yukawa theory. However, the coupling constant is ~ 1000 times bigger than the electromagnetic interaction and is too big for perturbational treatment. This was the primary difficulty in constructing the theory of strong interactions. It was only after the quark model was established that a mathematical framework of the strong interaction, i.e. QCD, was developed.

13.7

High-Energy Scattering

13.7.1

Black Sphere Model

As we have seen in Fig. 13.8, πN cross sections in the low-energy region ($T \lesssim 1$ GeV) show a rich structure with peaks and dips. At high energies, however, many resonances overlap and begin to show a smooth structure. Furthermore, contributions from inelastic channels dominate and the fraction of elastic scattering cross section diminishes (Fig. 13.25). Above 10 GeV, the πp total cross section is almost flat, rising slowly ($\sim \log^2 E$) above 100 GeV (Fig. 13.26). The cross section in this region ($10 < E < 100$ GeV) is ~ 25 mb. If it is replaced with the geometrical area of a circle πR^2 , $R = 0.9 \times 10^{-13}$ cm and reproduces the strong (nuclear) force range.

Besides, as can be seen from Fig. 13.26, cross sections of the particle and antiparticle ($\bar{p}$ and p , $\pi^- p$ and $\pi^+ p$, $K^- p$ and $K^+ p$) approach each other. There is an old prediction that their cross sections should agree at high energies, which is known

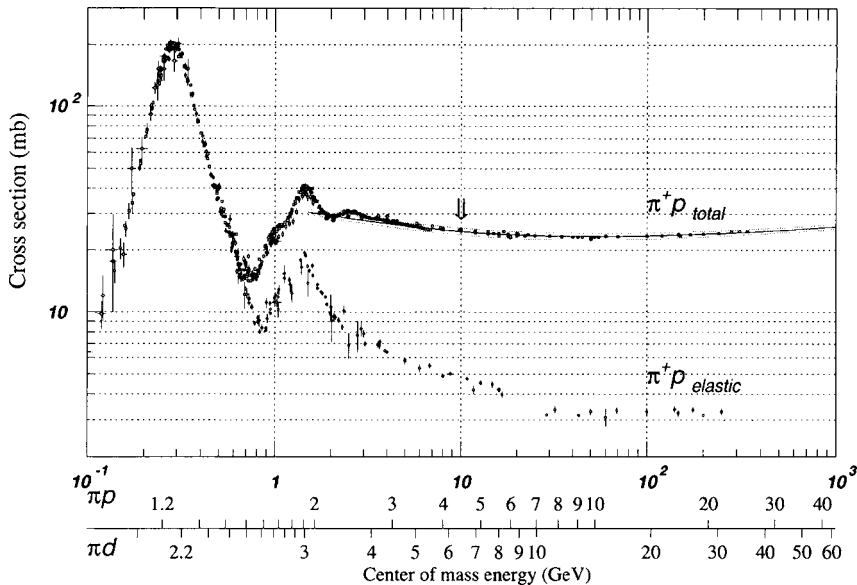


Figure 13.25 $\pi^+ p$ total and elastic cross sections and their behavior at high energy [311].

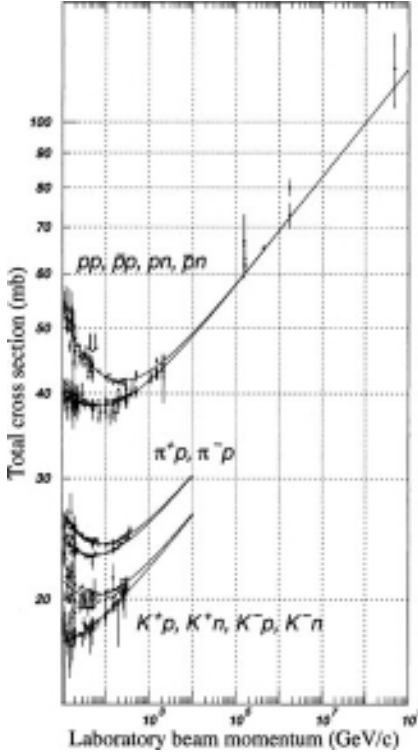


Figure 13.26 High-energy behavior of cross sections of various particles. The rise fits a $\ln^2 E$ dependence [313].

as Pomeranchuk's theorem [302, 325]. Combination with isospin invariance gives

$$\sigma(\pi^- n) = \sigma(\pi^+ p) = \sigma(\pi^- p) \quad (13.143)$$

Qualitatively we may expect that the black sphere approximation reasonably reproduces the high-energy behavior because inelasticity grows as the energy goes higher, becoming totally absorptive in the high-energy limit. The argument goes as follows. Ignoring spin effects, the elastic scattering amplitude can be expanded in terms of partial waves [Eqs. (13.53c), (13.91)]:

$$f(s, \theta) = \sum (2l + 1) f_l(s) P_l(\cos \theta) \quad (13.144a)$$

$$f_l = \frac{\eta_l e^{i\delta_l} - 1}{2ip} \quad (13.144b)$$

If there is no absorption $\eta_l = 1$ and for total absorption $\eta_l = 0$. In the black sphere model, total absorption in all partial waves is assumed ($\eta_l = 0$) and only sums up to $l \leq l_{\max}$ are taken. $l_{\max} \equiv pR$, where R is the force range, is introduced from a semiclassical argument. Then the forward scattering amplitude $f(\theta = 0)$ can be

obtained by inserting $P_l(\cos \theta = 1) = 1$

$$f(s, 0) = \frac{i}{2p} \sum_l^{l_{\max}} (2l+1) \sim \frac{i}{2} p R^2 \quad (13.145)$$

Application of the optical theorem [Eq. (9.56)] gives

$$\sigma_{\text{TOT}} = \frac{4\pi}{p} \text{Im} f(s, 0) = 2\pi R^2 \quad (13.146)$$

Equation (13.146) reproduces the approximate value and constancy of the observed cross section. However, quantum mechanically, the force range is not cut off abruptly at a certain value R but extends beyond its limit. From a purely analytical argument, $l_{\max} \sim \sqrt{s} \ln s$ can be deduced, which sets an upper limit to the total cross section:

$$\sigma_{\text{TOT}} < \text{constant} \times (\ln s)^2 \quad (13.147)$$

This is referred to as the Froissart bound [150]. Observed data above $\sqrt{s} > 100 \text{ GeV}$ seem to saturate and satisfy this limit (Fig. 13.26).

So qualitatively the black sphere model seems to work. However, it cannot reproduce the angular distributions in high-energy elastic scattering. Using approximations valid at small angles $\theta \sim 0$

$$\begin{aligned} t &= -4p^2 \sin^2 \theta / 2 \approx -(p\theta)^2 \\ P_l(\cos \theta) &\approx J_0((l+1/2)\theta) \end{aligned} \quad (13.148)$$

we can calculate

$$\begin{aligned} f(s, 0) &= \frac{i}{2p} \sum_{l=0}^{pR} (2l+1) P_l(\cos \theta) \approx \frac{i}{2p} \int_0^{pR} 2x dx J_0(x\theta) \\ &\approx \frac{iR}{\theta} J_1(pR\theta) = \frac{ipR}{\sqrt{-t}} J_1(R\sqrt{-t}) \end{aligned} \quad (13.149)$$

This is a typical diffraction pattern with dips and peaks.¹⁴⁾ Experimental data give, irrespective of particle species,

$$\frac{d\sigma}{dt} \approx \left[\frac{d\sigma}{dt} \right]_{t=0} e^{at+bt^2} \quad (13.150)$$

where a and b are constants dependent on particle species. This is different from the angular distribution given by the black sphere model.

¹⁴⁾ This is a striped pattern similar to the Fraunhofer diffraction that light waves make through a pinhole.

13.7.2

Regge Trajectory*

To understand the observed angular distributions, let us examine the quantum field theoretical possibilities. Interactions (scatterings) are, in general, mediated by exchange of particles. Let the spin of the exchanged t -channel particle be J . Then the high-energy scattering amplitude mediated by a particle exchange has the form

$$f(s, t) \sim g_J^2 \frac{(-s)^J}{t - m_J^2} \quad s \rightarrow \infty \quad (13.151)$$

It is smooth as a function of s and has the property of becoming small as $|t|$ grows. g_J^2 can be considered as the square of the coupling constant. Indeed, the one-pion exchange model is known to reproduce high-energy small-angle scattering behavior qualitatively. We expand the notion and consider that the scattering amplitude can be expressed as a sum of particle exchange amplitudes of various kinds. As these particles give poles in the complex t -plane, a dominant contribution to the scattering amplitude with given t naturally comes from the nearest pole. For s -channel scattering, t is negative and as far as small-angle scattering ($t \approx 0$) is concerned, the contribution of a particle with the smallest mass, namely that of π , should dominate. By taking into account other heavier particles, we are led to the idea of writing the scattering amplitude as a sum of poles generated by t -channel particles and resonances. Such formalism is naturally realized in the crossed t -channel amplitude introduced in Section 6.4.2. As an example, let us consider the s -channel πp amplitude starting from the t -channel $\pi\pi$ scattering. We expand it in partial waves in the t channel. Note, however, the crossing symmetry can be applied only for Lorentz-invariant amplitudes, so we have to use $A(s, t) = -8\pi\sqrt{s}f(s, t)$ instead of $f(s, t)$ [see Eq. (13.53b)]. The partial wave expansion formula in the t channel is given by

$$A(s, t) = \sum (2l+1)A_l(t)P_l(\cos \theta_t) \quad (13.152a)$$

$$\cos \theta_t = -1 - \frac{2s}{t - 4\mu^2} \quad (13.152b)$$

where μ is the π mass. In order to convert this amplitude to the s -channel πp scattering amplitude, all we need to do is to analytically continue from the $t > \mu^2$, $s < 0$ region to $s > 0$, $t < 0$. However, when $\cos \theta_t$ is considered as a complex variable, the convergence area of the Legendre function $P_l(\cos \theta_t)$ is limited to inside an ellipse tangential to the nearest pole and with foci at $\cos \theta_t = \pm 1$. In order to circumvent the difficulty, Regge regarded the amplitude as a function of complex variable l . In this case, it can be rewritten in a form known as the Watson–Sommerfeld transformation

$$A(s, t) = \frac{i}{2} \int_{C_1} dl (2l+1) A(l, t) \frac{P_l(-\cos \theta_t)}{\sin \pi l} \quad (13.153)$$

The contour C_1 is described in 13.27a. The formula is equivalent to Eqs. (13.152), because $\sin \pi l$ has poles at $l = 0, \pm 1, \pm 2, \dots$ with residue $(-1)^l/\pi$. Then we deform the contour as C_2 in Fig. 13.27b. The properties of the Legendre function

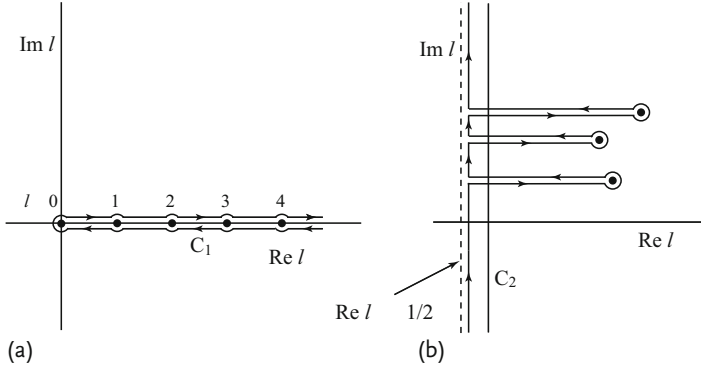


Figure 13.27 (a) Integration contour C_1 in the Watson-Sommerfeld transformation. (b) Deformed integration contour C_2 consists of a few circles around Regge poles and a line along the dashed line $\text{Re}\{l\} = -1/2$.

$P_l(\cos \theta_t)$ are well known. It is convergent as long as the contour C_2 stays to the right of $\text{Re}\{l\} = -1/2$. The problem is the behavior of $A(l, t)$, but Regge showed that for nonrelativistic potentials (superposition of Yukawa potentials) it has at most a few poles, which are referred to as Regge poles. Writing the Regge pole terms as

$$A(l, t) \sim \frac{\beta_n(t)}{l - \alpha_n(t)} \quad (13.154)$$

Eq. (13.153) can be rewritten as

$$\begin{aligned} A(s, t) = & \frac{i}{2} \int_{C_2} dl (2l + 1) A(l, t) \frac{P_l(-\cos \theta_t)}{\sin \pi l} \\ & - \pi \sum_n \frac{(2\alpha_n(t) + 1)\beta_n(t)}{\sin \pi \alpha_n(t)} P_{\alpha_n}(-\cos \theta_t) \end{aligned} \quad (13.155)$$

As l, α_n are complex numbers, convergence is not lost if we go to the region $t < 0$, $s > 0$, and thus analytic continuation is possible.

Properties of Regge poles for potential scattering are well studied. As the energy increases, the pole position of $\alpha_n(t)$ on the $\text{Re}\{l\}$ – $\text{Im}\{l\}$ plane moves with t as depicted in Fig. 13.28a; this is called a Regge trajectory. Starting from $\text{Re}\{l\} = -1/2$, the Regge pole moves to the right on the real axis, departs from it at some value of t (at the threshold $t = 4\mu^2$ in the $\pi\pi$ channel where the $\pi\pi$ scattering becomes physical), goes to the upper right and eventually changes direction to go back to $\text{Re}\{l\} = -1/2$. When $\alpha_n(t)$ is an integer on the real axis, it represents a bound state. After passing the threshold, when it has integer real part and a small imaginary part, it represents a resonance. To see this we calculate the contribution of the Regge pole to the partial-wave amplitude. Using the relation

$$\frac{1}{2} \int_{-1}^1 dz P_l(z) P_\alpha(-z) = \frac{\sin \pi \alpha}{\pi(\alpha - l)(\alpha + l + 1)} \quad (13.156)$$

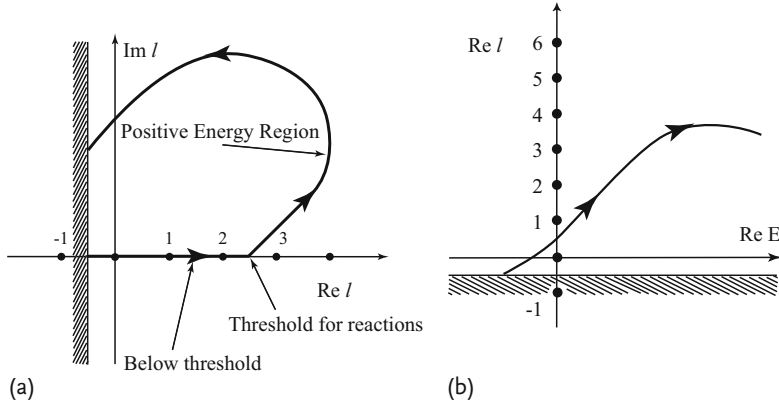


Figure 13.28 A Regge pole moves in the complex l plane as the energy changes; its track is called a Regge trajectory. A pole on the real axis corresponds to a bound state while

$\text{Im}\{l\} > 0$ corresponds to a resonance. (a) The behavior of a Regge trajectory in a nonrelativistic potential. (b) A nonrelativistic Chew-Frautschi plot.

we can obtain

$$\begin{aligned}
 A_l(t) &= \frac{1}{2} \int_{-1}^1 d \cos \theta A(t, \cos \theta) P_l(\cos \theta) \\
 &= \frac{(2\alpha_n + 1)\beta_n(t)}{(\alpha_n(t) - l)(\alpha_n(t) + l + 1)}
 \end{aligned} \tag{13.157}$$

If the Regge pole is located near the real axis, we can expand the above equation in the neighborhood of $\text{Re}\{\alpha_n(t_r)\} = l$ and the partial amplitude becomes

$$A_l \approx \frac{\beta_n(t_r)}{(t - t_r) \left. \frac{d\text{Re}\{\alpha_n(t)\}}{dt} \right|_{t=t_r} + i\text{Im}\{\alpha_n(t_r)\}} \tag{13.158}$$

which is a well-known resonance formula.

A diagram in which $\text{Re}\{l\}$ is displayed as a function of energy is called a Chew-Frautschi plot. For potential scattering (see Fig. 13.28b), $\text{Re}\{l\}$ diminishes at high energy. The field theoretically extended version or phenomenologically obtained plot seems to extend linearly without limit (Fig. 13.29 [98]).

In applications to field theory, we need to take into account exchange forces. For instance, when the t channel is $\pi\pi$, the bosonic nature of π together with isospin independence restricts l to either odd or even numbers. In other words, we have to replace the complex Legendre polynomial by

$$P_{\alpha_n}(\cos \theta) \rightarrow \frac{1}{2} [P_{\alpha_n}(\cos \theta) \pm P_{\alpha_n}(-\cos \theta)] \tag{13.159}$$

For the Regge trajectory including 1^{--} vectors such as ρ, ω, ϕ , the $(-)$ sign has to be adopted and for $0^{++}, 2^{++}$ mesons including f_0, a_2 , the $(+)$ sign has to

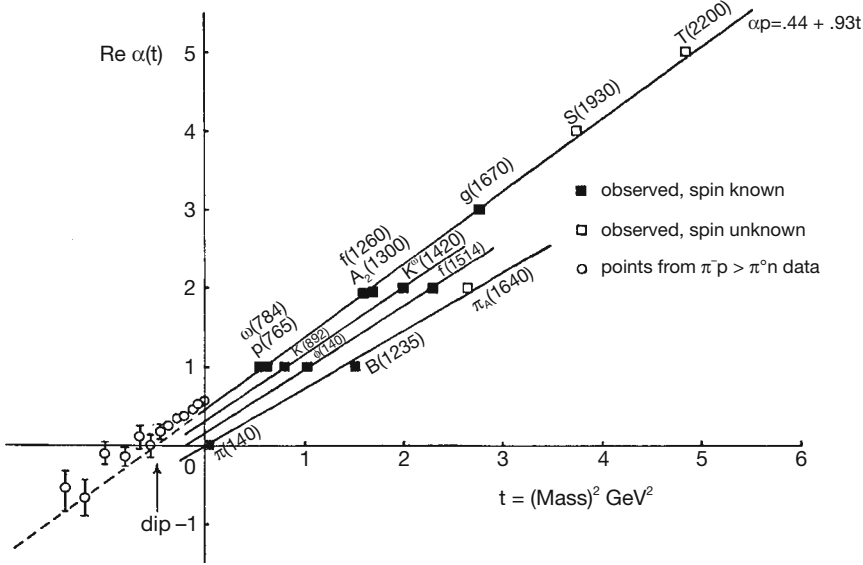


Figure 13.29 A Chew–Frautschi plot. Meson resonances are shown [98]. The data in the negative t region can be obtained by fitting the data to the t -channel particle exchange amplitudes. Circles were obtained by fitting to $\pi^- p \rightarrow \pi^0 n$ data should lie on the extended ρ trajectory.

be chosen. Therefore for one Regge trajectory, resonances appear at every other integer with $\Delta l = 2$. For a parity-degenerate trajectory, $\Delta l = 1$.

Now let us see how the s channel amplitude behaves if analytically continued from the t channel. We are interested in πp scattering for values $s \rightarrow \infty$, $t < 0$. As $P_l(-\cos \theta) \rightarrow s^l$ in the limit $s \rightarrow \infty$, a Regge pole positioned at the extreme right (the highest trajectory in the Chew–Frautschi plot) gives the dominant contribution. Therefore, instead of particles with discrete integer spin, we are talking of a scattering amplitude due to exchange of Regge poles with continuous l value:

$$\begin{aligned}
 A(s, t) &\approx -\pi \sum_n (\alpha_n(t) + 1/2) \beta_n(t) \left[\frac{\tau + e^{-i\pi\alpha_n(t)}}{\sin \pi\alpha_n(t)} \right] P_{\alpha_n}(-\cos \theta_t) \\
 &\Rightarrow -\sum_n \left(\frac{s}{s_0} \right)^{\alpha_n(t)} \gamma_n(t) \left[\frac{\tau + e^{-i\pi\alpha_n(t)}}{\sin \pi\alpha_n(t)} \right]
 \end{aligned} \quad (13.160)$$

where s_0 is a parameter having the same dimension as s and $\alpha_n(t)$ denotes a Regge pole representing bound states or resonances in the t channel. The numerator in the square brackets is a generalization of Eq. (13.159) and, depending on $\tau = \pm$, it is called an even (odd) trajectory. Using $d\sigma/dt \sim |A(s, t)|^2/s^2$, we obtain

$$\frac{d\sigma}{dt} = \left[\frac{d\sigma}{dt} \right]_{t=0} e^{2[\alpha(t)-1]\ln(s/s_0)} \quad (13.161a)$$

$$\left[\frac{d\sigma}{dt} \right]_{t=0} = (1 + r^2) \frac{(\sigma_{\text{TOT}})^2}{16\pi}, \quad r = \frac{\text{Re}\{A\}}{\text{Im}\{A\}} \quad (13.161b)$$

Since we can write $\alpha(t) \approx \alpha(0) + \alpha' t$ in the vicinity of $t \approx 0$, we have reproduced the observed t dependence Eq. (13.150).

In order to have an approximately constant total cross section ($\sigma_{\text{TOT}} \sim -\text{Im}A(t=0)/s$), $A(s, t) \sim s$ is necessary, which, in turn, means $\alpha_P(0) = 1 + \epsilon$, $\epsilon \simeq 0$. If this Regge trajectory has the same quantum number as the vacuum, the particle and antiparticle share a common dominant term, leading to Pomeranchuk's theorem. Therefore, the trajectory α_P is called the Pomeranchuk trajectory or Pomeron.

A t -channel Regge trajectory exchanged in a charge exchange scattering such as $\pi^- p \rightarrow \pi^0 n$ has quantum number $I = 1$. It becomes the ρ resonance at $t = m_\rho^2$ and $l = 1$ [i.e. $\alpha_\rho(m_\rho^2) = 1$], and is called the ρ trajectory. It crosses the point where $\alpha_\rho(0) = 1/2$ (see Fig. 13.29), and makes the cross section behave as s^{-1} .

Other trajectories corresponding to mesons having different quantum numbers, for example, ω ($I = 0, J^P = 1^-$) and K^* ($I = 1/2, J^P = 1^-$) follow more or less the same path.

Crossing between $s \leftrightarrow u$ converts one member to an antiparticle, but its effect on the Regge trajectory is simply an exchange $\cos \theta_t \leftrightarrow -\cos \theta_t$ and it changes it only by a multiplicative factor τ , as can be seen from Eq. (13.160). Consequently, by taking the difference of the cross sections of the particle and antiparticle [for instance, $\sigma(\pi^- p) - \sigma(\pi^+ p)$], we can extract the Regge trajectory having $\tau = -$ (here ρ trajectory) and by taking the sum of the two we can separate that with $\tau = +$ (Pomeron, f_0 , etc.). If we assume the scattering amplitude takes the form given by Eq. (13.151), the residue of a Regge pole exchanged in the reaction $A + B \rightarrow A + B$ can be written as a product $\gamma_n \rightarrow \gamma_{\text{RAA}} \gamma_{\text{RBB}}$ (factorization). Here, $\gamma_{\text{RAA}}, \gamma_{\text{RBB}}$ are quantities proportional to the coupling constant of the Regge trajectory with the particle A or B, and can be related by $SU(3)$ symmetry, as explained in the next chapter. Thus all hadronic total cross sections may be expressed by a small number of parametrized Regge poles. Furthermore, in the t region below threshold, the phase of the Regge trajectories depends only on the factor

$$S = -\frac{\tau + e^{-i\pi\alpha(t)}}{\sin \pi\alpha(t)} \quad (13.162)$$

and we have

$$\text{Re}[f(s, 0)] = \tan \left[\frac{\pi}{2} \alpha(0) \right] \times \text{Im}\{f(s, 0)\} \quad \tau = - \quad (13.163a)$$

$$= -\cot \left[\frac{\pi}{2} \alpha(0) \right] \times \text{Im}\{f(s, 0)\} \quad \tau = + \quad (13.163b)$$

Using parameters obtained by fitting the formula to the data, we can predict the real part of the forward scattering amplitude. As trajectories other than the Pomeron contribute to the real part, and dominant trajectories (ρ and f_0 , etc.) among them have $\alpha(0) = 1/2$, we expect the real part to decrease as $\sim s^{-0.5}$ at high energies. Experimentally, the real part can be determined by measuring the interference effect of the strong scattering amplitude with the Coulomb scattering; the result confirms our prediction [45, 46].

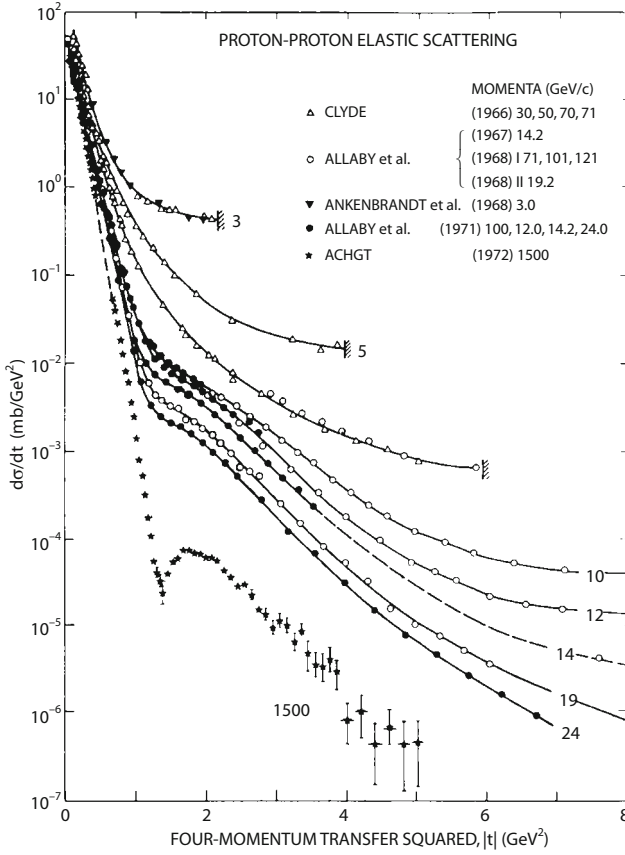


Figure 13.30 pp elastic cross sections as a function of $|t|$ for a variety of incident momenta [183]. At high energy, a diffraction peak appears, which shrinks as the energy increases.

Phenomenologically, the behavior of the high-energy total cross sections can be well reproduced by using two effective Regge poles, the Pomeron and Reggeon [312]. The latter is an effective Regge trajectory with ρ , ω , f , etc. combined:¹⁵⁾

$$\begin{aligned}\sigma^{ab} &= \sigma^{\bar{a}b} = X^{ab}s^\epsilon + Y^{ab}s^{-\eta} \\ \epsilon &= \alpha_P(0) - 1 = 0.093 \pm 0.002 \\ \eta &= 1 - \alpha_R(0) = 0.560 \pm 0.017\end{aligned}\quad (13.164)$$

where s is in GeV^2 . X^{ab} and Y^{ab} are parameters that depend on particle species. The first term represents the Pomeron and the second the Reggeon.

15) [312] gives three terms, separating the Reggeon contribution into $C = +1$ and -1 parts, thus reproducing the small differences between the particle (σ^{ab}) and antiparticle ($\sigma^{\bar{a}b}$) cross sections, but with somewhat poor fitting the difference can be neglected. Recently, owing to accumulation of more precise data, the Pomeron contribution has been modified to $Xs^\epsilon \rightarrow Z + B \log^2(s/s_0)^2$ [311].

According to Eq. (13.161), the coefficient of t in the exponent grows with s . This means that the forward scattering peak shrinks as s becomes larger. This is observed in pp scattering (Fig. 13.30), but in many other channels it is not the case. This means either s is not large enough to be in the asymptotic region or a simple approximation with a few Regge poles is not sufficient to reproduce the data.

In summary, the Regge model is very attractive as a way to explain the high-energy behavior and is frequently used in representing the asymptotic behavior of the scattering cross sections. In the language of the Standard Model, the Pomeron is not due to a specific resonance but a dynamical composite effect of multiple gluons. In the Regge representations, a scattering amplitude in terms of t -channel low-energy resonances describes the high-energy part of the s channel, and the converse is also true. Thus, there is a duality in the scattering amplitude. Extending the idea, each scattering amplitude in the crossed channel may be totally expressed as a sum over resonances in s , t , u channels and indeed such a model was created (Veneziano model [378]). The duality in which one scattering amplitude can be considered as both an s and a t channel amplitude induced a (hadronic) string model [288, 329, 358], which, in turn, triggered the modern string models. In this sense, the analysis with Regge trajectories may be said to be an important turning point in the history of particle physics.

14

The Quark Model

Nucleons and pions have been thought of as elementary particles for a long time. It took over 30 years of trial and error to realize they are composites made of quarks. During the process, there were two turning points. One was the discovery of the strange particles, which began in the late 1940s, and the other that of charmed particles in 1974. By the time the charmed particles were discovered, particle taxonomy was fairly advanced and the significance of the discovery was immediately recognized, creating a turbulent but transient transformation of the particle view referred to as “the November revolution”. On the other hand, the characteristics of the strange particles have been a central riddle, something researchers had to pursue for a long time. In retrospect, when the characteristics of the strange particles were understood, the essence of the quark model was already in hand. The process was slow and did not create a fever like that of the charmed particles, but was far more important in its contribution to the progress of particle physics.

14.1

$SU(3)$ Symmetry

Historically, the discovery of strange particles added to isospin another conserved internal quantum number, the strangeness S . Just as isospin ($SU(2)$ symmetry) was introduced when the equality of both the mass and the properties of the proton and neutron was noticed, the near equality ($m_p \simeq m_n \simeq m_A$) of three basic elements of the Sakata model [339] led to $SU(3)$ symmetry, known as IOO symmetry [212] at the time of introduction. Based on the Sakata model and $SU(3)$, the meson octet was successfully explained. The discovery of the Ω particle firmly established the validity of $SU(3)$. However, one misplacement of p, n, A as the three basic elements, hindered the Sakata model and the “eightfold way” took over, which, in turn, led to the quark model. With the discovery of the charmed particles, $SU(3)$ was extended to $SU(4)$, but all the essential features of so-called flavor symmetry are contained in $SU(3)$ and the quark model. In this chapter we explain that static properties of the hadrons, including the mass and magnetic moment, are well reproduced by the quark model. We defer discussions of the dynamic

properties such as the inner structure of the hadron, which can be investigated by scattering experiments, until Chap. 17.

14.1.1

The Discovery of Strange Particles

Around the late 1940s, Rochester and Butler [332], while investigating reactions induced by high-energy cosmic rays, observed strange particles that left V-shaped tracks in cloud chambers (Fig. 14.1).

Some of them were neutral, decaying into $\pi^- p$, or to $\pi^- \pi^+$ (Fig. 14.1a) and others were charged and decayed into $\pi^- n$ or $\pi^\pm \pi^0$ (Fig. 14.1b):

$$\begin{aligned} V_1 &\rightarrow \pi^- + p \\ V_2 &\rightarrow \pi^- + \pi^+ \\ V_3 &\rightarrow \pi^- + n \\ V_4 &\rightarrow \pi^+ + \pi^0 \end{aligned} \quad (14.1)$$

The pion was a new particle when it was discovered, but it was theoretically anticipated as a necessary ingredient of the nuclear force carrier. In contrast, theorists anticipated neither the existence nor the necessity of the “V” particles. Namely, they could not understand their properties. The most outstanding puzzle was that they were produced frequently in the cloud chamber, meaning the production was a fast process, but they left visible tracks, which is a slow process. The fact that the production was fast meant it was induced by the strong force. Then if the “V” particle were a strongly interacting particle, its decay time should be 10^{-24} – 10^{-22} s (sometimes referred to as the strong time scale) and it could not leave visible tracks. Visible tracks as long as a few centimeters mean a decay time of 10^{-10} – 10^{-8} s, which is a typical weak interaction time scale. The riddle was that a particle interacted strongly at one time and weakly at another. Then it was realized that all the “V” particles were produced in association with another “V” particle [310]. Nakano and Nishijima [286] and Gell-Mann [159] independently proposed a new quantum number “strangeness S ” to explain the controversy and that it is conserved in the strong interaction but not in the weak interaction.

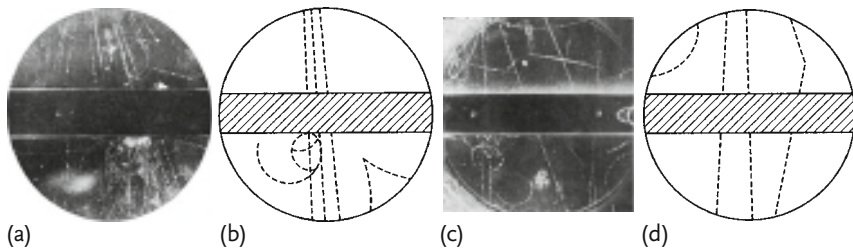


Figure 14.1 Production of new particles in the Wilson cloud chamber. Among the many showers penetrating through a thick lead absorber, a V-shaped track (a,b) and a kink (c,d) are observed [332].

Let us rename V_1, V_2 as Λ, K^0 and assume they have strangeness $S = -1$ and $S = +1$, respectively. All the hitherto known particles, including π, N , are assumed to have $S = 0$. Then the strangeness number in the production reaction can be written as

$$\begin{array}{ccccccc} \pi^- + & p & \rightarrow & K^0 & + & \Lambda & \\ S : & 0 & & 0 & & +1 & -1 \end{array} \quad (14.2)$$

S is conserved before and after the reaction. On the other hand, in the decays

$$\begin{array}{ccccccc} \Lambda & \rightarrow & \pi^- & + & p & & \\ S; & -1 & & 0 & & 0 & \\ K^0 & \rightarrow & \pi^+ & + & \pi^- & & \\ S; & +1 & & 0 & & 0 & \end{array} \quad (14.3)$$

S is not conserved and the strong interaction is forbidden. Similarly, if V_3, V_4 are Σ^-, K^+ , the reaction becomes

$$\pi^- + p \rightarrow K^+ + \Sigma^- \quad (14.4)$$

The reaction is clearly visualized in a bubble chamber picture exposed to the Bevatron beam at LBL (Lawrence Berkeley Laboratory) (Fig. 14.2).

Interestingly, these strange particles have partners with different electric charge but with nearly the same mass. Namely they also form isospin multiplets, just like nucleons and pions. Nishijima and Gell-Mann proposed a relation between the strangeness S , the electric charge Q (in units of the proton charge), and the third component of the isospin:

$$\begin{aligned} Q &= I_3 + \frac{B + S}{2} \equiv I_3 + \frac{Y}{2} \\ Y &= B + S \end{aligned} \quad (14.5)$$

Here, B is the baryon number and Y is called the hypercharge. This is a simple and yet very powerful rule that is used to predict missing partners of isospin multiplets. The law applies to all the hadrons discovered later. Table 14.1 lists a sample of hadrons. Particles with $B = 0$ are called mesons, $B = 1$ baryons and $B = 1, S \neq 0$

Table 14.1 Quantum numbers of representative hadrons.

B (baryon number)	0	0	1
J^P	0^-	1^-	$(1/2)^+$
$Y = 1$	(K^+, K^0)	(K^{*+}, K^{*0})	(p, n)
$Y = 0$	(π^+, π^0, π^-)	(ρ^+, ρ^0, ρ^-)	$(\Sigma^+, \Sigma^0, \Sigma^-)$
	η, η'	ω, ϕ	Λ
$Y = -1$	$(\bar{K}^0, K^-)$	$(\bar{K}^{*0}, K^{*-})$	(Ξ^0, Ξ^-)

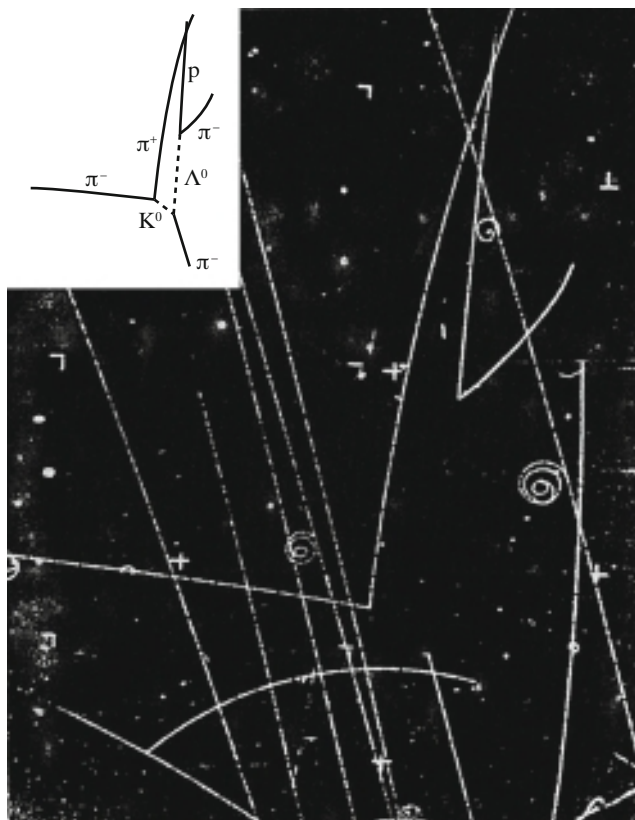


Figure 14.2 A bubble chamber picture taken at the 6 GeV Bevatron/LBL. $\pi^- + p \rightarrow K^0 + \Lambda$, $K^0 \rightarrow \pi^- + \pi^+$, $\Lambda \rightarrow \pi^- + p$. Photo LBL [254]. A bubble chamber is a vessel filled with a superheated transparent liquid (hydrogen, propane). Ions made by a passing charged particle act as evaporation centers,

around which bubbles form, creating a track. The mechanism is similar to that of the cloud chamber with gas and mist replaced by liquid and bubbles. The high density of the liquid makes the bubble chamber a very effective particle detector. It was invented by Glaser in 1952 [172].

hyperons. All the hundreds of hadrons so far discovered satisfy the Nishijima–Gell-Mann law.

In the 1960s, after the Cosmotron (3 GeV) at BNL and the Bevatron (6 GeV) at LBL began operation, detailed measurements of KN scattering and systematic searches for KN , KKN , πY , KY channels in multiparticle productions were carried out, clarifying the existence and properties of $S = 0, -1, -2$ particles and resonances. We have explained part of the story already in the previous chapter (Fig. 13.14–Fig. 13.19). Hundreds of hadrons were thus discovered. Then, one naturally raises the question of whether they are really elementary. If they are considered as composites of more fundamental particles, what properties should these have?

14.1.2

The Sakata Model

Hadrons have a variety of spins and isospins ($s, I = 0, 1/2, 1, 3/2, \dots$). As any spin (or isospin) can be constructed from $s = 1/2$ ($I = 1/2$) bases, the fundamental particles must comprise $s = I = 1/2$ particles, which requires at least a pair of $s = 1/2$ particles with different electric charge. Conversely, two kinds of species with this attribute can compose any value of spin s and isospin I . But to have the strangeness S , a third fundamental particle with nonzero S is necessary. The Sakata [339] model chose (p, n, Λ) as the three fundamental particles. Noticing the nearly same value of the mass of these particles, Ikeda, Ohnuki and Ogawa [212] proposed that the strong interaction is invariant under the exchange $p \leftrightarrow n \leftrightarrow \Lambda \leftrightarrow p$ (IOO symmetry). This is the introduction of $SU(3)$.¹⁾ However, just as isospin conservation ($SU(2)$ symmetry) was an approximate symmetry, $SU(3)$ symmetry (sometimes called unitary spin) is even more so. The mass difference between p and n is $\sim 0.1\%$ but that between the nucleon and Λ is $\sim 20\%$. $SU(3)$ symmetry is considered valid only within this accuracy.

Let us define three basis vectors of the $SU(3)$ group ψ and write them in matrix form:²⁾

$$\psi = \begin{bmatrix} p \\ n \\ \Lambda \end{bmatrix} \equiv \begin{bmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{bmatrix} = \xi^i e_i \quad (14.6)$$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

A triplet (p, n, Λ) represented as a 3×1 column vector is a three-dimensional representation of $SU(3)$ having $(I_3, Y) = (1/2, 1), (-1/2, 1), (0, 0)$ and is denoted as $\mathbf{3}$. The antiparticles can be represented as a 1×3 row vector $(\bar{p}, \bar{n}, \bar{\Lambda}) \equiv (\xi_1, \xi_2, \xi_3)$. The antitriplet has quantum numbers opposite to $\mathbf{3}$, i.e. $(I_3, Y) = (-1/2, -1), (1/2, -1), (0, 0)$ and belongs to $\mathbf{3}^*$.³⁾ Denoting U as a unitary uni-modular ($\det U = 1$) 3×3 matrix, the $SU(3)$ transformation can be expressed in terms of eight $(= 3^2 - 1)$ hermitian traceless matrices λ_i as

$$U = \exp \left(-\frac{i}{2} \sum_{i=1}^8 \lambda_i \theta_i \right) \quad (14.7)$$

$$\text{Tr}(\lambda_i) = 0$$

- 1) To be exact, this is flavor $SU(3)$, to be distinguished from color $SU(3)$ to be described later. The latter is the symmetry of the color charge that the quark has.
- 2) For the definition of groups and terminologies, refer to Appendices A and G.
- 3) Note $\mathbf{3}^* \neq \mathbf{3}$ in $SU(3)$, but $\mathbf{2}^* = \mathbf{2}$ in $SU(2)$, because they occupy the same point on the

I_3 map. For $SU(2)$, the quantum numbers of the antidoublet are the same as those of the basis doublet. Hence by a suitable transformation $(p, n) \rightarrow (\bar{n}, -\bar{p})$, they can be made to belong to the same $I = 1/2$ representation. [See Eq. 9.220]

(where θ_i are eight independent real numbers and the factor $1/2$ in the exponent is a historical convention) just as the $SU(2)$ transformation matrix is expressed in terms of three ($= 2^2 - 1$) Pauli matrices. For λ_i , the convention is to adopt Gell-Mann matrices, which are defined as

$$\lambda_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \lambda_2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \lambda_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (14.8a)$$

$$\lambda_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \quad (14.8b)$$

$$\lambda_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \quad (14.8c)$$

The number of diagonal matrices is two, which is the rank of $SU(3)$. The eight λ_i 's satisfy the following commutation relations:

$$\left[\frac{\lambda_i}{2}, \frac{\lambda_j}{2} \right] = i f_{ijk} \frac{\lambda_k}{2} \quad (14.9a)$$

$$\left\{ \frac{\lambda_i}{2}, \frac{\lambda_j}{2} \right\} = \frac{1}{3} \delta_{ij} + i d_{ijk} \frac{\lambda_k}{2} \quad (14.9b)$$

f_{ijk} is called the structure constant of $SU(3)$, where $\{a, b\} = ab + ba$ and f_{ijk} , d_{ijk} are totally antisymmetric or symmetric with respect to ijk permutations and have the values listed in Table 14.2.

Table 14.2 Structure constants of $SU(3)$.

$f_{123} = 1$
$f_{147} = f_{246} = f_{257} = f_{345} = f_{516} = f_{637} = \frac{1}{2}$
$f_{458} = f_{678} = \frac{\sqrt{3}}{2}$
$d_{118} = d_{228} = d_{338} = -d_{888} = \frac{1}{\sqrt{3}}$
$d_{146} = d_{157} = d_{256} = d_{344} = d_{355} = \frac{1}{2}$
$d_{247} = d_{366} = d_{377} = -\frac{1}{2}$
$d_{448} = d_{558} = d_{668} = d_{779} = -\frac{1}{2\sqrt{3}}$

14.1.3

Meson Nonets

Composites of a baryon and an antibaryon have the baryon number $B = 0$, hence they can be identified as mesons. Since under $SU(3)$ transformation

$$\psi \rightarrow \psi' = U\psi, \quad \bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} U^{\dagger 4)} \quad (14.10)$$

the combination $\psi^\dagger \psi = \xi_i \xi^i = \bar{p}p + \bar{n}n + \bar{\Lambda}\Lambda$ is invariant and hence

$$\eta' \equiv \frac{1}{\sqrt{3}}(\bar{p}p + \bar{n}n + \bar{\Lambda}\Lambda) \quad (14.11)$$

makes an irreducible representation by itself of dimension 1 (singlet). The remaining eight elements made up of the combinations

$$\begin{aligned} P_j^i &= \xi^i \xi_j - (1/3)\text{Tr}(\xi^i \xi_j) \\ &= \begin{bmatrix} \frac{1}{3}(2\bar{p}p - \bar{n}n - \bar{\Lambda}\Lambda) & \bar{n}p & \bar{\Lambda}p \\ \bar{p}n & \frac{1}{3}(2\bar{n}n - \bar{\Lambda}\Lambda - \bar{p}p) & \bar{\Lambda}n \\ \bar{p}\Lambda & \bar{n}\Lambda & \frac{1}{3}(2\bar{\Lambda}\Lambda - \bar{p}p - \bar{n}n) \end{bmatrix} \end{aligned} \quad (14.12)$$

mix with each other by the transformation and constitute the bases of an irreducible representation of dimension 8 (octet). The above process of multiplication and decomposition can be expressed in the simple form

$$3 \otimes 3^* = 1 \oplus 8 \quad (14.13)$$

The process is an extension of $SU(2)$ isospin addition to $SU(3)$. In $SU(2)$, when $I = 1/2$ and $1/2$ are added together, they can be decomposed to $I = 0$ and $I = 1$, which, in the above notation, can be expressed as $2 \otimes 2 = 1 \oplus 3$. By comparing the quantum numbers of the octet members, the observed mesons can be identified as

$$\begin{aligned} \pi^+ &= \xi^1 \xi_2, \quad \pi^0 = -(\xi^1 \xi_1 - \xi^2 \xi_2)/\sqrt{2}, \quad \pi^- = -\xi^2 \xi_1 \\ K^+ &= \xi^1 \xi_3, \quad K^0 = \xi^2 \xi_3, \quad \bar{K}^0 = -\xi^3 \xi_2, \quad K^- = \xi^3 \xi_1 \\ \eta &= (\xi^1 \xi_1 + \xi^2 \xi_2 - 2\xi^3 \xi_3)/\sqrt{6} \end{aligned} \quad (14.14)$$

The negative sign of $\pi^0, \pi^-, \bar{K}^0$ reflects the fact that $(\bar{n}, -\bar{p})$ makes an $I = 1/2$ doublet [see Eq. (9.220)]. The construction of neutral particles is not unique, but considering the isospin is a good quantum number, π^0 is made by applying $I_- = I_1 - iI_2$ to π^+ and η is made to be orthogonal to both π^0 and η' . Using Eq. (14.14), Eq. (14.12) can be rewritten as

$$P_j^i = \begin{bmatrix} -\frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & \pi^+ & K^+ \\ -\pi^- & \frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & K^0 \\ K^- & -\bar{K}^0 & -\sqrt{\frac{2}{3}}\eta \end{bmatrix} \quad (14.15)$$

4) Actually $\bar{\psi}$ means $\psi^\dagger$ as far as $SU(3)$ operation is concerned, but very often it also represents the Lorentz conjugate. So the latter's nomenclature is also used.

Mapping octet members on the I_3 - Y plane gives Fig. 14.3a. Note, in Fig. 14.3 the quark model (u, d, s) is used instead of (p, n, Λ) for the sake of later discussions. Meson octets have the same (I_3, Y) assignments in both the Sakata and quark models. Such a diagram is called a weight diagram. They reproduce a spectrum of observed spin = 0 mesons, although when it was first made, η and η' had not yet been discovered. For the spin = 1 members, replacements $\pi \rightarrow \rho$, $K \rightarrow K^*$, $\eta, \eta' \rightarrow \omega, \phi$ make the weight diagram as shown in Fig. 14.3b. Thus, the Sakata model prepared a mathematical tool $SU(3)$ and obtained the correct meson assignment, but because it put aside (p, n, Λ) as a special trio, it failed to reproduce the baryon octet correctly (see Fig. 14.4b). The correct assignments were made by the eightfold way of Gell-Mann [160] and Ne'eman [294]. From then on the Sakata

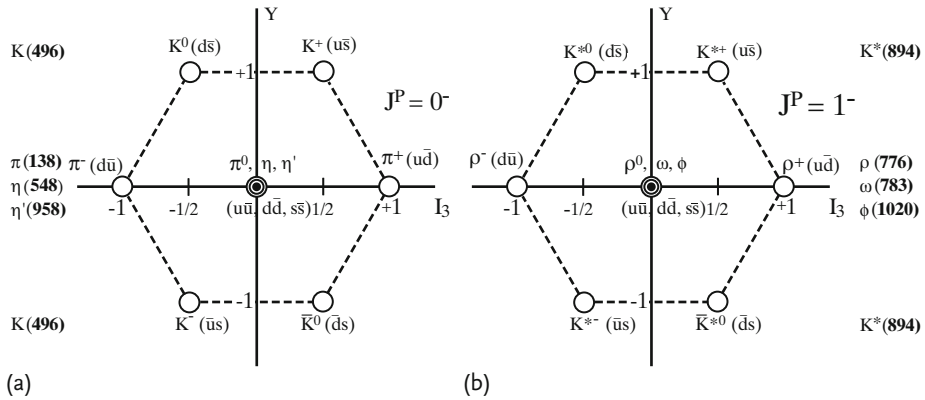


Figure 14.3 Nonet = singlet + octet made of (u, d, s) and ($\bar{u}, \bar{d}, \bar{s}$). Numbers in parentheses are mass values in MeV averaged within the isospin multiplet. (a) Correspondence between observed 0^- mesons and quark-antiquark pair combinations. (b) Same as (a) for 1^- mesons.

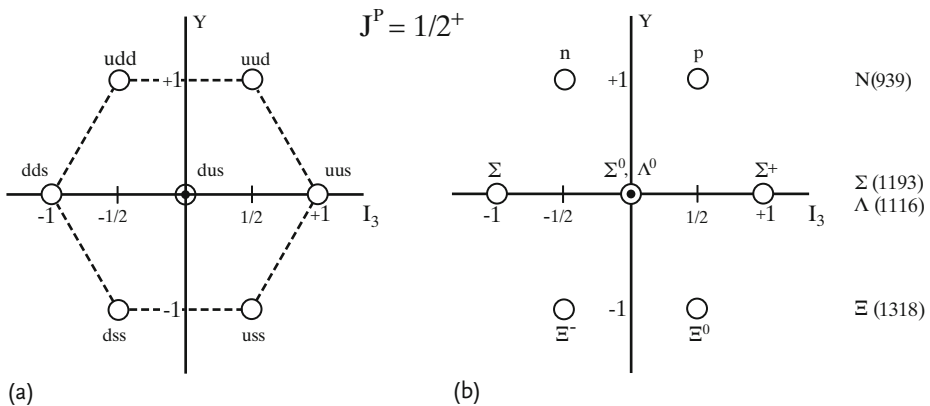


Figure 14.4 (a) An octet made of three quarks. (b) Observed $(1/2)^+$ baryons.

model was superseded by the quark model. The arguments for the meson we have developed so far are equally valid for the quark model, too.

14.1.4

The Quark Model

The eightfold way confers a more fundamental role on the octet since, experimentally, baryons with $J^P = (1/2)^+$ as well as mesons were observed as octets. If hypercharge $Y = S + B$ is adopted instead of the strangeness S as in the Sakata model, both baryons and mesons can be treated in a unified way. Though the eightfold way could reproduce the observed multiplets correctly, it lacked logical justification of why $SU(3)$ works. If one asks for the existence of the fundamental particles as bases of the octet, one needs three bases. They are similar to p, n, Λ of the Sakata model but have slightly different quantum numbers and are referred to as u, d, s quarks. If they have baryon numbers, $3^* \otimes 3 = 1 \oplus 8$ can reproduce meson octets. But to realize an octet with baryon number 1, a different combination has to be introduced. As it happens, a product of three 3's can make an octet that can be identified with the baryon octet. If that is so, the baryon number of the quarks must be $1/3$. Now one uses the same assignment of the Sakata model except for the baryon number and requires that Nishijima–Gell-Mann's law Eq. (14.5) applies to quarks, too. Then their quantum number can be assigned as in Table 14.3.

According to Table 14.3, the electric charge of the quarks would become $2/3$ or $-1/3$, not integers. This was an extraordinary concept at that time and even the proposer Gell-Mann [161] claimed for a while that the quarks were not real but should be considered as a convenient mathematical tool. The other proposer, Zweig [400], is said to have thought them real. According to the modern interpretation, all the strong interactions originate from color charges that all the quarks carry and their strength does not depend on the quark species. As $m_u \sim m_d \sim m_p/3 \sim 330$ MeV, $m_s \sim m_u + (m_\Lambda - m_p) \sim 500$ MeV can be inferred from phenomenological considerations [see Eq. (14.57)], it is clear that the isospin [$SU(2)$] and unitary spin [$SU(3)$] are approximate symmetries whose validity is limited to phenomena where the mass difference is negligible on the energy scale in question. Namely, $SU(3)$ symmetry under exchange of u, d, s is rather accidental. As the quark species $u, d, s, \dots$ are said to differ in “flavor”, this symmetry is now referred to as flavor $SU(3)$ to distinguish it from color $SU(3)$, which plays a fundamental role in the Standard

Table 14.3 Quantum numbers of the quarks.

Name	spin	I_3	S	B	Y	Q
u (up)	$1/2$	$1/2$	0	$1/3$	$1/3$	$2/3$
d (down)	$1/2$	$-1/2$	0	$1/3$	$1/3$	$-1/3$
s (strange)	$1/2$	0	-1	$1/3$	$-2/3$	$-1/3$

Model. Historically, however, flavor $SU(3)$ was instrumental in reaching the quark structure of hadrons.

14.1.5

Baryon Multiplets

To construct baryons, let us first look at the transformation property of nine products of two quarks $T^{ij} = \xi^i \xi^j$. They are not irreducible, but if they are divided into symmetric and antisymmetric parts

$$T^{ij} = \frac{1}{2}(\xi^i \xi^j + \xi^j \xi^i) + \frac{1}{2}(\xi^i \xi^j - \xi^j \xi^i) = A^{ij} + S^{ij} \quad (14.16)$$

each part transforms to itself because the symmetry property is conserved in the transformation (Problem 14.1). In other words, nine product tensors T^{ij} are separated into three antisymmetric and six symmetric tensors:

$$3 \otimes 3 = 3^* \oplus 6 \quad (14.17)$$

Why the three A^{ij} 's belong to 3^* and not to 3 is clearly seen by looking at the quantum numbers they possess. Alternatively, the transformation property of $\eta_k \equiv 2A_{ij} = \epsilon_{ijk} \xi^i \xi^j$ can be shown to be the same as that of $(\xi^k)^*$ (Problem 14.2). Multiplying Eq. (14.17) by one more 3 , the first term on the rhs of Eq. (14.17) gives

$$3^* \otimes 3 = 1 \oplus 8 \quad (14.18)$$

just like the meson nonets. However, unlike the meson singlet, 1 here is a totally antisymmetric combination of A^{ij} and ξ^k , i.e. $\psi_1 = \epsilon_{ijk} \xi^i \xi^j \xi^k$ (Problem 15.3), and the remaining eight tensors become an octet. From the second term 6 , 18 tensors can be made. Among them, ten are totally symmetric and constitute a decuplet whereas the remaining eight tensors make another octet. Namely we have obtained

$$6 \otimes 3 = 8 \oplus 10 \quad (14.19)$$

In summary, we have derived

$$3 \otimes 3 \otimes 3 = 1_A \oplus 8_{MA} \oplus 8_{MS} \oplus 10_S \quad (14.20)$$

Among them, 1 is totally antisymmetric and 10 is totally symmetric with respect to three indices. The two octets, 8_{MA} , 8_{MS} occupy the same position on the I_3 - Y plane, but as can be seen from the way they are constructed, they have mixed symmetry, one antisymmetric and the other symmetric in the first two indices. Experimentally, $(1/2)^+$ baryons constitute 8 (Fig. 14.4) and $(3/2)^+$ baryons constitute 10 (Fig. 14.5). This means the real baryon $(1/2)^+$ octet is a combination of the two octets with mixed symmetry.

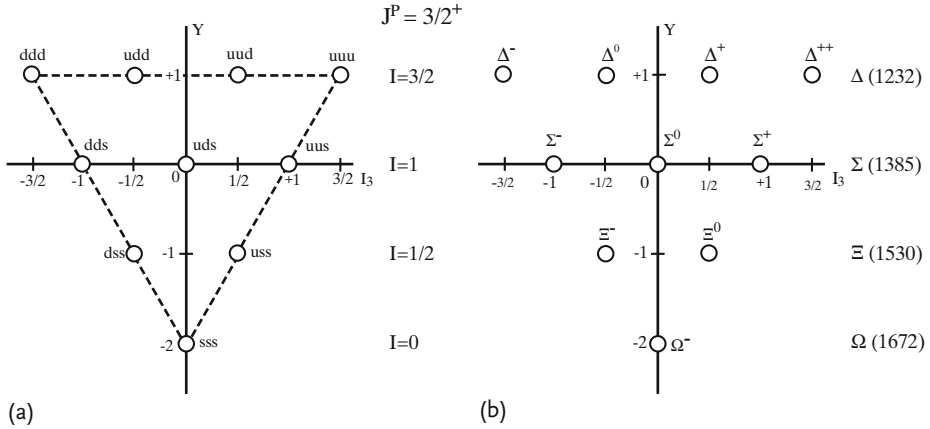


Figure 14.5 (a) 10 can be constructed as totally symmetric combinations of three quarks. (b) Observed spectrum of $(3/2)^+$ resonances. The numbers on the right denote their mass in MeV.

Problem 14.1

Show that the symmetry property of Eq. (14.16) does not change by transformation. Namely what was (anti)symmetric before is also (anti)symmetric after the transformation.

Problem 14.2

Show that A^{ij} in Eq. (14.16) belongs to 3^* in two ways: 1) by looking at the quantum number and 2) by looking at its transformation property.

Problem 14.3

Show that a totally antisymmetric wave function $\psi_1 = \epsilon_{ijk} \xi^i \xi^j \xi^k$ is invariant under $SU(3)$ transformation and hence constitutes a singlet.

14.1.6

General Rules for Composing Multiplets

In order to understand the basic rules for constructing multiplets from basis vectors, we have shown some simple examples. When the dimension of the multiplets increases, the method by hand like the one we have used rapidly becomes complex. There are general formulae in $[SU(2)]$ for constructing any value of angular momentum J out of spin $1/2$ bases using the Clebsch–Gordan coefficients. There is a similar formula in $SU(3)$ and tables of the corresponding Clebsch–Gordan coefficients are available (see for instance [311]). We refer to Appendix G for more

detailed arguments. We restrict arguments here merely to a simple extension of $SU(2)$ to $SU(3)$ and how to use tables derived from general addition rules of $SU(3)$. In $SU(2)$, a multiplet is specified by the angular momentum j .⁵⁾ Elements in a multiplet are specified by the eigenvalue of $J_z = m = (j, j-1, \dots, -j)$. A group consisting of $(2j_1+1)(2j_2+1)$ elements made by multiplying the $2j_1+1$ elements of a multiplet specified by j_1 by the $(2j_2+1)$ elements of another multiplet j_2 generally forms a reducible group. But such a group can be decomposed into a sum of irreducible groups (multiplets) as

$$j_1 \otimes j_2 = (j_1 + j_2) \oplus (j_1 + j_2 - 1) \oplus \dots \oplus (|j_1 - j_2|) \quad (14.21)$$

Elements of the irreducible group belonging to (J, M) can be expressed in terms of Clebsch–Gordan coefficients as (see Appendix E)

$$|JM\rangle = \sum_{m_1, m_2} C_{j_1 m_1, j_2 m_2}^{JM} |j_1 m_1\rangle |j_2 m_2\rangle \quad (14.22)$$

$$m_1 + m_2 = M$$

In $SU(3)$, specifying the dimension alone is not enough, as we saw that $3^* \neq 3$. An additional indicator is necessary. This is because there are two Casimir operators in $SU(3)$. It is convenient to use a pair of numbers (p, q) to specify a multiplet where p, q are effective numbers of quarks and antiquarks⁶⁾. The dimension is given by

$$D(p, q) = \frac{1}{2}(p+1)(q+1)(p+q+2) \quad (14.23)$$

We list some of the correspondences between the notation **1, 3**, etc. and $D(p, q)$:

$$\begin{aligned} D(0, 0) = \mathbf{1}, \quad D(1, 0) = \mathbf{3}, \quad D(0, 1) = \mathbf{3}^*, \quad D(1, 1) = \mathbf{8} = \mathbf{8}^*, \\ D(2, 0) = \mathbf{6}, \quad D(0, 2) = \mathbf{6}^*, \quad D(3, 0) = \mathbf{10}, \quad D(0, 3) = \mathbf{10}^*, \dots \end{aligned} \quad (14.24)$$

The decomposition corresponding to Eq. (14.21) in $SU(2)$ can be made using Young diagrams and expressed as

$$D(p_1, q_1) \otimes D(p_2, q_2) = \sum_n n(p, q) D(p, q) \quad (14.25)$$

$$\text{An example: } \mathbf{8} \otimes \mathbf{8} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \oplus \mathbf{10}^* \oplus \mathbf{27}$$

The difference from $SU(2)$ is that in Eq. (14.21), specification of J is enough to uniquely determine the representation, while in $SU(3)$ there is more than one representation with the same dimension and $n(p, q)$ becomes an integer larger than 1. In the above example, two octets appear with differing symmetry. The composition

- 5) In mathematical language, by the eigenvalues of the Casimir operator, which commutes with all the transformations in the group except the trivial identity operator. In $SU(2)$ it is J^2 with eigenvalue $j(j+1)$.
- 6) An antisymmetric pair of quarks (or antiquarks) are equivalent to an antiquark (or a quark). See arguments following Eq. (14.17) and Appendix G for details.

formula corresponding to Eq. (14.22) is

$$|\mu_\gamma, \nu\rangle = \sum_{\nu_1, \nu_2} \begin{pmatrix} \mu_\gamma & \mu_1 & \mu_2 \\ \nu & \nu_1 & \nu_2 \end{pmatrix} |\mu_1 \nu_1\rangle |\mu_2 \nu_2\rangle \quad (14.26)$$

Here, μ specifies the multiplet expressed in (p, q) and the subscript γ attached to μ differentiates μ with the same dimension. ν denotes different members within a multiplet, usually specified as (Y, I, I_3) with the condition $Y = Y_1 + Y_2$, $I_3 = (I_1)_3 + (I_2)_3$.

Wigner–Eckart Theorem Corresponding to the Wigner–Eckart theorem in $SU(2)$, which is expressed as

$$\langle j_2 m_2 | T(JM) | j_1 m_1 \rangle = \delta_{m_2, M+m_1} C_{j_1 M, j_1 m_1}^{j_2 m_2} \langle j_2 || T^J || j_1 \rangle \quad (14.27)$$

its analogue in $SU(3)$ is given by

$$\langle \mu_2 \nu_2 | T(\mu \nu) | \mu_1 \nu_1 \rangle = \sum_{\gamma} \begin{pmatrix} \mu_\gamma & \mu_1 & \mu_2 \\ \nu & \nu_1 & \nu_2 \end{pmatrix} \langle \mu_2 || T^\mu || \mu_1 \rangle \quad (14.28)$$

where the reduced matrix element $\langle \mu_2 || T^\mu || \mu_1 \rangle$ depends only on the representations involved. The Gell-Mann–Okubo mass formula described in the following section is an application of the Wigner–Eckart theorem.

14.2

Predictions of $SU(3)$

14.2.1

Gell-Mann–Okubo Mass Formula

If $SU(3)$ symmetry were exact, all the quarks would have equal masses ($m_u = m_d = m_s$). Although $m_u \simeq m_d$ is approximately true [$m_p(938 \text{ MeV}) \simeq m_n(939 \text{ MeV})$], $\Delta m = m_s - m_u \simeq m_\Lambda(1116 \text{ MeV}) - m_p(938 \text{ MeV})$ is fairly large. If we consider the mass difference is the origin of $SU(3)$ symmetry breaking, the Hamiltonian mass term would be

$$\begin{aligned} H_m &= m_u(\bar{u}u + \bar{d}d) + m_s\bar{s}s \\ &= \frac{2m_u + m_s}{3}(\bar{u}u + \bar{d}d + \bar{s}s) - \frac{\Delta m}{3}(\bar{u}u + \bar{d}d - 2\bar{s}s) \\ &\equiv m_0\bar{\psi}\psi + \frac{\Delta m}{\sqrt{3}}\bar{\psi}\lambda_8\psi \equiv H_S + H_M S \end{aligned} \quad (14.29)$$

The first term transforms as a singlet of $SU(3)$ and the second as an eighth member ($I = Y = 0$) of an octet and hence is the symmetry-breaking term. Using

Eq. (14.29) Gell-Mann [160] derived a sum rule that holds for particles in an octet (Problem 14.4).

$$\underbrace{\frac{1}{2}(m_N + m_\Xi)}_{1128.5 \text{ MeV}} = \underbrace{\frac{1}{4}(3m_\Lambda + m_\Sigma)}_{1135.0 \text{ MeV}} \quad (14.30a)$$

$$\underbrace{m_K^2}_{0.246 \text{ MeV}^2} = \underbrace{\frac{1}{4}(3m_\eta^2 + m_\pi^2)}_{0.231 \text{ MeV}^2} \quad (14.30b)$$

The agreement with experimental data is very good. Then Okubo, using Eq. (14.29), derived a mass formula applicable to any multiplet [300]:

$$M = a + bY + c \left[I(I+1) - \frac{1}{4}Y^2 \right] \quad (14.31)$$

a, b, c are constants that differ for each multiplet. Equation (14.31) is referred to as the Gell-Mann–Okubo mass formula. The reason for using mass squared rather than mass itself in Eq. (14.30b) is motivated by the Lagrangian mass term of the Klein–Gordon equation. The agreement with experiments is certainly better with the mass squared.

Problem 14.4

Derive Eqs. (14.30) from Eq. (14.31).

Problem 14.5*

Derive Eqs. (14.30) from Eq. (14.29).

14.2.2

Prediction of Ω

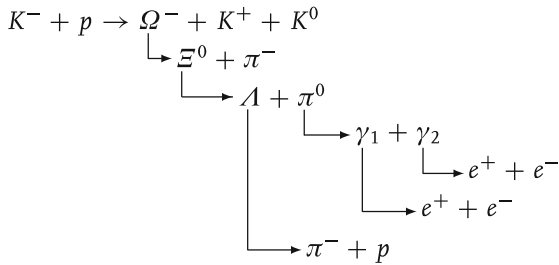
When flavor $SU(3)$ symmetry was proposed, three members of the $(3/2)^+$ decuplet, $\Delta(1232)$, $\Sigma(1385)$, $\Xi(1530)$ were known.⁷ They fit as members of the decuplet. However, the last member Ω was not known at the time. If $SU(3)$ symmetry is correct, Ω with $J^P = (3/2)^+$ should exist and have $I = 0$, $Q = -1$, $S = -3$. Furthermore, applying Eq. (14.31) the following relations are obtained:

$$M(\Omega) - M(\Xi^*) = M(\Xi^*) - M(\Sigma^*) = M(\Sigma^*) - M(\Delta) \quad (14.32)$$

We have attached an asterisk to denote that they are members of the decuplet. Inserting observed values $M(\Delta) = 1232 \text{ MeV}$, $M(\Sigma^*) = 1385 \text{ MeV}$, $M(\Xi^*) =$

⁷ The notation for baryonic resonances with different J^P is as follows. If they have the same (I, Y) as $[N(1/2, 1), \Delta(3/2, 1), \Lambda(0, 0), \Sigma(1, 0), \Xi(1/2, -1) \dots]$, they are expressed as $[N(1440), \Delta(1232), \Lambda(1520), \Sigma(1385), \Xi(1530), \dots]$ with the mass value in parentheses.

1530 MeV, we obtain $M(\Omega) \simeq 1680$ MeV. The Ω mass value is significant. Ω cannot decay by the strong interaction, because a composite baryon with $S = -3$ can be constructed from $(\Xi \bar{K})$, $(\Lambda \bar{K} \bar{K})$, $(N \bar{K} \bar{K} \bar{K})$, but they all have mass values larger than Ω : $M(\Xi \bar{K}) \geq 1809$ MeV, $M(\Lambda \bar{K} \bar{K}) \geq 2107$ MeV, $M(N \bar{K} \bar{K} \bar{K}) \geq 2423$ MeV. Consequently, Ω cannot decay to them because of energy conservation. Decays to $(\Lambda \bar{K})$, $(\Xi \pi)$ are allowed, but only weakly because they have $S = -2$, violating strangeness conservation. As a result of this prediction, a bubble chamber team at BNL launched a search for Ω , and out of 10^5 pictures, they found one with evidence of Ω production (Fig. 14.6). The picture clearly shows the following sequence of reactions:



K^0 is not seen in the picture, but from energy-momentum conservation the missing mass M_{miss} can be calculated; $M_{\text{miss}} = m_K$ indicates that it is a K^0 . The Ω^- leaves a visible track and its lifetime and mass were calculated to be $\sim 10^{-10}$ s and 1686 ± 12 MeV, confirming the $SU(3)$ prediction. This was a result of $SU(3)$ sym-

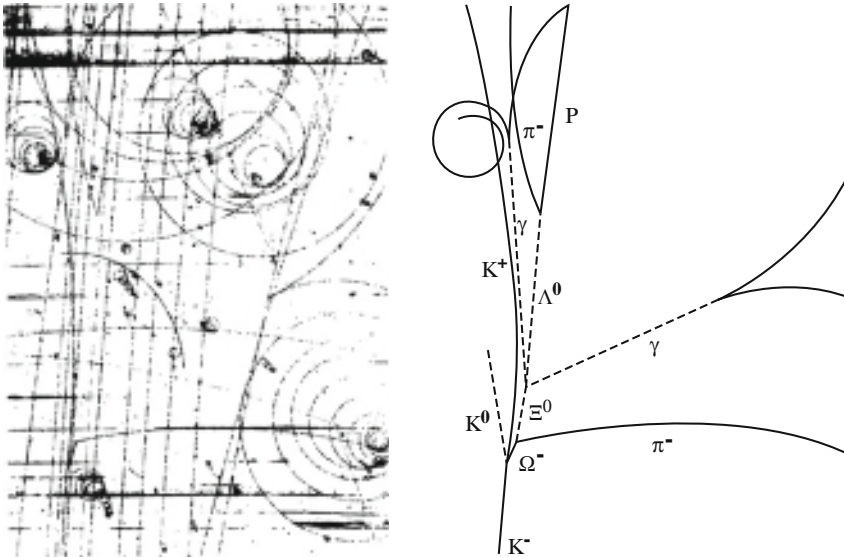


Figure 14.6 A picture of Ω production obtained in a bubble chamber at BNL [49].

metry and not of the quark model, but historically the latter was inspired by the former.

Note that so far resonances that appear in the strong interaction having lifetime $\sim 10^{-23}$ s were thought of as dynamical phenomena, not in the same category as quasistable particles with long lifetime $\gtrsim 10^{-10}$ s. But from the fact that the Ω^- belongs to the same decuplet as other resonances, it is clear that the lifetime does not play any essential role in the classification of particle species. This is analogous to treating all the excited atomic or nuclear energy levels, stable or unstable, on equal footings. As quark composites, it is legitimate to treat hadronic resonances equally with other hadrons such as p and π .

14.2.3

Meson Mixing

As stated before, flavor $SU(3)$ is not a rigorous symmetry and the Hamiltonian contains a symmetry-breaking interaction as in Eq. (14.29). It has the same transformation property as the neutral member of an octet and has nonzero matrix elements between particles belonging to different multiplets but having the same Y and I_3 . For example, it induces mixing between a singlet meson $\phi_0 = (u\bar{u} + d\bar{d} + s\bar{s})/\sqrt{3}$ and the $Y = I_3 = 0$ member of an octet $\phi_8 = (u\bar{u} + d\bar{d} - 2s\bar{s})/\sqrt{6}$. For 1^- mesons, we write

$$\begin{aligned}\phi &= \cos \theta \phi_8 - \sin \theta \phi_0 \\ \omega &= \sin \theta \phi_8 + \cos \theta \phi_0\end{aligned}\tag{14.33}$$

ϕ , ω are observed mesons and are mass eigenstates of a 2×2 matrix M^2 :

$$M^2 \begin{bmatrix} \phi \\ \omega \end{bmatrix} = \begin{bmatrix} m_\phi^2 & 0 \\ 0 & m_\omega^2 \end{bmatrix} \begin{bmatrix} \phi \\ \omega \end{bmatrix}\tag{14.34a}$$

On the other hand, the mass matrix $M^{2'}$ with bases (ϕ_8, ϕ_0) can be expressed as

$$M^{2'} \begin{bmatrix} \phi_8 \\ \phi_0 \end{bmatrix} = \begin{bmatrix} M_8^2 & M_{80}^2 \\ M_{08}^2 & M_0^2 \end{bmatrix} \begin{bmatrix} \phi_8 \\ \phi_0 \end{bmatrix}, \quad M_{80}^2 = M_{08}^2\tag{14.34b}$$

Using $\text{Tr} M^2 = \text{Tr} M^{2'}$, $\det M^2 = \det M^{2'}$, we obtain

$$\begin{aligned}m_\phi^2 + m_\omega^2 &= M_8^2 + M_0^2 \\ m_\phi^2 m_\omega^2 &= M_8^2 M_0^2 - M_{08}^2 M_{80}^2\end{aligned}\tag{14.35}$$

Gell-Mann and Okubo's formula Eq. (14.30), which is valid for no mixing, means

$$M_8^2 = \frac{4m_{K^*}^2 - m_\rho^2}{3}\tag{14.36}$$

Using Eqs. (14.33) and (14.36), the following relations are obtained:

$$\tan^2 \theta = \frac{M_8^2 - m_\phi^2}{m_\omega^2 - M_8^2} = \frac{4m_{K^*}^2 - m_\rho^2 - 3m_\phi^2}{3m_\omega^2 - 4m_{K^*}^2 + m_\rho^2} \quad (14.37a)$$

$$\tan 2\theta = \frac{2M_{08}^2}{M_0^2 - M_8^2} \quad (14.37b)$$

Table 14.4 gives the relevant mass values and mixing angles.

Problem 14.6

Derive Eqs. (14.37).

When $\sin \theta = 1/\sqrt{3} (\simeq 0.577, \theta \simeq 35.3)$,

$$\phi = \sqrt{\frac{2}{3}} \frac{u\bar{u} + d\bar{d} - 2s\bar{s}}{\sqrt{6}} - \sqrt{\frac{1}{3}} \frac{u\bar{u} + d\bar{d} + s\bar{s}}{\sqrt{3}} = -s\bar{s} \quad (14.38a)$$

$$\omega = \sqrt{\frac{1}{3}} \frac{u\bar{u} + d\bar{d} - 2s\bar{s}}{\sqrt{6}} + \sqrt{\frac{2}{3}} \frac{u\bar{u} + d\bar{d} + s\bar{s}}{\sqrt{3}} = \frac{u\bar{u} + d\bar{d}}{\sqrt{2}} \quad (14.38b)$$

This is called ideal mixing. Table 14.4 shows that for vector 1^- and tensor 2^+ mesons the mixing is almost ideal. For the pseudoscalars, the mixing angle is small and negative. This is because $m_{\eta'} > M_{08}$ and is attributed to the chiral anomaly contribution [97, 208, 361], which is an interesting topic theoretically but we do not discuss it here. If ideal mixing is realized, $\phi, f_2' \sim s\bar{s}$, $\omega, f_2 \sim u\bar{u} + d\bar{d}$, the former is expected to decay predominantly to $K\bar{K}$ and the latter to $2\pi, 3\pi$ (see OZI rule below). The observed branching ratios are [311]

$$\begin{aligned} \text{BR}(\phi \rightarrow K\bar{K}) &\simeq 84\%, & \text{BR}(\phi \rightarrow 3\pi) &\simeq 15\% \\ \text{BR}(\omega \rightarrow 3\pi) &\simeq 90\%, \\ \text{BR}(f_2' \rightarrow K\bar{K}) &\simeq 100\%, \\ \text{BR}(f_2 \rightarrow 2\pi + 4\pi) &\simeq 95\%, & \text{BR}(f_2 \rightarrow K\bar{K}) &\simeq 4\% \end{aligned} \quad (14.39)$$

which confirms the expectation.

Table 14.4 Meson mixing angle [311].

	0^-	1^-	2^+
Mass	η'	ω	f_2
[MeV]	η	ϕ	f_2'
	π^0	ρ^0	a_2
	K^0	K^*	K_2^*
M_8	569.3	928.7	1459.6
θ	-11.3°	40.0°	32.1°

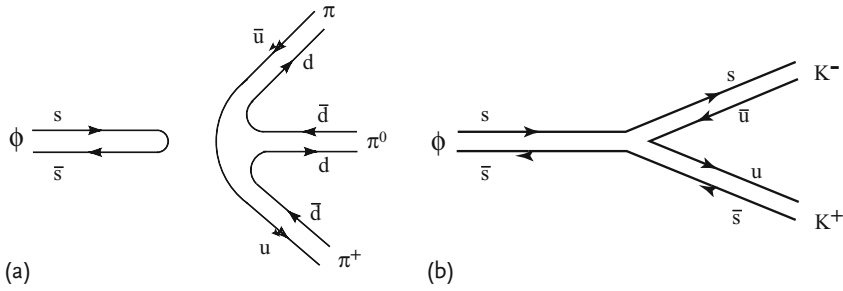


Figure 14.7 (a) $\phi \rightarrow 3\pi$, (b) $\phi \rightarrow K^+K^-, K^0\bar{K}^0$

OZI Rule The suppression of $\phi \rightarrow 3\pi$ and $\omega, f_2 \rightarrow K\bar{K}$ can be explained by the so-called OZI (Okubo–Zweig–Iizuka) rule [211, 301, 400], which claims that decays with disconnected quark diagrams are suppressed. It is represented pictorially in Fig. 14.7.

Since ϕ is predominantly composed of $s\bar{s}$, it can go to $K\bar{K}$ with connected quark lines, whereas the quark lines have to be cut and regenerated to go to 3π 's [note: $\phi \rightarrow 2\pi$ is forbidden by G-parity ($G = (-1)^I C = -1$), see paragraph in Sect. 13.2]. This is an empirical rule. The reason for it was clarified with the advent of QCD (see Sect. 14.5.4).

Problem 14.7

Experimentally, $\sigma(K^-p \rightarrow \phi A) \simeq 100\sigma(\pi^-p \rightarrow \phi n)$. By drawing the quark diagram, show that the OZI rule applies here.

Problem 14.8

The decay process of a vector meson 1^- to lepton pairs $l\bar{l}$ goes through a photon intermediate state and its decay width is proportional to the square of the electric charges $Q^2 = (\sum_i Q_i)^2$

$$16\pi\alpha^2 \langle Q \rangle^2 \frac{|\psi(0)|^2}{m_V^2} \quad (14.40)$$

where α is the fine structure constant, ψ is the wave function of the vector meson and m_V is its mass. Assuming $m_\rho = m_\omega = m_\phi$ and ideal mixing, show the relative ratios of the decays are

$$\Gamma_\rho : \Gamma_\omega : \Gamma_\phi = \frac{1}{2} : \frac{1}{18} : \frac{1}{9} \quad (14.41a)$$

The experimental data are

$$\Gamma_\rho = 7.04 \pm 0.06 \text{ keV}, \quad \Gamma_\omega = 0.60 \pm 0.02 \text{ keV}, \quad \Gamma_\rho = 1.27 \pm 0.04 \text{ keV} \quad (14.41b)$$

and agree fairly well with the expectation.

Problem 14.9

Derive Eq. (14.40)*.

Hint: $V \rightarrow l\bar{l}$ can be considered as $q\bar{q} \rightarrow l\bar{l}$ scattering in which quark flux is prepared by the vector meson (see Fig. 14.8). Use:

1. $\sigma(e^-e^+ \rightarrow \mu^-\mu^+)$ in Eq. (C.17) with appropriate replacement of the quark mass and charge,
2. The quark flux prepared by the meson is $F(q\bar{q}) = 2\beta_q |\psi(0)|^2$, where β_q is the quark velocity in the total center of mass frame energy $\sqrt{s} = m_V$,
3. The result would still be a factor 3 off, which can be corrected by considering the color factor described below.

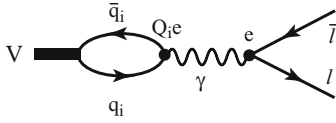


Figure 14.8 A diagram that contributes to $V \rightarrow \ell\bar{\ell}$ decay.

14.3

Color Degrees of Freedom

The success of $SU(3)$ immediately raised a new problem. Ω^- and Δ^{++} are composed of sss and uuu , which are symmetric. Their spin is $3/2$, which means the configuration is $\uparrow\uparrow\uparrow$ and is also symmetric. Since ground states of the quark composites are expected to have orbital angular momentum $L = 0$ (states with $L \geq 1$ are generally at higher energy levels), the particles' wave functions in the decuplet are totally symmetric. This contradicts quarks having spin $1/2$, which should obey Fermi statistics. Therefore, we need to introduce another degree of freedom to make the wave function totally antisymmetric, which we refer to as "color". In order to make three particles totally antisymmetric, the color freedom has to be three. We have already learned that a singlet made from $3 \otimes 3 \otimes 3$ in $SU(3)$ is totally antisymmetric [see arguments following Eq. (14.18)]. We name the three degrees

of color freedom R (red), B (blue) and G (green). The names are arbitrary. We could just as well call them η_1 , η_2 and η_3 or use numbers. However, it is a very convenient name, since all the observed hadrons seem to be color singlets, and can be called white or color neutral figuratively.

Evidence for Color Degrees of Freedom There is a great deal of evidence for the existence of three colors ($N_c = 3$).

1) The spin statistics of Ω^- and Δ^{++} that we have just stated.

2) $\pi^0 \rightarrow \gamma + \gamma$: As the neutral pion π^0 is a bound state of $q\bar{q}$, i.e. $|\pi^0\rangle = (-|u\bar{u}\rangle + |d\bar{d}\rangle)/\sqrt{2}$ [see Eq. (14.14)], the process goes through quark loop intermediate states as depicted in Fig. 14.9. The decay matrix element is the sum of the two diagrams in Fig. 14.9. If the quark has N_c colors, the number of diagrams is N_c and the decay rate has to be multiplied by N_c^2 . The quark model calculation with $N_c = 3$ gives [7, 55]

$$\Gamma(\pi^0 \rightarrow 2\gamma) = N_c^2 (Q_u^2 - Q_d^2)^2 \frac{\alpha^2 m_{\pi^0}^3}{32\pi^3 f_\pi^2} \quad (14.42)$$

where f_π is the pion decay constant defined in Eq. (15.44). Calculation gives $\Gamma(\pi^0 \rightarrow 2\gamma) = 7.64 \text{ eV}$, in good agreement with the observed lifetime $\tau_{\text{exp}} = 8.4 \pm 0.6 \times 10^{-17} \text{ s}$, which gives $\Gamma_{\text{exp}} = \hbar/\tau = 7.84 \pm 0.6 \text{ eV}$. Without color, the theoretical value would be a factor of 9 less than the experimental value.

3) Drell-Yan process ($p + p \rightarrow l\bar{l} + X$): In the $p(\pi)-p$ reactions to make a lepton pair with large invariant mass, the production cross section can be shown to be proportional to N_c (see Sect. 17.9.1) and a comparison with experiment shows $N_c = 3$.

4) R -value: The R -value is defined by

$$R \equiv \frac{\sigma(e^-e^+ \rightarrow \text{all hadrons})}{\sigma(e^-e^+ \rightarrow \mu^-\mu^+)} = N_c \sum_i Q_i^2 \quad (14.43)$$

$$= 3 \times \left[\left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \dots \right]$$

$$= \begin{cases} 2 & u, d, s & \sqrt{s} < 2m_c \\ \frac{10}{3} & u, d, s, c & 2m_c < \sqrt{s} < 2m_b \\ \frac{11}{3} & u, d, s, c, b & 2m_b < \sqrt{s} \end{cases} \quad (14.44)$$

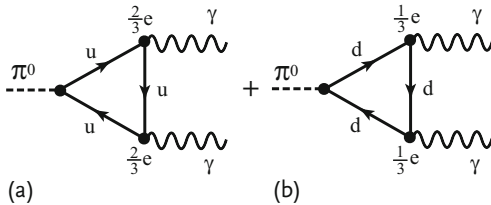


Figure 14.9 Triangular diagrams: $\pi^0 \rightarrow \gamma + \gamma$ goes through quark loops.

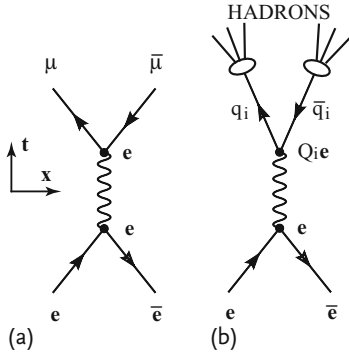


Figure 14.10 One-photon exchange diagrams for $\mu\bar{\mu}$ and hadron production in e^-e^+ reaction. $R = \sigma(e^-e^+ \rightarrow \text{hadrons})/\sigma(e^-e^+ \rightarrow \mu^-\mu^+) = 3 \sum_i Q_i^2$.

Feynman diagrams for $e^-e^+ \rightarrow \mu^-\mu^+$ and $e^-e^+ \rightarrow \text{all hadrons}$ are shown in Fig. 14.10.

Quarks are confined and it means the quarks produced in e^-e^+ reaction are converted 100% to hadrons. From Fig. 14.10 it is clear that production of a quark pair is a pure electromagnetic process and its cross section is proportional to Q_i^2 times that of $\mu\bar{\mu}$ production. Multiplying by N_c and summing over all the kinds of quarks that can be produced gives the R -value.

Figure 14.11 gives a compilation of R -value data. It clearly shows that at the threshold of going over $\sqrt{s} > 2m_q$, a variety of resonances appear, but that the R -value agrees with the predictions at $\sqrt{s} \gg 2m_q$, $q = c, b$ for each energy range with mass values given in Table 2.2.

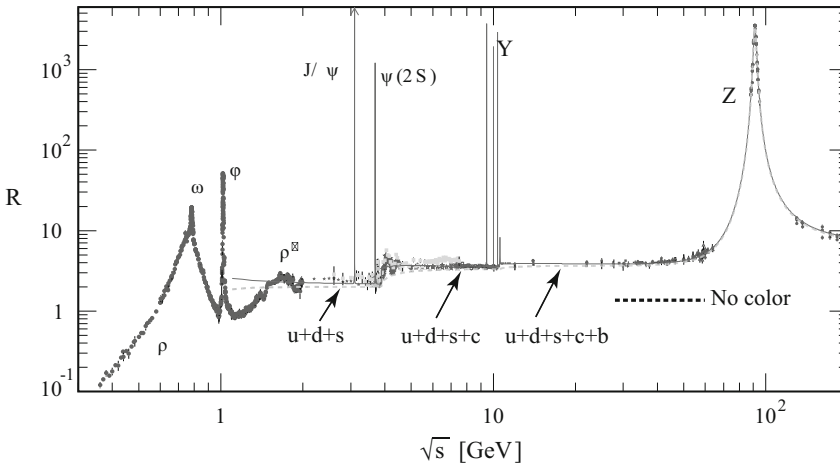


Figure 14.11 Compilation of R -value data [311]. Lines indicated by arrows represent the values 2, $10/3$ and $11/3$, respectively. The dashed line indicates the R -value for $u + d + s + c + b$ with no color. Lines J/ψ , Υ indicate sharp resonances.

5) The weak force carrier $W^\pm$ boson decays into

$$W \rightarrow e\nu_e, \quad \mu\nu_\mu, \quad \tau\nu_\tau, \quad \bar{u}d', \quad \bar{c}s' \quad (14.45)$$

and the quarks hadronize after the decay. d', s' are Cabibbo-rotated d, s (see Sect. 15.6.4). If the masses of the leptons and quarks are neglected compared to $m_W = 81 \text{ GeV}$, the partial decay rates are all the same because the coupling of the W boson is universal. Therefore, the branching ratio of $W \rightarrow e\nu_e$ (or any single leptonic decay) is $1/5$ without color, but $1/9$ if color is included. The experimental value is $\sim 11\%$ and supports the existence of the color degrees of freedom.

14.4

$SU(6)$ Symmetry

14.4.1

Spin and Flavor Combined

$SU(6)$ Multiplets As the quark has color degrees of freedom, a wave function of quark composites can be expressed as a product:

$$\text{Total wave function} = \text{color} \times \text{flavor} \times \text{spin} \times \text{space}$$

Assuming all hadrons are color singlets and observed low-lying hadrons are in the S states of orbital angular momentum, the remaining degrees of freedom are $3 \times 2 = 6$. If the six degrees of freedom in the flavor and spin are equal and if the strength of the interaction does not change by exchange among them, the symmetry to be considered is $SU(6)$. Base vectors are

$$(\xi^1, \xi^2, \dots, \xi^6) = (u \uparrow, u \downarrow, d \uparrow, d \downarrow, s \uparrow, s \downarrow) \quad (14.46)$$

We designate this as $\mathbf{6}$ or, writing flavor and spin separately, as $(3, 1/2)$. The way to make product representations from the above bases is quite similar to that of $SU(3)$. Meson multiplets are (see Appendix G)

$$\mathbf{6} \otimes \mathbf{6}^* = \mathbf{1} \oplus \mathbf{35} = (1, 0) \oplus (1, 1) \oplus (8, 0) \oplus (8, 1) \quad (14.47)$$

which means both mesons with $J^P = 0^-, 1^-$ appear in a singlet and an octet. It is customary to combine them and call them a nonet. Baryons appear in

$$\mathbf{6} \otimes \mathbf{6} \otimes \mathbf{6} = \mathbf{20}_A \oplus \mathbf{70}_{MA} \oplus \mathbf{70}_{MS} \oplus \mathbf{56}_S \quad (14.48)$$

However, the Fermi statistics of the quark allows only totally symmetric $\mathbf{56}_S$ representation. Separating the flavor and the spin, we can re-express it as

$$\mathbf{56} = (8, 1/2) \oplus (10, 3/2) \quad (14.49)$$

Therefore the baryon $(1/2)^+$ octet and the decuplet $(3/2)^+$ exhaust the available spectra. Namely, the baryons, unlike the mesons, do not constitute a nonet. Note, in Sect. 14.1.5 we stated that there are two baryon octets with different symmetry, now we know that combined with spin wave functions they should make a totally symmetric wave function (see Problem 14.11).

Wave Functions in $SU(6)$ Wave functions of the baryons in flavor octet and those of spin are of mixed symmetry. The combined wave functions have to be totally symmetric in the S states. As an example, let us consider $|p\rangle \sim |uud\rangle$. With respect to the first two quarks it is symmetric, and hence the spin wave function of the first two must also be symmetric. A symmetric spin wave function is expressed as

$$|1, 1\rangle = \uparrow\uparrow, \quad |1, 0\rangle = \frac{1}{\sqrt{2}}(\uparrow\downarrow + \downarrow\uparrow), \quad |1, -1\rangle = \downarrow\downarrow \quad (14.50)$$

In order to make a spin $1/2$ wave function from the product of Eq. (14.50) and the third quark, we use the Clebsch–Gordan coefficients to give

$$\begin{aligned} |1/2, 1/2\rangle &= \sqrt{\frac{2}{3}} \uparrow\uparrow\downarrow - \sqrt{\frac{1}{3}} \frac{(\uparrow\downarrow + \downarrow\uparrow) \uparrow}{\sqrt{2}} \\ &= \frac{1}{\sqrt{6}} [2 \uparrow\uparrow\downarrow - (\uparrow\downarrow + \downarrow\uparrow) \uparrow] \equiv \chi_{MS} \end{aligned} \quad (14.51)$$

As the wave function must be totally symmetric, we make a cyclic replacement of the particle $1 \rightarrow 2 \rightarrow 3$ and add them all together. Therefore

$$\begin{aligned} |p \uparrow\rangle &\sim uud[2 \uparrow\uparrow\downarrow - (\uparrow\downarrow + \downarrow\uparrow) \uparrow] + \text{cyclic} \\ &= \frac{1}{\sqrt{18}} (2u \uparrow u \uparrow d \downarrow - u \uparrow u \downarrow d \uparrow - u \downarrow u \uparrow d \uparrow \\ &\quad + 2u \uparrow d \downarrow u \uparrow - u \downarrow d \uparrow u \uparrow - u \uparrow d \uparrow u \downarrow \\ &\quad + 2d \downarrow u \uparrow u \uparrow - d \uparrow u \uparrow u \downarrow - d \uparrow u \downarrow u \uparrow) \end{aligned} \quad (14.52)$$

The normalization constant is determined to make $\langle p|p\rangle = 1$.

Problem 14.10

Show that the same wave function is obtained if one starts from the antisymmetric wave function for the first two quarks

$$\phi_{MA} \chi_{MA} \equiv \frac{(ud - du)}{\sqrt{2}} u \frac{(\uparrow\downarrow - \downarrow\uparrow)}{\sqrt{2}} \uparrow \quad (14.53)$$

Problem 14.11

Using $\phi_{MS} = \frac{1}{\sqrt{6}}\{2uud - (ud + du)u\}$ and χ_{MS} , ϕ_{MA} , χ_{MA} given in Eqs. (14.51) and (14.53), prove that the $SU(6)$ wave function of Eq. (14.52) can be written as

$$|p \uparrow\rangle = \frac{1}{\sqrt{2}} [\phi_{MS} \chi_{MS} + \phi_{MA} \chi_{MA}] \quad (14.54)$$

ϕ_{MS} , ϕ_{MA} are wave functions belonging to $\mathbf{8}_{MS}$, $\mathbf{8}_{MA}$, as is clear from the symmetry. ϕ_{MS} , ϕ_{MA} of other baryons are given in Table G.2.

Magnetic Moment of the Proton The $SU(6)$ wave function is given by Eq. (14.52), but the relative spin direction of u and d is the same if cyclic permutation is performed. Therefore the first term is enough for considering the magnetic moment. The normalization, however, is important and has to be redefined each time. Consequently we consider

$$|p\rangle = uud \frac{1}{\sqrt{6}} [2\uparrow\uparrow\downarrow - (\uparrow\downarrow\uparrow + \downarrow\uparrow\uparrow)] \quad (14.55)$$

The magnetic moment of the proton can be calculated as a vector sum of that of each quark. From the first term, the proton magnetic moment receives a contribution from the quarks $\mu_p = \mu_u + \mu_u - \mu_d$ with probability $(2/\sqrt{6})^2 = 2/3$. From the second term, the proton receives $\mu_p = \mu_u - \mu_u + \mu_d$ with probability $2 \times (1/\sqrt{6})^2 = 1/3$

$$\therefore \mu_p = \frac{2}{3}(2\mu_u - \mu_d) + \frac{1}{3}\mu_d = \frac{1}{3}(4\mu_u - \mu_d) \quad (14.56a)$$

Similarly for the neutron,

$$\mu_n = \frac{1}{3}(4\mu_d - \mu_u) \quad (14.56b)$$

If $SU(6)$ is an exact symmetry, $m_u = m_d$, hence $\mu_d = -\mu_u/2$, which makes $\mu_n/\mu_p = -(2/3)$. The agreement with the observed value (-0.685) is very good and was considered a great success of $SU(6)$ symmetry. Here, we go one step further and calculate all the magnetic moments of the $1/2$, $3/2$ baryons, but regard m_u, m_d, m_s as parameters to be adjusted with the observed values. Table 14.5 lists the magnetic moments of other baryons together with fitted and observed values.

Problem 14.12

Derive the expressions in the second column of Table 14.5.

We see that all nine magnetic moments are well reproduced with only three adjustable variables. The quark masses determined from the above table using $\mu_i = q_i e/2m_i$ are

$$m_u = 338 \text{ MeV}, \quad m_d = 322 \text{ MeV}, \quad m_s = 510 \text{ MeV} \quad (14.57)$$

We see $m_u \simeq m_d \simeq m_p/3$, $m_s - m_u \simeq m_\Lambda - m_p$. They are called the constituent mass (effective static quark mass) in contrast to current mass (see Table 14.5), which is determined by the scattering data. The constituent masses are known to reproduce static properties of the hadron, including the magnetic moment and the mass. The riddle of two different quark masses will be resolved later when we learn about asymptotic freedom (see Sect. 18.5.1). It states that quarks with the current mass are freely flying at almost the speed of light but confined inside a sphere with the effective size of hadrons ($\lesssim 10^{-15}$ m). In describing static properties, however,

Table 14.5 Magnetic moments derived using $SU(6)$ wave function [117].

Baryon	Magnetic moment	Theory	Observed value ^a
p	$(4\mu_u - \mu_d)/3$	input	$2.792847386 \pm (63)$
n	$(4\mu_d - \mu_u)/3$	input	$-1.91304275 \pm (45)$
Λ	μ_s	input	-0.613 ± 0.004
Σ^+	$(4\mu_u - \mu_s)/3$	2.67	2.458 ± 0.01
Σ^0	$(2\mu_u + 2\mu_d - \mu_s)/3$	0.79	–
$\Sigma^0 \rightarrow \Lambda$	$(\mu_d - \mu_u)/\sqrt{3}^b$	–1.63	-1.61 ± 0.08
Σ^-	$(4\mu_d - \mu_s)/3$	–1.09	-1.160 ± 0.025
Ξ^0	$(4\mu_s - \mu_u)/3$	–1.43	-1.250 ± 0.014
Ξ^-	$(4\mu_s - \mu_d)/3$	–0.49	-0.651 ± 0.0025
Ω^-	$3\mu_s$	–1.84	-2.02 ± 0.05

^a In units of proton Bohr magneton ($e/2m_p$). Data are taken from [311].

^b Transition magnetic moment $\Sigma^0 \rightarrow \Lambda$.

the contributions of the quark kinetic energy and the potential made by the color force field add up to make the constituent mass. The constituent quarks are almost at rest and nonrelativistic treatment reproduces the static properties of the hadron quite well.

14.4.2

$SU(6) \times O(3)$

Excitations of orbital angular momentum with $L = 0(S)$, $L = 1(P)$, $L = 3(D)$, ... are possible. Using the spectroscopic terminology $^{2S+1}L_J(J^{PC})$ to denote quantum numbers of the excited states, the known 0^- , 1^- are denoted as $^1S_0(0^{-+})$, $^3S_1(1^{--})$. For $L > 0$ meson states, $^1P_1(1^{+-})$, $^3P_0(0^{++})$, $^3P_1(1^{++})$, $^3P_2(2^{++})$, ... have been observed (Fig. 14.12).

14.5

Charm Quark

14.5.1

J/ψ

In November 1974, an incident remembered as the “November Revolution” occurred. A precursor was the discovery that summer of an anomalously large peak by Ting and group at $m_{ee} \simeq 3 \text{ GeV}$ in a reaction $pp \rightarrow e^-e^+ + X$ carried out at BNL (Brookhaven National Laboratory). Its spectrum was so anomalous, the events showing a large, sharp peak with practically no backgrounds, that the group was

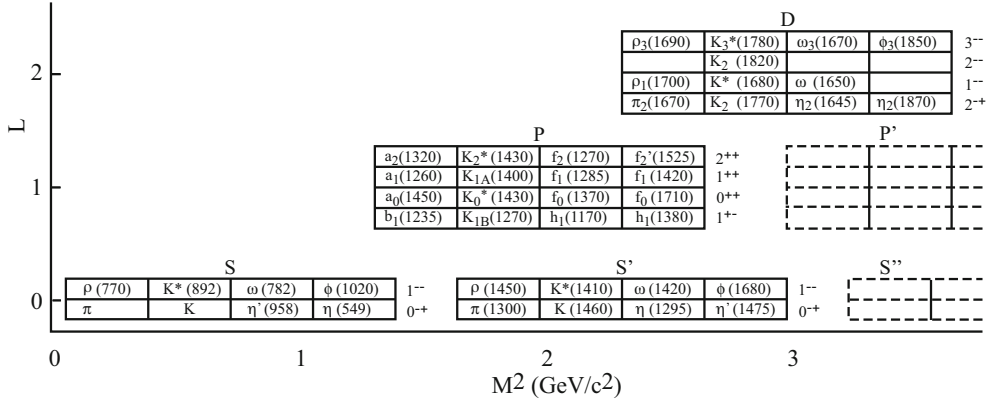


Figure 14.12 A table of observed mesons classified by $SU(6) \times O(3)$ [251, 311]. Particles in the blank spaces have not yet been observed. At the right of each box J^{PC} 's are indicated. S' , S'' , P' are radial excitations of S and P .

not certain about the validity of their observation and repeated tests without ever making it public. On the other hand, at SLAC (Stanford Linear Accelerator Center) on the west coast, Richter's group had been measuring the total cross sections of the e^-e^+ reaction since the summer at the electron-positron collider accelerator SPEAR and were frustrated by anomalous behavior of the observed reaction rate, which went up and down every time they re-measured with no sign of stabilization. Eventually, they noticed that the lack of reproducibility might be due to the inaccuracy of the energy setting. They repeated the experiment with improved accuracy and they also observed a very narrow peak at 3 GeV. Ting named it the J -particle and Richter named it ψ , and today it is known as J/ψ . The reason why people made a fuss about J/ψ was that despite its large mass value of 3 GeV the width was far smaller than the experimental resolution of a few MeV. In conventional thinking, when the parent's mass is large compared to the offspring's mass (if they are a few pions, the total mass is at most 1/10 that of the parent), the parent is expected to have a short lifetime because of the large Q -value. In other words, the decay width should be at least as large as that of resonances with comparable mass (for instance ρ has mass 770 MeV and width $\Gamma_\rho \sim 150$ MeV). The situation is very similar to the discovery of the strange particles. Hot debates ensued, but it was soon realized that J/ψ is composed of the fourth, "charm" quark; Ting and Richter received the Nobel prize in 1975. The short duration between the discovery and the award of the prize was second only to the discovery of parity violation by Lee and Yang, indicating the large impact that made on the high-energy community.

14.5.2

Mass and Quantum Number of J/ψ

J/ψ was produced in the following reactions [33, 34]:

$$\begin{aligned} \text{BNL: } p + Be &\rightarrow J/\psi + \cdots \\ J/\psi &\rightarrow e^- e^+ \end{aligned} \quad (14.58a)$$

$$\text{SLAC: } e^- + e^+ \rightarrow J/\psi \rightarrow e^- e^+, \mu^- \mu^+, \text{ hadrons} \quad (14.58b)$$

Figures 14.13 and 14.14 show J and ψ production cross section results. Note the scale of the abscissa. In Figure 14.15 the cross section goes down just before rising to the peak of the resonance. This is due to an interference effect between the photon and J/ψ intermediate states in $e\bar{e} \rightarrow \mu\bar{\mu}$ and indicates that J/ψ shares the quantum number $J^{PC} = 1^{--}$ with the photon. The shape of the resonance before and after the peak is not symmetric. This is due to the radiative effect, in which the electron and positron with larger energy than the resonance mass emit photons and lose energy, and the remaining less energetic electron–positron pair forms the resonance.

Isospin and G-Parity of J/ψ If $I = 1$, the decay to $\rho^0 \pi^0$ is forbidden. This is easily seen by looking at the Clebsch–Gordan coefficients of making $I = 1$ from two $I = 1$ objects. $I = 0$ was confirmed by the decay mode to $A\bar{A}$ and

$$\frac{\text{BR}(J/\psi \rightarrow \rho^0 \pi^0)}{\text{BR}(J/\psi \rightarrow \rho^+ \pi^-) + \text{BR}(J/\psi \rightarrow \rho^- \pi^+)} \sim \frac{1}{2} \quad (14.59)$$

Since $C = (-)$, this gives $G = (-1)^I C = (-)$. This is also seen by the predominant decay into an odd number of pions rather than an even number of pions.

14.5.3

Charmonium

Potentials of the Quarks Following the discovery of J/ψ (3096.9 ± 0.1 MeV), a higher state $\psi' = \psi$ (3686.0 ± 0.1 MeV) and its decay products χ_{c0} (3415.1 ± 1.0), χ_{c1} (3510.6 ± 0.5), χ_{c2} (3556.3 ± 0.4), etc. from $\psi' \rightarrow \gamma\chi$ were discovered successively. When one makes a chart in terms of J^{PC} (Fig. 14.16), it resembles bound state levels of the positronium made of an electron and a positron and thus they are named charmonium states. From the similarity of the levels, their structure is expected to be almost identical. Namely, it is inferred that a pair of point particles with spin 1/2 interacting in a Coulomb-like potential has formed a bound state. The level structure is better reproduced (Fig. 14.17) if one uses

$$V(r) = -\frac{4}{3} \frac{\alpha_s}{r} + k r^\gamma \quad (14.60)$$

7) $V = r^n$, $-1 \leq N \leq 2(1/r, \ln r, r, r^2)$ produce similar spectra. However, only $V \sim r$ gives the Regge trajectories ($J \propto m^2$) shown in Fig. 14.18.

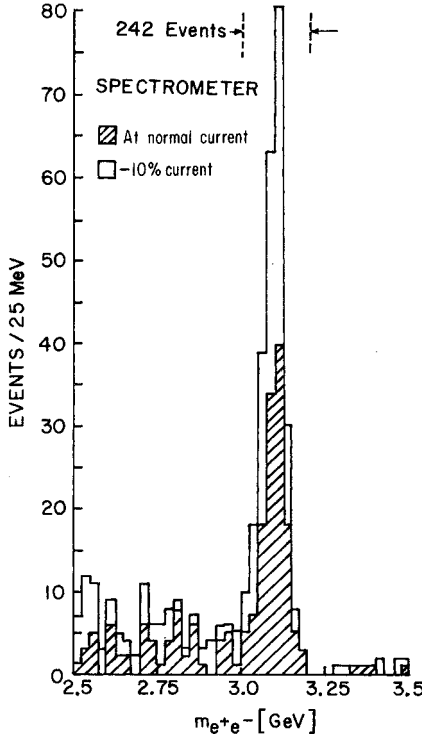


Figure 14.13 J -particle production in p -Be collisions. A very narrow resonance is observed in the e^-e^+ invariant mass distribution. The experiment was carried out with the 28 GeV AGS at Brookhaven National Laboratory [33].

The coefficient $4/3$ of the first term is introduced to be consistent with the QCD potential, which will be described later [see Eq. (14.86)]. The second term, if extrapolated to $r \rightarrow \infty$, gives the confining potential, because to separate the two quarks to $r = \infty$ requires an infinite amount of energy.

Evaluation of α_s For a short distance ($r \lesssim 1/m_\pi \simeq 10^{-15}$ m), the first term is dominant and we can drop the second term for a qualitative argument about the bound state.⁹⁾ Energy levels of a bound particle due to the Coulomb potential (i.e. the levels of hydrogen atoms and positronium) are expressed as

$$E_n = -\frac{\alpha^2 \mu}{2n^2} \quad n: \text{principal quantum number} \quad (14.61)$$

where α is the fine structure constant and $\mu = m_1 m_2 / (m_1 + m_2)$ is the reduced mass. The energy difference between the two levels of $^3S_1$ with $n = 1$ and $n = 2$

9) This is a good approximation for bottomonium to be discussed later. For the case of charmonium, the second term is comparable to the first term.

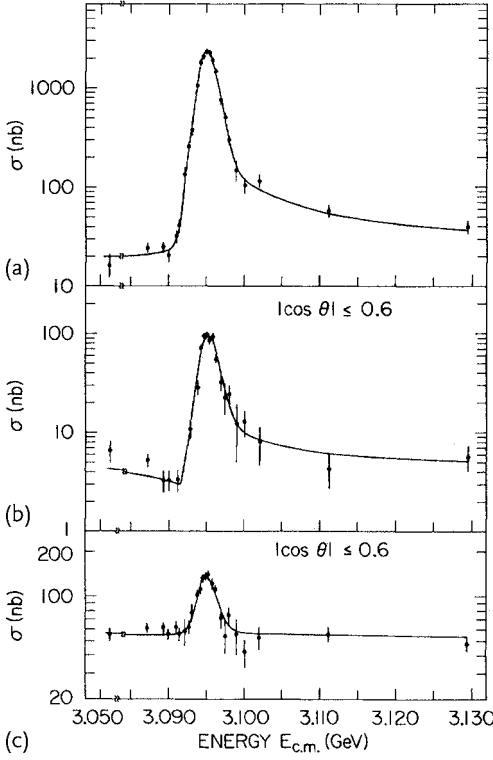


Figure 14.14 Production cross sections of $e^-e^+ \rightarrow \psi \rightarrow X$, where X is (a) hadrons, (b) $\mu^-\mu^+$ and (c) e^-e^+ . The experiment was carried out with PEP at Stanford Linear Accelerator Center [34, 73].

is given by

$$\Delta E \simeq 5 \text{ eV} = \frac{\alpha^2 m_e}{4} \left(\frac{1}{1^2} - \frac{1}{2^2} \right) = \frac{3}{16} \alpha^2 m_e \quad (14.62)$$

where the reduced mass $m_e/2$ is used for the positronium. For charmonium, we change $\Delta E \rightarrow 600 \text{ MeV}$, $\alpha \rightarrow (4/3)\alpha_s$, $m_e \rightarrow m_c$. We set $m_c \simeq 3700/2 \text{ MeV}$, because while the decay width of $\psi'(3686)$ is narrow, that of $\psi'' = \psi(3770) \rightarrow D\bar{D}$ ($D\bar{D}$ are hadrons which contain a charm quark) is wide ($\sim 25 \text{ MeV}$), indicating a strong decay. In other words, the mass of ψ exceeds $2m_c$ between the two levels. Inserting numerical values ($m_e \simeq 0.5 \text{ MeV}$, $\alpha \simeq 1/137$), we obtain $\alpha_s \simeq 1$. It is very surprising that levels at energy scales differing by a factor of 100 million exhibit the same structure. Thus, it was clarified that a Coulomb-like force, but 100 times stronger, is at work between the quarks.

Confinement and the Regge Trajectory The Regge trajectory represents the phenomenon that a series of particles with the same quantum numbers other than spin are on straight tracks with the same slope on the m^2 - J plane (Figs. 14.18

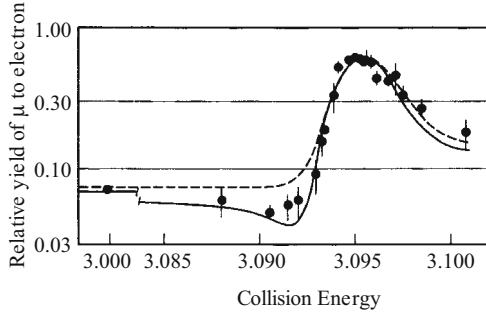


Figure 14.15 Interference effect between γ and J/ψ [73]. If J/ψ has the same quantum number as γ , the interference effect appears in the $e^-e^+ \rightarrow \mu^-\mu^+$ reaction as a change in the cross section just before the resonance

sets in (solid line). If the quantum number is different, no interference occurs (dashed line). The data show the existence of the interference.

and 13.29). We shall show that $V(r) = kr$ can reproduce the Regge trajectories. Assume the attractive potential $V(r) = kr^n$, that two massless quarks are rotating on a circle and that the centrifugal force is in balance with the potential. The mass of the two-particle system, i.e. the total energy, is given by

$$H = 2p + kr^n \quad (14.63)$$

As they are rotating, they have angular momentum $J = 2pr$. They are stationary at the distance where H is minimum,

$$\frac{\partial H}{\partial r} = 0 \rightarrow J = nkr^{n+1} \quad (14.64)$$

Inserting Eq. (14.64) in Eq. (14.63) we have

$$H \propto J^{n/(n+1)} \rightarrow J \propto M^{(n+1)/n} \quad (14.65)$$

To reproduce the Regge trajectory, n has to be 1. Inserting $n = 1$ and recalculating, we obtain

$$J = \frac{m^2}{4k} \sim \frac{(mc^2)^2}{4k\hbar c} \quad (14.66)$$

where we have recovered $\hbar$ and c in the last equation. From Figure 14.18 the slope of the Regge trajectories is obtained to be $\sim 1 \text{ (GeV)}^{-2}$ and

$$k = 0.25 \text{ GeV}^2/\hbar c \simeq 1.3 \text{ GeV}/10^{-15} \text{ m} \quad (14.67)$$

which is often referred to as the string tension. The work necessary to separate the quarks to a distance of 1 m is $\sim 10^{15} \text{ GeV} \sim 150 \text{ tons} \times 1 \text{ m}$. It is impossible to give that amount of energy to a quark.

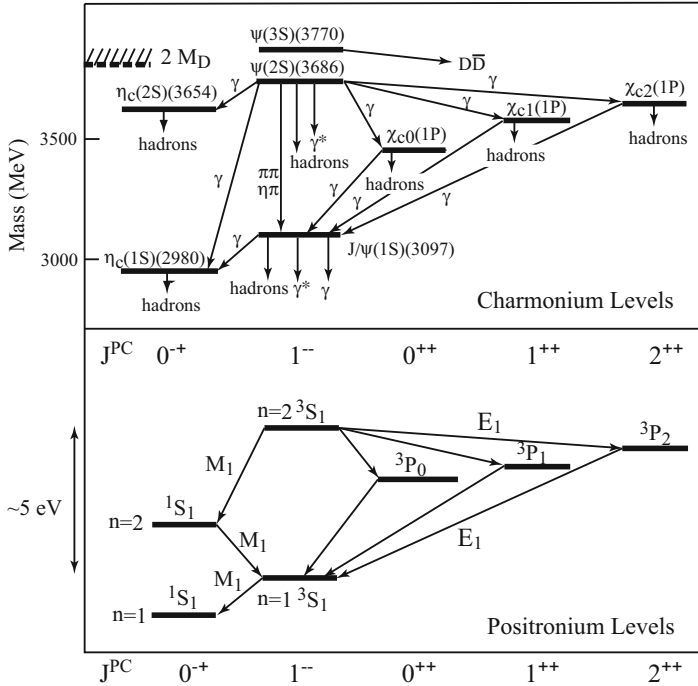


Figure 14.16 Comparison of levels between charmonium and positronium [311].

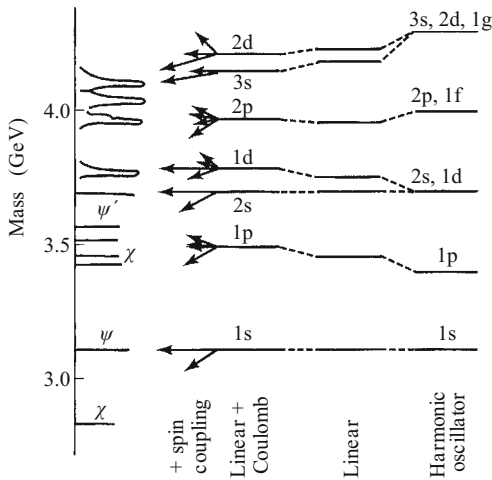


Figure 14.17 A comparison of level structure by using various kind of potentials [101]. The rightmost column shows the levels of a harmonic oscillator ($V \sim r^2$), the second from right is due to a linear potential and the third

is a combination of the linear and Coulomb potentials [Eq. (14.60)]. Introducing spin couplings reproduces the observed spectrum, shown in the fourth column (at the left).

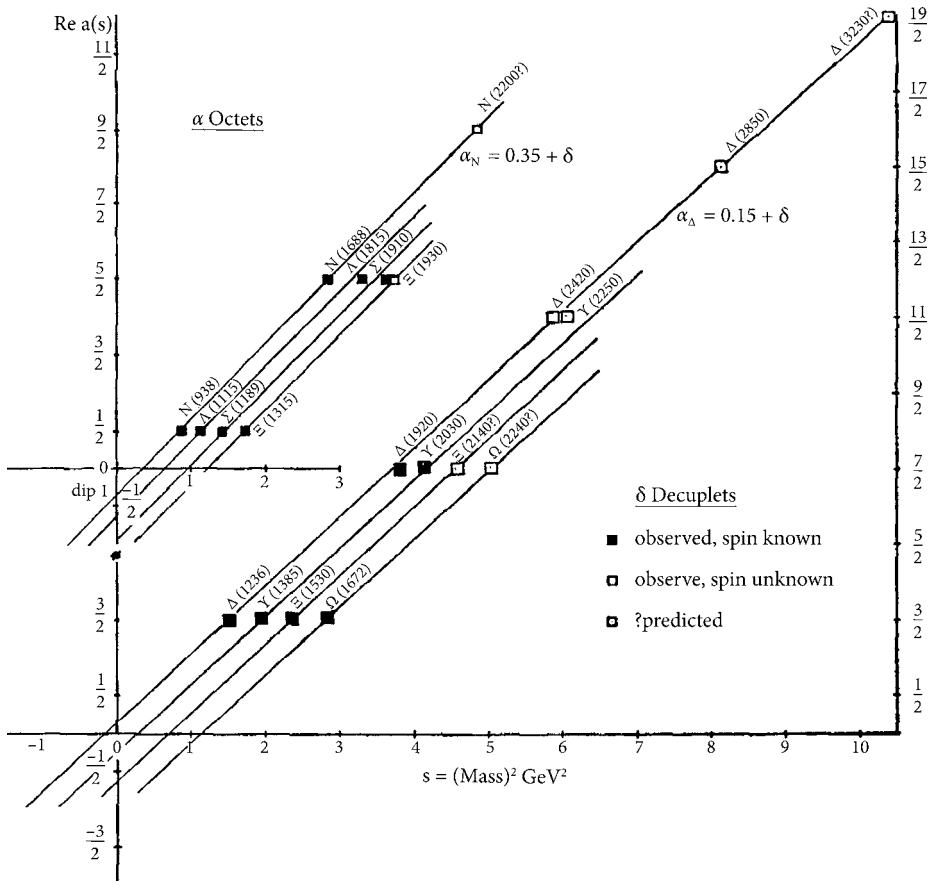


Figure 14.18 Regge trajectories: A group of particles with the same quantum numbers except their spin are on straight tracks ($J \propto m^2$) with the same gradient on the m^2 - J plane. This is called the Chew–Frautschi plot [98].

$V(r) = kr$ means the quarks are attracted to each other by a string with constant tension. Hence k is referred to as the string constant. As $1.3 \text{ GeV} \sim 9m_\pi$, an energy equivalent of 9π 's is stored if the two quarks are $\sim 10^{-15} \text{ m}$ apart. One notices that this is about the size and mass of a hadron. If the string between the quarks is stretched, a large amount of energy is stored in it. According to the Heisenberg uncertainty principle, pairs of particles and antiparticles are constantly produced and annihilated in the vacuum. If a pair is produced on the line, they make two color singlets with the original quarks. Then the force between the pair is weaker than that between the members of the singlet and the string is cut off. This means that making hadrons by extracting quark pairs from the vacuum is more advantageous than stretching the string further (Fig. 14.19). Thus, the string stretched between two energetic quarks is cut into pieces, producing many hadrons, which is referred to as quark hadronization. This is why the quark is believed to be unable to separate

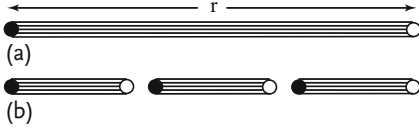


Figure 14.19 Energy stored between two quarks increases in proportion to the separation and behaves like a string with constant tension. If a pair of quarks is produced in vacuum and combines with the original pair to make a new pair of color singlets, the orig-

inal color force is terminated there and the string is cut. In other words, for $r > 10^{-15}$ m it is more advantageous to cut the string and hadronize than to stretch the string further. Many hadrons are created.

as a single entity (confinement). An energetic quark (and gluon also) is converted into many hadrons flying as a bundle in the same direction, which is observed as a jet phenomenon. Figure 14.20a shows a typical two-jet event ($e^- + e^+ \rightarrow q + \bar{q}$) observed by the VENUS detector of TRISTAN e^-e^+ collider [285, 380]. One immediately observes that many tracks and associated energies are concentrated in two opposite directions. Figure 14.20b shows an azimuthal angle distribution of the jet axis produced by polarized electrons at SPEAR [195]. The theoretical angular distribution of a particle in the e^-e^+ reaction is expressed as

$$\frac{d\sigma}{d\Omega} \propto 1 + \alpha \cos^2 \theta + P^2 \alpha \sin^2 \theta \cos(2\varphi) \quad (14.68)$$

where P is beam polarization and $\alpha = +1$ for spin $1/2$ particles and $\alpha = -1$ for spin 0 particles. The data show $\alpha = 0.78 \pm 0.12$, confirming the quark's spin to be $1/2$.

14.5.4

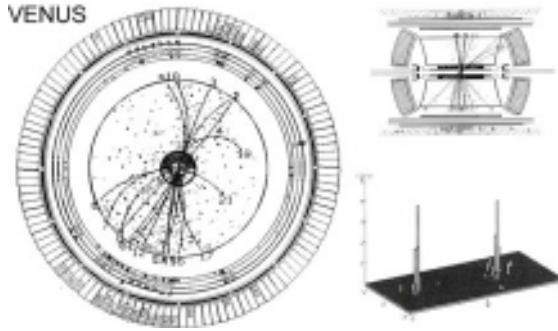
Width of J/ψ

The resonance width of the J/ψ was narrower than the experimental resolution. We discuss a method to determine the width. When there is a resonance, the total cross section to a final state f is given by [see Eq. (13.85)]

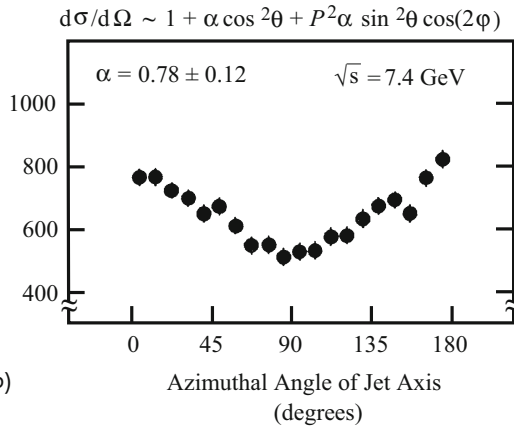
$$e^-e^+ \rightarrow J/\psi \rightarrow f$$

$$\sigma = \frac{2J+1}{(2s_1+1)(2s_2+1)} \frac{4\pi}{p^2} \left(\frac{\Gamma_{ee}}{\Gamma} \right) \left(\frac{\Gamma_f}{\Gamma} \right) \frac{\Gamma^2/4}{(E-E_R)^2 + \Gamma^2/4} \quad (14.69)$$

Here, Γ_e/Γ and Γ_f/Γ are the branching ratios of the J/ψ decaying to electrons or the final f state, $s_1 = s_2 = 1/2$ are electron spins and $J = 1$. The observed value is a cross section convoluted with the experimental resolution δ_{exp} . Under normal circumstances, $\Gamma \gg \delta_{\text{exp}}$ and the functional shape of Eq. (14.69) can be reproduced faithfully. If, on the other hand, Γ is much smaller than δ_{exp} (here, $\delta_{\text{exp}}/E \simeq 0.1\%$, $\delta_{\text{exp}} \simeq 3 \text{ MeV} \gg 93 \text{ keV}$), the measured cross section is nothing but the error function (see Fig. 14.21).



(a)



(b)

Figure 14.20 Quark jets produced in $e^- + e^+ \rightarrow q + \bar{q}$ reaction. (a) Two views of tracks in the central drift chamber. Numbers in rectangles at outer circle are energy deposits in the lead glass calorimeter. Lower right: Back-to-back emission of two sharply

peaked jets is evident in the Lego plot, a 3-dimensional display of energy deposits in $\phi - z$ plane [285]. (b) Azimuthal angular distribution of the jet axis produced by a polarized beam. $\alpha = 1$ for spin $1/2$ and $\alpha = -1$ for spin $= 0$ particles [195].

When $f = ee$, the integral of Eq. (14.69) over energy with $p = p_R = 1548 \text{ MeV}/c$ fixed at the resonance peak gives

$$\int_0^\infty \sigma(E) dE = \frac{3\pi^2}{2p_R^2} \left(\frac{\Gamma_{ee}}{\Gamma} \right)^2 \Gamma \quad (14.70)$$

which should be equal to the area under the resonance curve of $\sigma(ee \rightarrow e^- e^+)$ (Fig. 14.14c). As the branching ratio $\Gamma_{ee}/\Gamma = 0.0594$ can be measured independently, comparison with the data gives $\Gamma = 93.2 \pm 2.1 \text{ keV}$, $\Gamma_{ee} = 5.52 \pm 0.14 \text{ keV}$. The smallness of $\Gamma = 93 \text{ keV}$ may be recognized by comparing it with hadronic

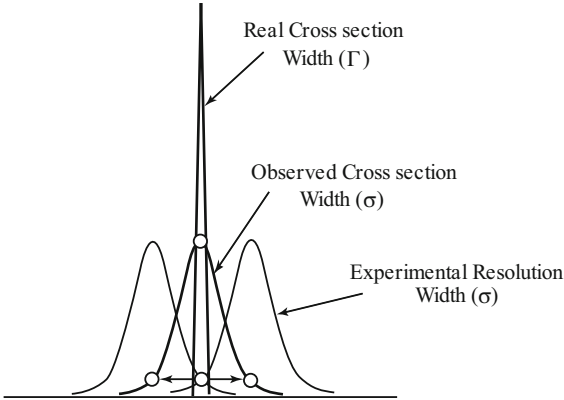


Figure 14.21 When the experimental resolution is much worse than the resonance width, $\delta_{\text{exp}} \gg \Gamma$, the observed shape of the resonance is that of the error function.

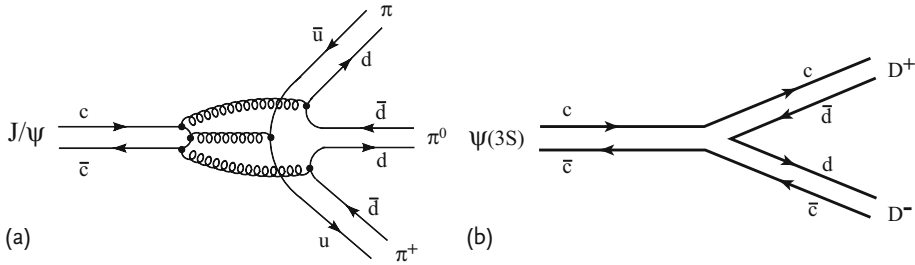


Figure 14.22 (a) $J/\psi \rightarrow 3\pi$ has disconnected charm lines. (b) In $\psi(3770) \rightarrow D^+\bar{D}^-$ the charm lines are connected. By the OZI rule, (a) is heavily suppressed. Compare with Fig. 14.7.

decays of comparable resonances:

$$\begin{array}{lll}
 \rho : & M = 768 \text{ MeV}, & \Gamma_\rho = 152 \text{ MeV}, & \Gamma_{\rho \rightarrow ee} = 6.8 \text{ keV} \\
 \omega : & M = 782 \text{ MeV}, & \Gamma_\omega = 8.4 \text{ MeV}, & \Gamma_{\omega \rightarrow ee} = 0.6 \text{ keV} \\
 \phi : & M = 1020 \text{ MeV}, & \Gamma_\phi = 4.4 \text{ MeV}, & \Gamma_{\phi \rightarrow ee} = 1.37 \text{ keV}
 \end{array} \quad (14.71)$$

We see the width of J/ψ is 1000 times smaller than typical hadron widths and almost comparable with the electromagnetic decay widths.

This was why people were surprised by the discovery of J/ψ . As the mass of J/ψ is far larger than p, n, Λ, Σ , etc., it should decay strongly to pions and other light particles, giving $\Gamma \sim O(100) \text{ MeV}$, unless some selection is at work. Indeed, $\psi(3S)(3772.92 \pm 0.35)$ has a broad width $\sim 27.3 \pm 1.0 \text{ MeV}$ and predominantly decays to $D^+ D^-$ or $D^0 \bar{D}^0$ (Fig. 14.22b). The D particles were discovered in $e^- + e^+ \rightarrow D^+ D^-$, $D^0 \bar{D}^0$ [114, 177, 178, 321] and they are quasistable particles. Their

masses and lifetimes are

$$M(D^\pm) = 1869.6 \pm 0.2 \text{ MeV}, \quad \tau(D^\pm) = 10.4 \pm 0.07 \times 10^{-13} \text{ s} \quad (14.72a)$$

$$M(D^0) = 1864.8 \pm 0.17 \text{ MeV}, \quad \tau(D^0) = 4.10 \pm 0.015 \times 10^{-13} \text{ s} \quad (14.72b)$$

These phenomena cannot be explained if these particles contain only the known quarks u, d, s . They can be understood if $J/\psi, \psi(3770)$ are assumed to be made of new quarks $c\bar{c}$. The fact that D particles are heavy ($m \sim 2 \text{ GeV}$) and yet quasistable can be understood if they contain c or $\bar{c}$. We now know that $D^+ = c\bar{d}$, $D^- = \bar{c}d$, $D^0 = c\bar{u}$, $\bar{D}^0 = \bar{c}u$. The c quark has a new quantum number “charm” that is conserved in the strong interaction, hence it cannot decay strongly to ordinary particles containing only u, d, s , but it can decay weakly. J/ψ has total charm quantum number zero and the decay mode conserves the charm quantum number, but drawing quark lines of the decay, the charm line has to be closed (see Fig. 14.22a). This is because $m(J/\psi) < 2m_D$ and no process satisfying the OZI rule (see Sect. 14.2.3) exists. On the other hand the charm quark line in the decay $\psi(3770) \rightarrow D\bar{D}$ is connected and hence the OZI rule is satisfied.

In the language of QCD, where the quark lines are disconnected, they are connected by gluon lines. Each time a gluon is exchanged, a factor $\alpha_s/2\pi \sim 0.1$, where $\alpha_s = g^2/4\pi$, has to be multiplied. For J/ψ three gluons are exchanged (Problem 14.13). Therefore the decay rate is suppressed by $(\alpha_s/2\pi)^6$ compared to ordinary strong decays.

Problem 14.13

Explain why there have to be at least three gluons in the intermediate state in Fig. 14.22.

14.5.5

Lifetime of Charmed Particles

The fourth quark was proposed as early as 1964 based on baryon–lepton symmetry [196, 270],¹⁰ but it was not until 1970 that it was taken seriously, when Glashow, Iliopoulos and Maiani showed that the absence of strangeness-changing neutral weak current can be explained if the existence of the charm quark is assumed (GIM mechanism [174], see Sect. 15.7.1 for details). It constitutes a doublet in the weak interaction paired with the s quark. If D^0 is made of the charm quark in a weak doublet, it can decay predominantly to the s quark by emitting a W^+ weak boson,

10) The charm quark itself was discovered by Niu [297] prior to the November revolution. His group found a short-lived particle with mass and lifetime $m \sim 2\text{--}3 \text{ GeV}$, $\tau \sim 10^{-14} \text{ s}$, decaying to π^0 and a charged particle in an emulsion stack placed in an

airplane. The Nagoya group analyzed it regarding it as the fourth quark. However, as it was a cosmic ray event with poor statistics, and the particle identification was not definitive, the discovery was largely ignored outside Japan.

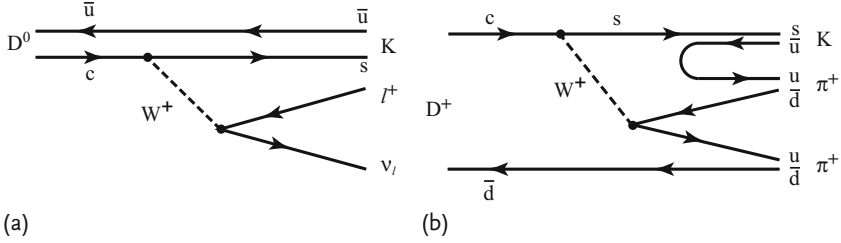


Figure 14.23 (a) $D^0 \rightarrow K^- + l^+ + \nu_l$, ($l = e, \mu$), where $\bar{u}$ acts only as a spectator. Here the decay can be considered as $c \rightarrow s + l^+ + \nu_l$. (b) An example of hadronic decay $D^+ \rightarrow K^- \pi^+ \pi^+$.

producing a variety of decay products (Fig. 14.23). Let us consider a semi-leptonic decay $D^0 \rightarrow K^- + l^+ + \nu_l$ ($l = e, \mu$). Here, the $\bar{u}$ quark does not participate in the decay, playing only a spectator role, and the process can be considered as $c \rightarrow s + l^+ + \nu_l$, which has identical kinematical structure to $\mu \rightarrow e + \nu_e + \nu_\mu$, differing only in the mass values. Since the lifetime of the muon is $\propto (G_F^2 m_\mu^5)^{-1}$ [see Eq. (15.84)], the lifetime of D is estimated to be roughly

$$\begin{aligned} \tau_D &\sim \text{BR}(D \rightarrow K^- e^+ \nu_e) \left(\frac{m_\mu}{m_c} \right)^5 \tau_\mu \\ &\sim 0.036 \times (0.1/1.5)^5 \times 2.2 \times 10^{-6} \text{ s} \sim 10^{-13} \text{ s} \end{aligned} \quad (14.73)$$

which reproduces the experimental value ($\tau_{D\text{exp}} = (4.10 \pm 0.02) \times 10^{-13} \text{ s}$) approximately. This shows that the D meson decays weakly.

14.5.6

Charm Spectroscopy: $SU(4)$

If there are four quarks, the flavor symmetry $SU(3)$ is extended to $SU(4)$. However, the mass of the charm quark $\sim 1.3 \text{ GeV}^{11)}$ is far greater than others, and $SU(4)$ is not considered as a good symmetry. Nevertheless, it is useful in classifying and predicting the existence of hadrons containing the c quark. We briefly describe multiplets that are generated by $SU(4)$ (see Appendix G for $SU(4)$ algebra). The meson nonets in $SU(3)$ become 16-plets in $SU(4)$.

Charmed Mesons

$$4 \otimes 4^* = 1 \oplus 15 \quad (14.74)$$

Among the 15 members of **15**, we already know eight members, which do not contain c . The remaining seven members contain a c quark and are called charmed

11) This is the current mass, which is somewhat smaller than the constituent mass, but not as drastically small as u, d, s .

mesons. They are

$$J^P = 0^-, 1^- \begin{cases} C = 1 : D^0(c\bar{u}), D^+(c\bar{d}), D_s(c\bar{s}) \\ C = 0 : \eta_c \sim c\bar{c} \\ C = -1 : \bar{D}^0(\bar{c}u), D^-(\bar{c}d), \bar{D}_s(\bar{c}s) \end{cases} \quad (14.75)$$

Their configuration in I_3 - Y - C space can be expressed as in Fig. 14.24, (a) for 0^- and (b) for 1^- . All of them have already been discovered.

Charmed Baryons

$$4 \otimes 4 \otimes 4 = 4_A \oplus 20_{MA} \oplus 20_{MS} \oplus 20_S \quad (14.76)$$

The two 20_S with mixed symmetry correspond to $(1/2)^+$ and the totally symmetric 20_S corresponds to $J^P = (3/2)^+$. Consequently, there are 12 charmed baryons with $(1/2)^+$ and 10 charmed baryons with $J^P(3/2)^+$ (Fig. 14.24c,d). Λ_c^+ , Σ_c , Ξ_c and Ω_c have been observed.

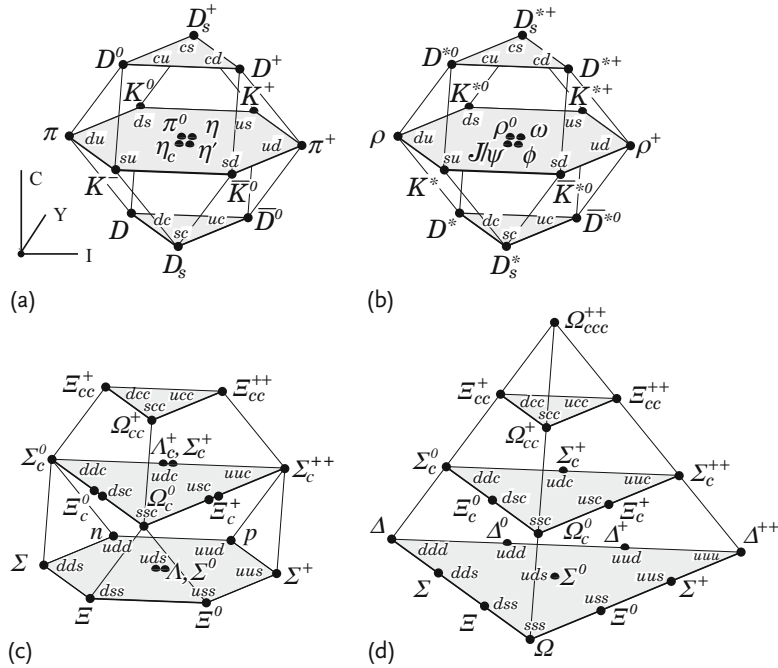


Figure 14.24 Charm spectrum: meson 16 and baryon 20. (a) 0^- (b) 1^- . The central planes of (a) and (b) represent $C = 0$ mesons and except for η_c and J/ψ are $SU(3)$ multiplets.

(c) $(1/2)^+$, (d) $(3/2)^+$. The plane at the bottom represents $C = 0$ baryons, which already appeared in $SU(3)$.

14.5.7

The Fifth Quark b (Bottom)

A repeat of almost the same procedure was used to discover upilon $\Upsilon(b\bar{b})$ bound state) in 1977 [213] shortly after the charm discovery

$$\begin{aligned} p + p &\rightarrow \Upsilon + X \\ \Upsilon &\rightarrow \mu^- + \mu^+ \end{aligned} \quad (14.77)$$

and later, a series of upsilons were discovered [323] at DESY (Deutsches Elektronen-Synchrotron) and at CESR (Cornell Electron Storage Ring) [24, 71]:

$$e^- + e^+ \rightarrow \Upsilon(9460), \quad \Upsilon'(10023), \quad \Upsilon''(10355), \quad \Upsilon'''(10580) \quad (14.78)$$

Among them the first three had narrow widths and Υ''' had a broad width, indicating $m_b \sim 5$ GeV. Figure 14.25 shows the first three upsilons.

The level structure of the Υ 's is quite similar to that of the charmonium and collectively they are called bottomonium. However, corresponding to the larger mass $m_b \simeq 5$ GeV, bound states are formed deep in the potential, and up to $\Upsilon(3S)$ are quasistable (Fig. 14.26). For the observed hadrons that include the b quark, refer to the Particle Data Group table [311].

14.6

Color Charge

The discovery of the color degrees of freedom was the key to the true understanding of the strong force. It was also the discovery of the “color charge”. In electromag-

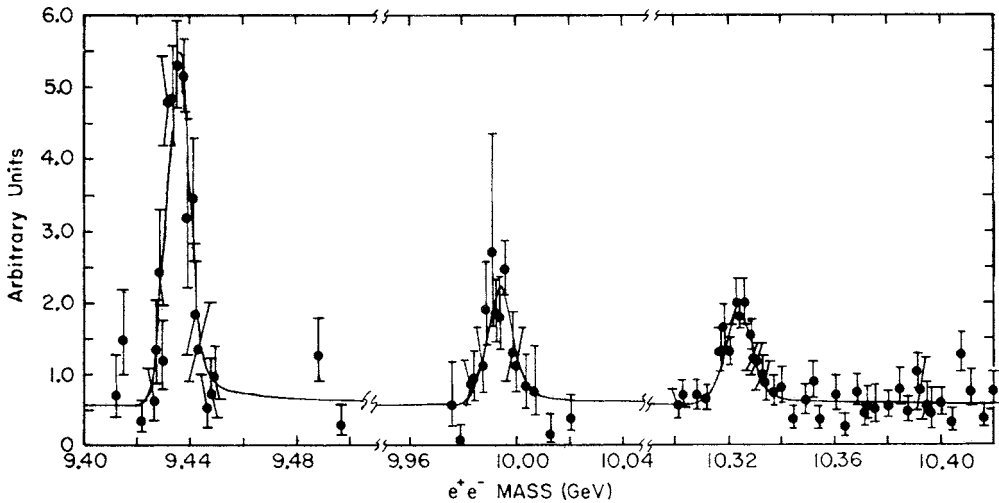


Figure 14.25 $\Upsilon, \Upsilon', \Upsilon''$ observed in $e^-e^+ \rightarrow \Upsilon$'s [24, 71].

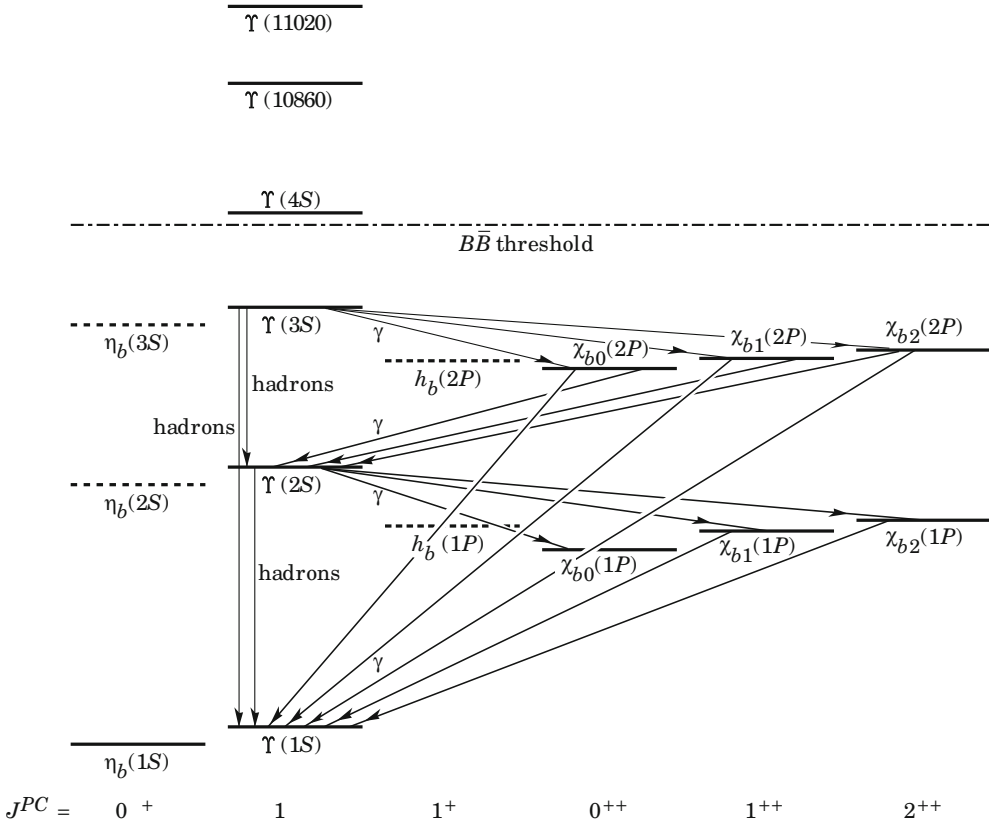


Figure 14.26 Level structure of bottomonium [311].

netism, the electric force is generated by an electric charge and the magnetic force by its current. Since current is nothing but a moving charge, the electric and magnetic forces are unified by the relativity principle. The current is a Lorentz four vector and its 0th or scalar component generates the electric force and is spin independent in its rest frame. The space component, a vector, generates the magnetic force, which acts on (couples to) the current (the moving charge) or more generally angular momentum, including spin. This characteristic is closely related to the fact that the photon is a Lorentz vector. Just like the electromagnetic force, the strong force (or more rigorously color force to discriminate it from the nuclear force) is generated by the color charge and its current. This is why color force acts like the Coulomb force at short distances. Since the characteristics of the two forces are very much alike, the color force is expected to obey the same mathematical framework as that of the electromagnetic force, namely gauge theory. The difference lies in the number of charges. There are three color charges (R, G, B), which generate the strong force with equal strength. The action (or emission/absorption of the force carrier, which is referred to as the gluon) changes the color. So we can infer

that the characteristics of the color force are basically similar to those of the electromagnetic force with added effects due to color exchange. We have already learned that the nuclear exchange force stems from the exchange of pions, which carry isospin. Unlike isospin, whose symmetry is $SU(2)$, color comes in three kinds and its symmetry is $SU(3)$. This is the origin of color $SU(3)$ and, unlike flavor $SU(3)$, it is considered to be an exact symmetry.¹²⁾

From the above arguments, it is natural that a gauge theory based on $SU(3)$ has emerged as a strong candidate for governing the strong interaction. Because of its similarity to QED (quantum electrodynamics), the mathematical framework of the color force is called QCD (quantum chromodynamics). A strict formalism for constructing the QCD Lagrangian will be given in Chap. 18, but here we show that a lot of insight into the QCD features can be obtained by simple color counting if we restrict our arguments to the lowest order corrections in the gluon emission or exchange, namely as long as we do not touch the nonlinear aspect of QCD. Much of the elementary dynamics can be obtained using QED formula we already know with additional color factors. To obtain a glimpse of the power of QCD, we consider a color exchange potential and how low-lying mass levels of the hadron are reproduced.

The force carrier, the gluon, comes in eight kinds. Let us try to understand why. Analogy with the nuclear force helps because this is also an exchange force obeying $SU(2)$ symmetry that alters the identity of the force source – p (proton) or n (neutron) – by the force action. The nuclear force carrier pion comes in three kinds. If we omit the space-time structure, the nuclear force interaction is expressed as

$$H = g_N \sum_{i=1}^3 \bar{\psi} \frac{\tau_i}{2} \psi \pi_i = \frac{g_N}{2} \left[\sqrt{2}(\bar{p}n\pi^+ + \bar{n}p\pi^-) + (\bar{p}p - \bar{n}n)\pi^0 \right] \quad (14.79a)$$

$$\psi = \begin{bmatrix} p \\ n \end{bmatrix}, \quad (\pi^+, \pi^0, \pi^-) = \left(\frac{\pi_1 - i\pi_2}{\sqrt{2}}, \pi_3, \frac{\pi_1 + i\pi_2}{\sqrt{2}} \right) \quad (14.79b)$$

where τ_i is the Pauli matrix. The pion field operators (π^+, π^0, π^-) have the same isospin function as $\{\bar{p}n, (\bar{p}p - \bar{n}n)/\sqrt{2}, \bar{n}p\}$ and change the isospin component of the force source ψ . For instance, absorption of π^+ changes n to p (i.e. $\pi^+ \sim \bar{p}n$). Actually there are four combinations to change isospin components (p, n) to (p, n), but the combination $(\bar{p}p + \bar{n}n)/\sqrt{2}$ is totally neutral (isospin singlet) and does not count. Therefore, the number of force carriers is $2^2 - 1 = 3$. In mathematical language, the pion belongs to the regular representation of the $SU(2)$ group. In correspondence to the isospin exchange hamiltonian, one that generates the color exchange interaction can be expressed as

$$H = g_s \sum_{i=1}^8 \bar{\psi} \frac{\lambda_i}{2} \psi A_i, \quad \psi = \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (14.80)$$

12) At long range, the confinement force becomes effective, which is a nonlinear effect of the color charge.

λ_i is the Gell-Mann matrix [Eq. (14.8)] and is the regular representation of $SU(3)$. A_i is the force field, referred to as the gluon. There are $3^2 - 1 = 8$ gluons (A_i). The absorption or emission of the gluon changes the color charge or the color state of the quark ($R \leftrightarrow G \leftrightarrow B \leftrightarrow R$), just as the pion changes $p \leftrightarrow n$. The gluon is an eight-component vector in color space, just as the pion is an isovector in isospin space.

Unlike the pion, the gluon is a Lorentz vector, like the photon, and its spatial part induces the magnetic interaction. This is the origin of color and spin exchange forces, which will be described shortly.

14.6.1

Color Independence

We show the color independence of the quark–gluon coupling strength just like the charge independence of the pion–nucleon coupling (recall arguments in Sect. 13.3.1).

We first reformulate the gluon fields to grasp an intuitive image of the gluon's role. Just as we used (π^+, π^0, π^-) instead of abstract (π_1, π_2, π_3) , we define gluons, g_i 's ($i = 1, \dots, 8$), as follows. Omitting the space-time part of the field subscripts for simplicity, we write down the Gell-Mann matrices explicitly:

$$\begin{aligned} -\mathcal{L}_{\text{QCD}} &= g_s \sum \bar{\psi} (\lambda_a/2) \psi A_a \\ &= \frac{g_s}{\sqrt{2}} \left[\bar{R} G g_1 + \bar{G} R g_2 + \frac{1}{\sqrt{2}} (\bar{R} R - \bar{G} G) g_3 + \bar{R} B g_4 \right. \\ &\quad \left. + \bar{B} R g_5 + \bar{G} B g_6 + \bar{B} G g_7 + \frac{1}{\sqrt{6}} (\bar{R} R + \bar{G} G - 2\bar{B} B) g_8 \right] \end{aligned} \quad (14.81a)$$

where

$$\begin{aligned} g_1 &= \frac{1}{\sqrt{2}} (A_1 - iA_2), & g_2 &= \frac{1}{\sqrt{2}} (A_1 + iA_2), & g_3 &= A_3 \\ g_4 &= \frac{1}{\sqrt{2}} (A_4 - iA_5), & g_5 &= \frac{1}{\sqrt{2}} (A_4 + iA_5), \\ g_6 &= \frac{1}{\sqrt{2}} (A_6 - iA_7), & g_7 &= \frac{1}{\sqrt{2}} (A_6 + iA_7), \\ g_8 &= A_8 \end{aligned} \quad (14.81b)$$

The g_a 's are redefined gluon fields. Equation (14.81a) means, for instance, g_1 has the ability to change its color $G \rightarrow R$ if absorbed by G , in other words, as far as the color degrees of freedom are concerned, it plays an equivalent role to $\bar{R}G$. This is analogous to π^+ being $\sim \bar{n}p$.

Let us calculate the strength of the force that acts between two quarks using the Lagrangian in Eq. (14.81a). First of all, between RR we can draw two Feynman diagrams, shown in Fig. 14.27a,b. Here, at the vertices R is annihilated and recreated

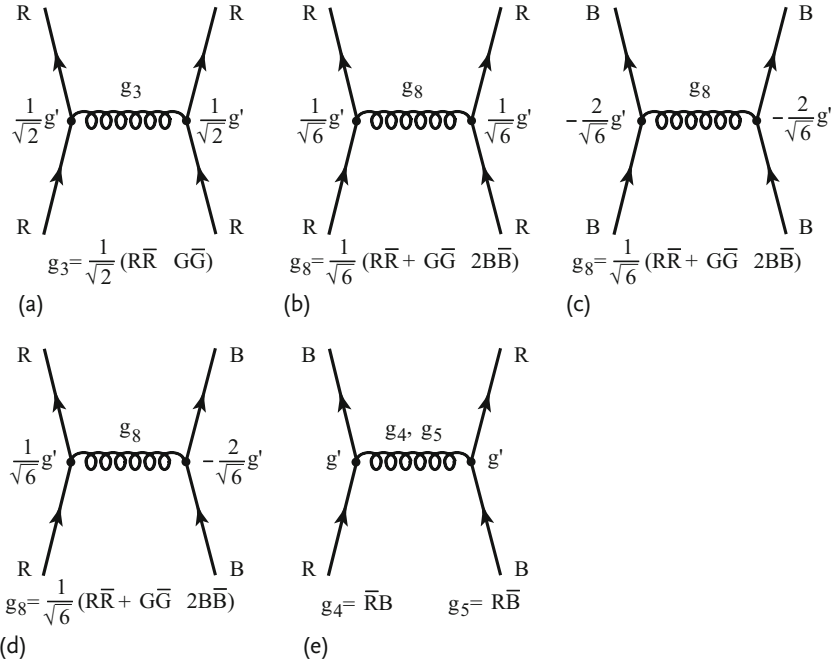


Figure 14.27 Color independence, i.e. universality of the quark–gluon coupling strength.

(i.e. $R \rightarrow R \sim \bar{R}R$). Among the gluons in Eq. (14.81b), one sees that g_3 and g_8 couple to $\bar{R}R$. The strength of the g_3 , g_8 exchange is given by

$$g'^2 \left[\left(\frac{1}{\sqrt{2}} \right)^2 + \left(\frac{1}{\sqrt{6}} \right)^2 \right] = \frac{2}{3} g'^2 \quad (14.82)$$

where $g' = g_s/\sqrt{2}$. Next we calculate the strength of the BB force (Fig. 14.27c). Looking at Eq. (14.81a), we see only g_8 couples to $\bar{B}B$, consequently

$$g'^2 \times \left(-\frac{2}{\sqrt{6}} \right)^2 = \frac{2}{3} g'^2 \quad (14.83)$$

This has the same strength as the RR force. Considering the universality of the forces acting between R, B, G , it is logical to obtain the same strength. Next we consider the RB force. First, in addition to the contribution (Fig. 14.27d) due to g_8 , which couples to both R and B , we have to consider the RB exchange force due to g_4, g_5 (Fig. 14.27e). Combining the two

$$g'^2 \left[\left(\frac{1}{\sqrt{6}} \right) \times \left(-\frac{2}{\sqrt{6}} \right) + (1) \times (1) \right] = \frac{2}{3} g'^2 \quad (14.84)$$

one sees that the universality is maintained. Next, let us consider the force in a singlet meson $(\bar{R}R + \bar{G}G + \bar{B}B)/\sqrt{3}$.

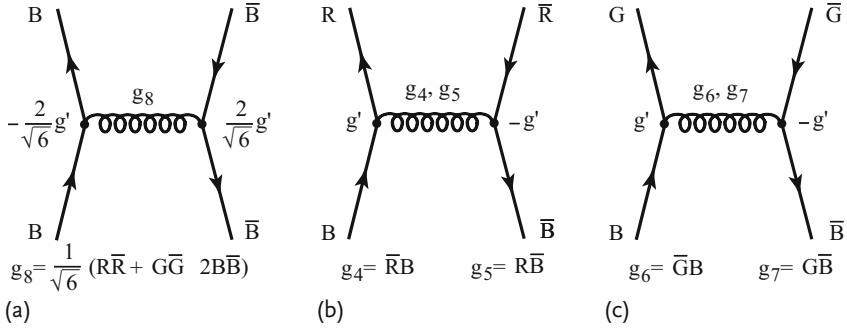


Figure 14.28 Gluon exchanges between $B\bar{B}$ in a color singlet meson.

Since each force that acts on the three pairs is equal, we only consider an initial state $B\bar{B}$ and multiply by 3. But as the weight of $\bar{R}R$ in the singlet meson is $(1/\sqrt{3})^2 = 1/3$, all we need to do is to calculate the $\bar{B}B$ contribution. Considering $B\bar{B}$ can be converted to $B\bar{B}$, $R\bar{R}$, $G\bar{G}$, there are three diagrams to consider (Fig. 14.28). As the gluon is a vector and changes its sign for the antiparticle, we have three contributions from each diagram:

$$g'^2 \left[\left(\frac{2}{\sqrt{6}} \right) \left(-\frac{2}{\sqrt{6}} \right) + (1)(-1) + (1)(-1) \right] = -\frac{8}{3}g'^2 = -\frac{4}{3}g_s^2 \quad (14.85)$$

The formula tells us that there is an attractive force between the quarks forming a singlet meson. This is the strength of the color force that acts between q and $\bar{q}$, corresponding to that of the Coulomb force in QED. Writing $\alpha_s = g^2/4\pi$, we have

$$\frac{4}{3}\alpha_s \quad (\text{color force in QCD}) \leftrightarrow \alpha \quad (\text{Coulomb force in QED}) \quad (14.86)$$

Now we have reproduced the coefficient of the Coulomb-like force that was introduced in Eq. (14.60) to explain the charmonium spectrum.

14.6.2

Color Exchange Force

The calculations we have done are intuitive and easy to understand but rather complicated. In general, we rely on group theory. The force acting on qq or $q\bar{q}$ can be expressed as a potential. Expressing it in scattering matrix form, we have

$$\begin{aligned} \mathcal{M}(c + c' \rightarrow b + b') &\sim g_s^2 \sum_A \langle bb' | [\bar{\psi}_b(\lambda_A/2)_{bc} \psi_c] [\bar{\psi}_{b'}(\lambda_A/2)_{b'c'} \psi_{c'}] | cc' \rangle \\ &\sim g_s^2 \sum_A \langle bb' | F_{1A} F_{2A} | cc' \rangle \end{aligned} \quad (14.87)$$

where we have suppressed all irrelevant factors except color indices. F_{1A} , F_{2A} are representation matrices of $SU(3)$ generators that act on particles 1 and 2. If we take

the sum over color indices A , Eq. (14.87) reduces to

$$\begin{aligned} \left\langle \sum_A F_{1A} F_{2A} \right\rangle &\equiv \langle \mathbf{F}_1 \cdot \mathbf{F}_2 \rangle = \frac{1}{2} \langle (\mathbf{F}_1 + \mathbf{F}_2)^2 - \mathbf{F}_1^2 - \mathbf{F}_2^2 \rangle \\ &= \frac{1}{2} \langle \mathbf{F}^2 - \mathbf{F}_1^2 - \mathbf{F}_2^2 \rangle \end{aligned} \quad (14.88)$$

which is a Casimir operator of $SU(3)$ [see Eq. (G.50) in Appendix G]. This is the $SU(3)$ extension of the isospin exchange force $\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2$ [see Eq. (13.27)]. If we assume that three-body forces are the sum of two-body forces, the sign and strength of its potential are proportional to

$$C_F = \sum_{i < j} \langle \mathbf{F}_i \cdot \mathbf{F}_j \rangle = \frac{1}{2} \left\langle \mathbf{F}^2 - \sum_{i=1}^3 \mathbf{F}_i^2 \right\rangle \quad \mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 \quad (14.89)$$

Values of C_F [refer to Eq. (G.51)] are given in Table 14.6. By surveying the table, one notices that a strong attractive force acts on color singlets and qualitatively reproduces the feature that only color singlets are observed in nature. The force acting on $qq[3^*]$ has to be attractive, being a part of qqq and making a singlet by coupling with $q[3]$. The coupling among $qqqq[3]$ is the same as for $qqq[1]$, meaning that attaching an extra q to qqq does not give a stronger attraction. Values of the force strength given in Table 14.6 are based on one-gluon exchange only. They cannot be considered as quantitatively correct as they do not take into account nonlinear couplings, which are special features of QCD. Nevertheless we are pleased to see that they give qualitatively correct results.

14.6.3

Spin Exchange Force

Before we go into the details of the color force, we review the space-time structure of the force. From what we have studied so far, the color force acting on the quarks

Table 14.6 Color factor $\sum \langle \mathbf{F}_j \cdot \mathbf{F}_j \rangle$.

Configuration	Multiplets	Strength
$q\bar{q}$	1	$-4/3$
$q\bar{q}$	8	$+1/6$
qq	3^*	$-2/3$
qq	6	$+1/3$
qqq	1	-2
qqq	8	$-1/2$
qqq	10	$+1$
$qqqq$	3	-2

is like the Coulomb force, at least at distances comparable with the hadron size. This suggests that we may acquire some insight into the level structure by studying that of the hydrogen atom, which is a bound state as a result of the Coulomb force (Fig. 4.1). The levels are obtained by solving the Dirac equation, but the principal part can be obtained from the Schrödinger equation with a central force α/r to give

$$E_n = -\frac{\alpha^2}{2} \frac{\mu}{n^2} \quad n = 1, 2, \dots \quad (14.90)$$

where μ is reduced mass. For hydrogen $m_1 = m_p \gg m_2 = m_e$ and $\mu \simeq m_e$.

The solutions of the Schrödinger equation having the same principal quantum number n but different L and spin have the same energy, for instance (2S,2P), (3S,3P,3D), ... are degenerate. Solving the Dirac equation, which takes into account spin-orbital angular momentum interaction (LS force) and spin-spin interaction (SS force), we find the levels split. Splitting due to the LS force is called the fine structure, and that due to spin-spin coupling the hyperfine structure. The LS force arises from the magnetic interaction between the magnetic moment μ_L produced by the circulating electron and the electron spin magnetic moment μ_e . The Hamiltonian is given by

$$H_{LS} \sim -\boldsymbol{\mu}_L \cdot \boldsymbol{\mu}_e = -\left(\frac{e}{2m_e} \mathbf{L}\right) \cdot \left(g_e \frac{e}{2m_e} \mathbf{s}_e\right) \sim -\frac{\alpha}{m_e^2} \mathbf{L} \cdot \mathbf{s} \quad (14.91)$$

The hyperfine structure arises from the spin-spin interaction, which is a magnetic interaction produced by spin magnetic moments of the electron and proton:

$$H_{ss} \sim -\mu_e \cdot \mu_p = \left(g_e \frac{e}{2m_e} \mathbf{s}_e\right) \cdot \left(g_p \frac{e}{2m_p} \mathbf{s}_p\right) \sim \frac{\alpha}{m_e m_p} \mathbf{s}_e \cdot \mathbf{s}_p \quad (14.92)$$

The interactions exchange the spin component of the interacting particle and are called spin exchange forces. As is clear from Eqs. (14.91) and (14.92), the different magnitude of the level splitting between the fine and hyperfine structures stems from the different masses of the electron and proton. In fact, the positronium shows the same structure as the hydrogen atom, but m_p is replaced by m_e , thus the LS and SS forces generate approximately the same magnitude of splitting. Calculation of the hyperfine splitting using quantum mechanics gives (see for instance [63])

$$E_{ss} = \frac{8\pi}{3} \frac{e_i e_j}{m_i m_j} |\psi(0)|^2 \mathbf{s}_i \cdot \mathbf{s}_j \quad (14.93)$$

$\psi(0)$ is the value of the wave function at the origin. Since

$$\mathbf{s}_i \cdot \mathbf{s}_j = \frac{1}{2} \left[(\mathbf{s}_i + \mathbf{s}_j)^2 - s_i^2 - s_j^2 \right] = \frac{1}{2} [s(s+1) - s_i(s_i+1) - s_j(s_j+1)] \quad (14.94)$$

where s is the total spin, the spin exchange force produces the energy difference between different s , hence the hyperfine splitting.

Since the gluon is a Lorentz vector just like the photon, we expect the same kind of forces to be generated by the gluon. That is to say, the spin-independent color electrostatic force is expected to generate the principal level splitting differentiated by n ($= 1, 2, \dots$) and the color magnetic force is expected to generate the fine structure, which is different depending on the value of L and S (not to be confused with the strangeness). The LS force is absent for $L = 0$ states.

14.6.4

Mass Formulae of Hadrons

Now we extend the level splitting force to include the color exchange. When the $SU(3)$ operator is written as $F(= F_1, F_2, \dots, F_8)$, the gluon exchange includes both color and the (real) spin exchange force. The Hamiltonian of the two-quark interaction is expected to have the form

$$H_{\text{color}} \sim (F_a \cdot F_b)(s_a \cdot s_b) \quad (14.95)$$

$$F_a \cdot F_b \equiv \sum_i^8 F_{ai} F_{bi}, \quad s_a \cdot s_b \equiv \sum_i^3 s_{ai} s_{bi}$$

Comparing this with the interaction of the hydrogen atom, we may interpret it as the color magnetic moment interaction that is generated by three color charges. Considering the mass and color charge strength, we see the Hamiltonian will have the form

$$H = H_0 - A \sum_{ab} \frac{\alpha_s}{m_a m_b} (F_a \cdot F_b)(s_a \cdot s_b) \quad (14.96)$$

where m_a, m_b are the quark masses and $\alpha_s = g_s^2/4\pi$ is the color coupling constant. In the above expression, we considered only low lying states with orbital angular momentum $L = 0$, hence there is no LS force and we are dealing with hyperfine splittings. H_0 generates the confining force, which has principal quantum numbers. The color exchange term $(F_a \cdot F_b)$ was evaluated in Table 14.6. Since all the hadrons are assumed to be color singlets, constituent quarks in the meson on which the operator $(F_a \cdot F_b)$ acts are in a singlet state. For baryons, it acts on a pair out of three quarks, but as the pair forms a singlet with the remaining q (3), the quark pair belongs to 3^* . We obtain

$$(F_a \cdot F_b) = \begin{cases} -4/3 & \text{meson} \\ -2/3 & \text{baryon} \end{cases} \quad (14.97)$$

If the quark masses are degenerate ($m_u = m_d = m_s$), we can sum over spin components. For mesons

$$\text{meson:} \quad (s_a \cdot s_b) = \begin{cases} -3/4 & 0^- \\ 1/4 & 1^- \end{cases} \quad (14.98a)$$

and for baryons

$$\text{baryon:} \quad \sum_{ab} (s_a \cdot s_b) = \begin{cases} -3/4 & (1/2)^+ \\ 3/4 & (3/2)^+ \end{cases} \quad (14.98b)$$

Therefore, the mass value differs depending on the spin. The above calculation gives $m(1^-) > m(0^-)$, $m(3/2) > m(1/2)$ and the relative separation is given by

$$m(1^-) - m(0^-) > m(3/2) - m(1/2) \quad (14.99)$$

which is consistent with observations. If there is no color exchange, the levels for mesons or baryons are inverted, contradicting the observations.

Problem 14.14

Derive Eqs. (14.98).

If we assume that the masses of hadrons and quarks are large compared to their binding energy (nonrelativistic approximation), the hadron mass in the first approximation [H_0 in Eq. (14.96)] can be expressed as a sum over quark masses. Then we have

$$\begin{aligned} \text{baryon:} \quad m_B &= m_1 + m_2 + m_3 + B \left[\frac{(s_1 \cdot s_2)}{m_1 m_2} + \frac{(s_2 \cdot s_3)}{m_2 m_3} + \frac{(s_3 \cdot s_1)}{m_3 m_1} \right] \\ &= \sum m_i + \sum_{ij} \frac{B}{m_i m_j} (s_i \cdot s_j) \end{aligned} \quad (14.100a)$$

$$\text{meson:} \quad m_M = m_1 + m_2 + \frac{C}{m_1 m_2} (s_1 \cdot s_2) \quad (14.100b)$$

For baryons there are a $(1/2)^+$ octet and a $(3/2)^+$ decuplet and considering the isospin degeneracy ($m_u \simeq m_d$), there are eight levels. Here, as an example, we calculate the mass of $\Lambda(1/2) - \Sigma(1/2) - \Sigma^*(3/2)$, which contain an s quark. We set $1 = s, 2, 3 = u, d$, and $m_2 = m_3 = m_u = m_d, m_1 = m_s$. Note, the total wave function including both flavor and spin should be totally symmetric. Since the quarks 2 and 3 make $I = 1$ (Σ, Σ^*) and $I = 0$ (Λ), the spin wave functions must be $s = 1$ ($I = 1$) and $s = 0$ ($I = 0$), respectively. For

$$\sum_{ij} \frac{s_i \cdot s_j}{m_i m_j} = \frac{s_2 \cdot s_3}{m_u^2} + \frac{s_1 \cdot (s_2 + s_3)}{m_u m_s} \quad (14.101)$$

we obtain the following:

$$\begin{aligned}
 2(s_2 \cdot s_3) &= (s_2 + s_3)^2 - s_2^2 - s_3^2 = s_{23}(s_{23} + 1) - 3/2 \\
 (s_2 \cdot s_3) &= \begin{cases} 1/4 & s_{23} = 1 \quad (\Sigma, \Sigma^*) \\ -3/4 & s_{23} = 0 \quad (\Lambda) \end{cases} \quad (14.102a) \\
 s_1 \cdot (s_2 + s_3) &= (1/2) \{s^2 - s_1^2 - (s_2 + s_3)^2\} \\
 &= (1/2) \{s(s+1) - 3/4 - s_{23}(s_{23} + 1)\}
 \end{aligned}$$

$$\begin{aligned}
 s(s+1) &= \begin{cases} 15/4 & s = 3/2 \quad (\Sigma^*) \\ 3/4 & s = 1/2 \quad (\Sigma, \Lambda) \end{cases} \\
 s_{23}(s_{23} + 1) &= \begin{cases} 2 & s_{23} = 1 \quad (\Sigma, \Sigma^*) \\ 0 & s_{23} = 0 \quad (\Lambda) \end{cases} \\
 \therefore s_1 \cdot (s_2 + s_3) &= \begin{cases} 1/2 & (\Sigma^*) \\ -1 & (\Sigma) \\ 0 & (\Lambda) \end{cases} \quad (14.102b)
 \end{aligned}$$

Inserting Eqs. (14.102) in Eqs. (14.101) and (14.100a) we obtain

$$m(\Sigma^*) = 2m_u + m_s + B \left(\frac{1}{4m_u^2} + \frac{1}{2m_u m_s} \right) \quad (14.103a)$$

$$m(\Sigma) = 2m_u + m_s + B \left(\frac{1}{4m_u^2} - \frac{1}{m_u m_s} \right) \quad (14.103b)$$

$$m(\Lambda) = 2m_u + m_s + B \left(-\frac{3}{4} \frac{1}{m_u^2} \right) \quad (14.103c)$$

Other baryons can be calculated similarly to get

$$m(\Lambda) = 3m_u + B \left(\frac{3}{4} \frac{1}{m_u^2} \right) \quad (14.104a)$$

$$m(N) = 3m_u + B \left(-\frac{3}{4} \frac{1}{m_u^2} \right) \quad (14.104b)$$

$$m(\Xi^*) = m_u + 2m_s + B \left(\frac{1}{4m_s^2} + \frac{1}{2m_u m_s} \right) \quad (14.104c)$$

$$m(\Xi) = m_u + 2m_s + B \left(\frac{1}{4m_s^2} - \frac{1}{m_u m_s} \right) \quad (14.104d)$$

$$m(\Omega) = 3m_s + B \left(\frac{3}{4} \frac{1}{m_s^2} \right) \quad (14.104e)$$

There are eight equations with three adjustable parameters. The mass formulae for mesons can be calculated in a similar manner. We compare our results with observed values in Fig. 14.29 and Table 14.7.

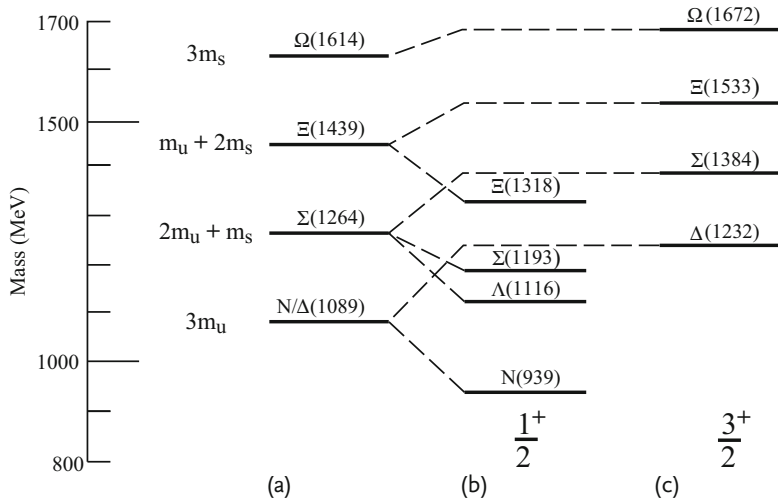


Figure 14.29 Levels of the baryon splitting due to spin–spin coupling [101]. (a) Levels due to quark mass splitting, (b) (c) spin-dependent force resolves the degeneracy.

Table 14.7 Mass values calculated including color–color and spin–spin couplings [296].

Baryon	Theory	Exp.	Meson	Theory	Exp.	
Ω	1682	1672	ϕ	1032	1020	For baryons
Ξ^*	1529	1533	K^*	896	892	$m_u = m_d = 363 \text{ MeV}$
Σ^*	1381	1384	ω	780	783	$m_s = 538 \text{ MeV}$
Δ	1239	1232	ρ	780	776	$B = 200 m_u^2 \text{ MeV}$
Ξ	1327	1318	η	559	549	For mesons
Σ	1179	1193	K	485	496	$m_u = m_d = 310 \text{ MeV}$
Λ	1114	1116	π	140	138	$m_s = 483 \text{ MeV}$
N	939	939				$C = 640 m_u^2 \text{ MeV}$

Problem 14.15

Derive Eqs. (14.104).

Problem 14.16

Calculate meson masses using Eq. (14.100b).

The quark masses determined from baryons and mesons may seem different, but considering the simple approximation we used, they are surprisingly close. Note that those determined from baryon magnetic moments [$m_u = 338 \text{ MeV}$, $m_d = 322 \text{ MeV}$, $m_s = 510 \text{ MeV}$, see Eq. (14.57)] are in between. These mass for-

mulae agree with the Gell-Mann–Okubo mass formula extended to $SU(6)$ [192]

$$m = A + BY + C[I(I+1) - Y^2/4] + DS(S+1) \quad (14.105)$$

to first order in power expansion of $\Delta m = m_s - m_u$.

Mass Formula Including Orbital Angular Momentum We have to include the spin–orbit coupling (LS force) to calculate the masses of states with orbital angular momentum $L > 0$. A representative formula, given by [117], has the form

$$H = L(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + \sum \left(m_i + \frac{\mathbf{p}_i^2}{2m_i} \right) - \frac{2}{3} \alpha_s \sum S_{ij} \quad (14.106a)$$

$$S_{ij} = \frac{1}{r} - \frac{1}{2m_i m_j} \left[\frac{\mathbf{p}_i \cdot \mathbf{p}_j}{r} + \frac{(\mathbf{r} \cdot \mathbf{p}_i)(\mathbf{r} \cdot \mathbf{p}_j)}{r^3} \right] \\ - \frac{1}{m_i m_j} \left[\frac{8\pi}{3} (s_i \cdot s_j) \delta^3(\mathbf{r}) + \frac{1}{r^3} \{ 3(s_i \cdot \hat{\mathbf{r}})(s_j \cdot \hat{\mathbf{r}}) - (s_i \cdot s_j) \} \right] \\ - \frac{1}{2r^3} \left[\frac{1}{m_i^2} (l_i \cdot s_i) - \frac{1}{m_j^2} (l_j \cdot s_j) + \frac{1}{m_i m_j} \{ (l_i \cdot s_j) - (l_j \cdot s_i) \} \right] \\ + \dots \text{(relativistic correction)} \quad (14.106b)$$

where $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$, $\mathbf{l} = \mathbf{r} \times \mathbf{p}$. The first term of Eq. (14.106a) gives the confinement potential and phenomenologically a harmonic oscillator potential is often used. The second term is the mass and kinetic energy term and the third term S_{ij} is

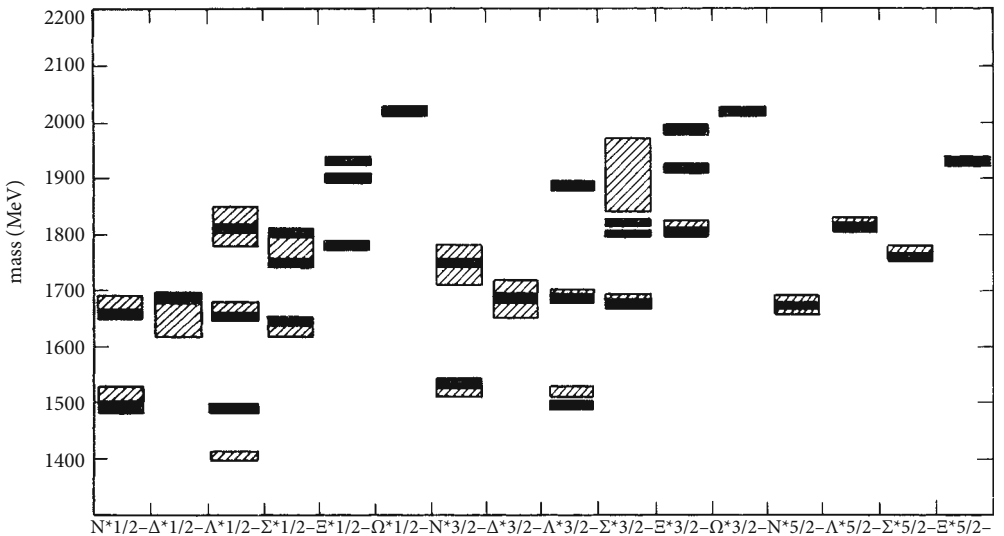


Figure 14.30 A comparison of the baryon mass prediction and experimental values based on the nonrelativistic quark model [214]. Hatched boxes are observed data of resonance masses and the black bars represent the prediction.

the “Breit interaction” (see for instance [58]) commonly used in nuclear physics. The second term in the first line of the Breit interaction is referred to as the Darwin term and the third term as the *zitterbewegung*, the second line represents the spin–spin interaction and the third line represents the spin–orbit interaction. Phenomenologically, however, it is known to reproduce the observed baryons’ mass and width when the spin–orbit interaction is neglected and only the spin–spin interaction is included. In the previous calculation, we used the first term in the spin–spin interaction term. As an example, Fig. 14.30 [204, 214] shows a comparison of the calculation with the observed baryon mass spectrum with $L = 1$ and negative parity.

With $\Lambda^*(1/2)^-$ as the only exception, the agreement is remarkable. The number of parameters was five.

It is not well understood why a nonrelativistic potential can reproduce the observed data so well.¹³⁾ However, the phenomenological success in hadron spectroscopy can certainly be considered strong evidence for the validity of the quark model and QCD.

13) Because of the asymptotic freedom, which will be explained in Chap. 18, the high-energy phenomena of QCD are fairly well understood, but the low-energy part of QCD needs nonperturbative approaches and remains a little-explored area to this day.

15

Weak Interaction

All hadrons and leptons are known to take part in the weak interaction, but its effect is felt only when the strong and electromagnetic interactions are forbidden by some selection rules. Since weak interaction phenomena were hard to reproduce experimentally, the progress in clarifying the weak process was slow at the beginning. But its very weakness made the perturbative approach effective and clear connections of theoretical ideas with observations were possible. Fermi's theory, which was proposed in 1934 and adapted to V–A interaction in 1957 to take parity violation into account, reproduced observed experimental data well. However, it is based on an effective Lagrangian valid only in first order and, lacking self-consistency, has no predictive power for higher order corrections. Nonetheless, its similarity with the electromagnetic interactions inspired the pursuit of a self-consistent mathematical framework modeled on QED – the only self-contained gauge theory at that time – and finally led to the unified theory of the electromagnetic and weak forces.

To help comprehend the logical flow that is described in the following, we start from known basic principles of the weak interaction, which serve as background knowledge, then follow in the footsteps of our great leaders, because it is instructive to learn how they gained insight from observations. It is especially illuminating to understand the similarities to and the differences from the electromagnetic force at each step. The purpose is to learn heuristic methods as well as to study the facts and logic of elementary particle physics in the weak force world.

15.1

Ingredients of the Weak Force

The weak force is generated by the “weak charge”, just as the electromagnetic force is generated by the electric charge. This means there is a “weak static force” as well as a “weak magnetic force”, the former produced by the static weak charge and the latter by its current, i.e. the moving charge. Correspondingly, the force carrier (denoted as W^μ) is a four-vector and couples to a four-current j_W^μ (weak current and $H_{W1} = g_W j_W^\mu W_\mu$) generated by the motion of the weak charge carrier (i.e. quarks and leptons) just as the photon is a Lorentz vector that couples to the

electric current ($H_{\text{EM}} = e j_e^\mu A_\mu$). The coupling constant g_W is the strength of the weak charge, corresponding to the unit electric charge e .

The first major difference is that the weak charge comes in two kinds, to be denoted temporarily as Q_{W1} , Q_{W2} , which we call the “weak isospin doublet”. For example, the electron–neutrino has Q_{W1} and the electron has Q_{W2} , and together they form a weak doublet (ν_e, e^-) . The term “weak doublet” is used to distinguish it from the isospin doublets that appeared in the strong interaction. The naming is not accidental, though. The generators of the nuclear force, the proton and neutron, constitute an isospin doublet (p, n) and change their identity by the nuclear force action (referred to as the exchange force), because the force carriers (i.e. the pions) are charged. Like the pion, the W boson also comes in three kinds and constitutes a weak isospin triplet (W^+, Z^0, W^-) .¹⁾ Since the lepton is a fermion and fermion number is known to be conserved, we assign lepton number 1 to all leptons and -1 to all antileptons. The sum of the lepton number is conserved in the weak as well as in other interactions. The action and reaction of the weak force (absorption and emission of the W) change the identity of the weak charge carriers. For instance, by emitting W^+ (or by absorbing W^-), ν_e changes into e^- . The six known leptons are classified into three doublets. A notable feature of the weak force is that it changes the lepton only into itself or into the other member of the same doublet. Namely, the three doublets in the lepton family (ν_e, e^-) , (ν_μ, μ^-) , (ν_τ, τ^-) are independent. They do not mix with each other, which is referred to as lepton flavor conservation.²⁾

$$\begin{array}{lll} \text{Quark doublets:} & \begin{bmatrix} u \\ d \end{bmatrix} & \begin{bmatrix} c \\ s \end{bmatrix} & \begin{bmatrix} t \\ b \end{bmatrix} \\ \text{Lepton doublets:} & \begin{bmatrix} \nu_e \\ e^- \end{bmatrix} & \begin{bmatrix} \nu_\mu \\ \mu^- \end{bmatrix} & \begin{bmatrix} \nu_\tau \\ \tau^- \end{bmatrix} \end{array}$$

The situation in the quark sector is more complicated. Its known six members also constitute three doublets in the weak interaction. However, mixing among them, known as CKM (Cabibbo–Kobayashi–Maskawa) mixing, occurs.

The second major difference between the weak and electromagnetic forces is their reach. The force range r_0 is determined by the Compton wavelength of the force carrier. Since the masses of $W^\pm$ and Z^0 are

$$M_W = 80.40 \pm 0.025 \text{ GeV}, \quad M_Z = 91.186 \pm 0.002 \text{ GeV} \quad (15.1)$$

r_0 is $\sim 2 \times 10^{-18}$ m. For practically all reactions we deal with in this chapter, the energy in a process is much smaller than m_W . This means that the mass of W , Z can be regarded as infinite or equivalently the force range as zero. Namely, the

1) The reason for denoting the neutral member of the W family not as W^0 but as Z^0 is because it is a mixture of W^0 with the photon, changing its identity and at the same time acquiring a different mass as well as coupling strength from its charged partners.

But in this chapter the distinction of W^0 from Z^0 does not matter.

2) The neutrino oscillation, discovered in 1998 (to be discussed in Sect. 19.2.1), mixes them, but mixing among the charged leptons has not been observed. In this book we assume lepton flavor is conserved.

weak force is effective only when the participating particles are in direct contact with each other.

15.2

Fermi Theory

15.2.1

Beta Decay

Neutrino Hypothesis The weak interaction was first discovered in nuclear beta (β) decays back in 1896 by Becquerel. In alpha (α) and gamma (γ) decays discovered at the same time, the emitted particles had a monochromatic energy spectrum, which agreed with the difference of the energy between the parent and the daughter nuclei. However, in β decay, the energy of the emitted β rays, i.e. of the electrons, was continuous and there was no sign of other particles. Namely, the law of conservation of energy seemed to be violated. Bohr was willing to abandon the law of conservation of energy, but Pauli proposed in 1930 that a yet unknown particle (neutrino) carries away the missing energy. In 1934, taking into account Pauli's neutrino hypothesis, Fermi proposed the first weak interaction theory to explain the phenomenon. Its formalism was made in analogy with the electromagnetic interaction. The gamma decay process, in which a proton in a nucleus emits a photon (Fig. 15.1a) can be described using an interaction Lagrangian

$$\mathcal{L}_{\text{EM}} = -q j^\mu A_\mu = -e \bar{p} \gamma^\mu p A_\mu \quad (15.2)$$

Here, $p, \bar{p}, A_\mu$ are field operators of the proton and the photon. Namely, in the quantum field theoretical interpretation, the operator p annihilates a proton and A_μ creates a photon and at the same time $\bar{p}$ recreates a proton. The β decay is a process in which a neutron in the nucleus is converted into a proton and at the same time emits an electron–neutrino pair instead of the photon (Fig. 15.1b):⁴⁾

$$n \rightarrow p + e^- + \bar{\nu}_e \quad (15.4)$$

In terms of the field operators, the process can be described by a four-fermion interaction Lagrangian

$$\mathcal{L}_\beta = -G_\beta (j_h^+)_{\mu} (j_l^-)^{\mu} = -G_\beta (\bar{p} \gamma^\mu n) (\bar{e} \gamma_\mu \nu) + h.c. \quad (15.5)$$

- 3) The proton is not a Dirac particle, as it has a large anomalous magnetic moment and is spatially extended. Therefore, interactions with the anomalous magnetic moment
- $$\mu_p = g_p e / 2m_p$$

$$\mathcal{L} = \frac{g_p - 2}{2} \frac{e}{2m_p} \sigma^{\mu\nu} (\partial_\mu A_\nu - \partial_\nu A_\mu) \quad (15.3)$$

- have to be included from the phenomenological viewpoint. Equation (15.2) is to be considered as a starting point.
- 4) It is actually an antineutrino in modern terminology.

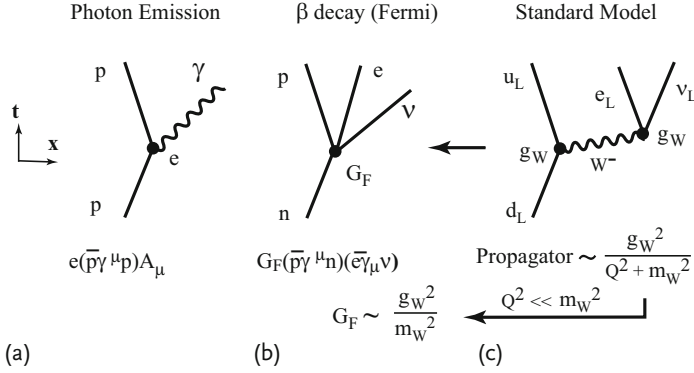


Figure 15.1 The weak interaction was conceived in analogy with the electromagnetic interaction. (a) A field theoretical illustration of a photon emission. (b) Fermi interaction: Instead of γ , an $(e^-\bar{\nu}_e)$ pair is emitted. (c)

The β decay in the Standard Model. A d quark in the neutron is converted to a u quark with the emission of $e^-\bar{\nu}_e$. Note the $W^\pm$ couples only to left-handed particles, hence the subscript L.

where $h.c.$ stands for the hermitian conjugate and is added to make the Lagrangian hermitian. We omit the subscript e from ν in the four-fermion interaction until the difference between ν_e and ν_μ is discovered. This was the first application of Heisenberg and Pauli's quantized field theory. In the case of photon emissions, quantum field theory could be considered as a reformulation of known formalism, but in the β decay the electron and the neutrino were literally created inside the nucleus. G_β is called the Fermi coupling constant and is a measure of the weak force strength. The Lagrangian is a product of two currents, hadronic and leptonic, and is called current-current coupling. To distinguish them from the electromagnetic current, they are called weak currents. The subscripts h,l attached to the currents are abbreviations of hadronic or leptonic. Unlike the electromagnetic current, which does not change the identity of particles, the weak currents change the electric charge [$n \rightarrow p$ and $\nu \rightarrow e^-$ or (vacuum $\rightarrow e^-\bar{\nu}$)] and are called charged currents. In this sense the electromagnetic current is a kind of neutral current.

In the Standard Model, the charged current couples to $W^\pm$:

$$\begin{aligned}\mathcal{L}_{WI} &= g_W \left[j_h^{+\mu} W_\mu^+ + j_l^{-\mu} W_\mu^- \right] + h.c. \\ &= g_W \left[(\bar{u}_L \gamma^\mu d_L) W_\mu^+ + (\bar{e}_L \gamma^\mu \nu_{eL}) W_\mu^- \right] + h.c.\end{aligned}\quad (15.6)$$

Here, $W_\mu^{+(-)}$ is a field operator to annihilate (create) $W^{+(-)}$ bosons. Beta decay in the Standard Model is the two-step process depicted in Fig. 15.1c, whose effective Lagrangian can be expressed as [see for instance Eq. (7.12)]

$$\mathcal{L}_{SM, 2nd\ order} = g_W^2 (\bar{u}_L \gamma^\mu d_L) \frac{g_{\mu\nu}}{q^2 - m_W^2} (\bar{e}_L \gamma^\nu \nu_{eL}) \quad (15.7)$$

where $q^2 = (p_d - p_u)^2$ is the momentum transfer squared. Since the energy difference of the nucleus before and after the decay is a few MeV and can be neglected

compared to $m_W = 80 \text{ GeV}$, the Lagrangian can be rewritten as

$$\mathcal{L}_{\text{SM, 2nd order}} \simeq -G_\beta (\bar{u}_L \gamma^\mu d_L) (\bar{e}_L \gamma_\mu \nu_{eL}) , \quad G_\beta = \frac{g_W^2}{m_W^2} \quad (15.8)$$

which essentially reproduces the Fermi interaction, except that only left-handed particles participate in the Standard Model, as is indicated by subscript L to the field operators. The left-handed nature of the weak interaction will be explained later in connection with parity violation. The Lagrangian of this type is called the four-fermion interaction.

Allowed Transitions To obtain the rate of β decay, it is necessary to calculate the transition amplitude from an initial to a final state:

$$T_{fi} = G_\beta \int d^4x \langle p | \bar{p}(x) \gamma^\mu n(x) | n \rangle \langle e \bar{\nu} | \bar{e}(x) \gamma_\mu \nu(x) | 0 \rangle \quad (15.9)$$

For a while, we follow the historical development, treating the proton and neutron as elementary particles. If we consider them as represented by plane waves, the nucleonic part of Eq. (15.9) is

$$\langle p | j_h^\mu(x) | n \rangle \sim \bar{u}(p') \gamma^\mu u(p) e^{-i(p-p') \cdot x} \quad (15.10)$$

where p, p' are four-momenta of the proton and the neutron. As the Q value (energy difference before and after the reaction) of the β decay is at most a few MeV and overwhelmingly small compared to the nuclear mass, the recoil of the nucleus can be neglected. Referring to Table 4.2, only the time component survives in the nonrelativistic limit and is proportional to 1:

$$\int d^3x \langle p | \bar{p}(x) \gamma^\mu n(x) | n \rangle \simeq \delta_{\mu 0} u^\dagger(p') u(p) \delta^3(\mathbf{p}' - \mathbf{p}) \equiv \delta_{\mu 0} \langle 1 \rangle \quad (15.11)$$

In practice, the proton and neutron are in a nucleus. The expectation value in Eq. (15.11) should be calculated not by using the free nucleon wave functions but by using initial and final wave functions of the nucleus. As the operation includes a change of a neutron to a proton, the operator is not exactly 1 but $1\tau_+$, where $\tau_\pm = (\tau_x \pm i\tau_y)/2$ is the raising (lowering) operator of the isospin.

As for the electron and neutrino, they must be treated relativistically because their mass is much smaller than typical β -decay energies. Namely, the leptonic current is expressed as

$$\langle e \bar{\nu} | j_e^\mu(x) | 0 \rangle \sim \bar{u}(p_e) \gamma^\mu v(p_\nu) e^{i(p_e + p_\nu) \cdot x} \quad (15.12)$$

Therefore, the transition matrix element becomes, after integration over the time variable,

$$\begin{aligned} T_{fi} &= 2\pi \delta(\Delta - E_e - E_\nu) G_\beta \int d^3x \Phi_F^*(x) \Phi_I(x) e^{-i(p_e + p_\nu) \cdot x} \\ &\quad \times [\bar{u}(p_e) \gamma^0 v(p_\nu)] \\ \Delta &= M(Z, A) - M(Z+1, A) \end{aligned} \quad (15.13)$$

where Φ_I , Φ_F are wave functions of the nucleus before and after the decay, respectively, and Δ is the mass difference of the parent and daughter nucleus. As $|\mathbf{p}_e|$, $|\mathbf{p}_\nu| \lesssim$ a few MeV, the de Broglie wavelength of the electron and neutrino is $\sim 10^{-13}$ m, which is much larger than the nuclear size $\sim 10^{-14}$ m. This means $(\mathbf{p}_e + \mathbf{p}_\nu) \cdot \mathbf{x} \ll 1$ in the overlap region of the nuclear wave functions. Therefore the higher powers in the expansion

$$e^{-i(\mathbf{p}_e + \mathbf{p}_\nu) \cdot \mathbf{x}} = 1 - i(\mathbf{p}_e + \mathbf{p}_\nu) \cdot \mathbf{x} + \dots \quad (15.14)$$

can be neglected in the first approximation. The allowed processes under this condition are called “allowed transitions”. If some selection rule is effective and the first term does not contribute, one needs to consider the higher order terms (forbidden transition), but we do not discuss them here as they are more relevant for nuclear physics. For the allowed transition, the nuclear part

$$\langle 1 \rangle = \int d^3x \Phi_F^* \Phi_I \quad (15.15)$$

is separated from the leptonic part and is a constant that does not include any dynamical variables.

ft Value Calculating the decay rate from Eq. (15.13), we have

$$d\Gamma = G_\beta^2 |\langle 1 \rangle|^2 \sum_{\text{spin}} L^{00} 2\pi \delta(\Delta - E_e - E_\nu) \frac{d^3p_e}{(2\pi)^3 2E_e} \frac{d^3p_\nu}{(2\pi)^3 2E_\nu} \quad (15.16a)$$

$$\begin{aligned} \sum_{\text{spin}} L^{00} &= \sum_{r,s} \bar{u}_r(p_e) \gamma^0 v_s(p_\nu) \bar{v}_s(p_\nu) \gamma^0 u_r(p_e) = \text{Tr}[\not{p}_e \gamma^0 \not{p}_\nu \gamma^0] \\ &= 4[2p_e^0 p_\nu^0 - g^{00}(p_e \cdot p_\nu)] = 4E_e E_\nu (1 + v_e \cos \theta) \end{aligned} \quad (15.16b)$$

where $v_e = p_e/E_e$ is the electron velocity and θ is the angle between the electron and the neutrino. $\sum_{\text{spin}} L^{\mu\nu}$ is given in Eq. (6.143). Integration over the phase space gives

$$\begin{aligned} \delta(\Delta - E_e - E_\nu) p_\nu^2 dp_\nu &= p_\nu E_\nu \\ p_\nu E_\nu &= (\Delta - E_e)^2 \quad \text{if } m_\nu = 0 \end{aligned} \quad (15.17)$$

which gives

$$d\Gamma = G_\beta^2 |\langle 1 \rangle|^2 (1 + v \cos \theta) \frac{1}{(2\pi)^5} p E (\Delta - E)^2 dE d\Omega_e d\Omega_\nu \quad (15.18)$$

Here, we have put $p = p_e$, $E = E_e$, $v = v_e$. From the expression we know that the electron energy spectrum is wholly determined by the phase space volume. We define

$$f \equiv \int_{m_e}^{\Delta} F(Z, E) p E (\Delta - E)^2 dE \quad (15.19)$$

$F(Z, E)$ is a correction factor due to the electron Coulomb wave function and is a known function [280]. When the electron is nonrelativistic it is approximated by

$$F(Z, E) \simeq \frac{2\pi Z\alpha/v}{1 - \exp(-2\pi Z\alpha/v)} \quad (15.20)$$

For light to medium elements $F \sim 1$, f grows in proportion to $\sim \Delta^5$ and for a small change of Δ (i.e. Q -value of the decay) the lifetime changes drastically from 10^{-8} to 100 s. If we use the half lifetime $t_{1/2} = \ln 2/\Gamma$, the total decay rate is written as

$$\Gamma = \int d\Gamma = \frac{G_\beta^2}{2\pi^3} |\langle \mathbf{1} \rangle|^2 f = \frac{\ln 2}{t_{1/2}} \quad (15.21)$$

We define the ft value by

$$ft_{1/2} = \frac{2\pi^3 \ln 2}{G_\beta^2 |\langle \mathbf{1} \rangle|^2} \quad (15.22)$$

As f , $|\langle \mathbf{1} \rangle|$ are calculable using the nuclear shell model and $t_{1/2}$ is an observable, we can determine G_β .

Fermi and Gamow–Teller Transitions In allowed transitions, the orbital angular momentum of the lepton pair relative to the nucleus vanishes and if there is no spin flip in the nuclear transition; the angular momentum transfer from the initial to the final state of the nucleus is $\Delta J = 0$ with no parity change. This is called the Fermi transition. In actual β decays, transitions with the angular momentum transfer equal to one ($\Delta J = 0, \pm 1$, but $0 \rightarrow 0$ is forbidden⁵⁾) with no parity change (Gamow–Teller transition) have been observed, hence the original Lagrangian proposed by Fermi was not sufficient to explain all the observed β decays. Furthermore, parity violation was discovered in 1957 and its effect must also be taken into account. The most general form of the four-fermion interaction Lagrangian for β decay can be written as follows:

$$\mathcal{L}_\beta = -G_\beta \sum [(\bar{p}\Gamma^i n)(C_i \bar{e}\Gamma_i \nu + C'_i \bar{e}\Gamma_i \gamma^5 \nu) + h.c.] \quad (15.23)$$

$$\Gamma^i = 1(S), \quad \gamma^\mu(V), \quad \sigma^{\mu\nu}(T), \quad \gamma^\mu \gamma^5(A), \quad i\gamma^5(P)$$

where S, V, T, A and P represent scalar, vector, tensor, axial vector and pseudoscalar, respectively. By setting $C_V = 1$ and $C_i = C'_i = 0$ for all others, Eq. (15.23) reduces to the original Fermi interaction Lagrangian Eq. (15.5). For more general arguments, interactions including derivatives have to be considered, but we omit them as the inclusion of derivatives changes the energy spectrum and there is no sign of such terms. Since C_i , C'_i are complex numbers, there are 20 parameters in total. Using the transformation property of each term (see Appendix F), one sees that $C'_i = 0$ for P invariance, $C_i(C'_i)$ are real (pure imaginary) for C invariance and all C_i , C'_i must be real for T invariance.

5) See Problem 13.1 (3).

Problem 15.1*

Prove the above statement.

In the real world, P and C are maximally violated and T (or CP if CPT invariance holds) invariance is almost true. In the following discussions, we assume T invariance until proven otherwise and set all C_i , C'_i as real, reducing the number of parameters to 10. For a while we proceed assuming P, C invariance as well, and set $C'_i = 0$. In this case, the number of parameters is 5. In the approximation neglecting nuclear recoils, the transition matrix elements can be calculated using Table 4.2, and we obtain

$$T_{fi} \sim G_\beta \left[\langle 1 \rangle \{ C_S \bar{u}(p_e) v(p_\nu) + C_V \bar{u}(p_e) \gamma^0 v(p_\nu) \} + \langle \sigma \rangle \cdot \{ C_A \bar{u}(p_e) \gamma^0 \sigma v(p_\nu) + C_T \bar{u}(p_e) \sigma v(p_\nu) \} \right] \quad (15.24a)$$

$$\langle 1 \rangle \equiv \int d^3x \Phi_F^*(s) \tau^+ \Phi_I(s) \quad (15.24b)$$

$$\langle \sigma \rangle \equiv \int d^3x \Phi_F^*(x) \sigma \tau^+ \Phi_I(x) \quad (15.24c)$$

Resultant expressions for the decay rates have the same form as Eqs. (15.18) and (15.21) but the matrix elements squared $|\langle 1 \rangle|^2$ and the angular correlation coefficient 1 are replaced by [156, 280]

$$|\langle 1 \rangle|^2 \rightarrow C_F^2 |\langle 1 \rangle|^2 + C_{GT}^2 |\langle \sigma \rangle|^2 \quad (15.25a)$$

$$C_F^2 = |C_S|^2 + |C_V|^2 + \frac{2m_e}{E_e} C_S C_V \quad (15.25b)$$

$$C_{GT}^2 = |C_A|^2 + |C_T|^2 - \frac{2m_e}{E_e} C_A C_T \quad (15.25c)$$

$$1 + \nu \cos \theta \rightarrow 1 + \lambda \nu \cos \theta$$

$$\lambda = \frac{(-|C_S|^2 + |C_V|^2) |\langle 1 \rangle|^2 + (1/3) (-|C_A|^2 + |C_T|^2) |\langle \sigma \rangle|^2}{C_F^2 |\langle 1 \rangle|^2 + C_{GT}^2 |\langle \sigma \rangle|^2} \quad (15.25d)$$

If the parity-violating terms are included, all the coefficients are replaced by

$$|C_i|^2 \rightarrow |C_i|^2 + |C'_i|^2 \quad (15.26)$$

Later, $|C'_i| = |C_i|$ was clarified, and consequently Eqs. (15.25) are equally valid after the discovery of parity violation if we use $G_\beta/\sqrt{2}$ instead of G_β . The reader may have noticed that P-type interaction is missing. It does not contribute to the allowed transitions. This is because P-type coupling is equivalent to derivative coupling in the low-energy limit (see Table 4.2 and discussions in Sect. 13.6.1); the matrix element is proportional to the recoil momentum of the nucleus, hence can be neglected. When the interaction type is S or V, the selection rule is the same as V, namely $\Delta J = \Delta P = 0$, and is referred to as the Fermi transition. When

the interaction is A or T, the electron and the neutrino carry away the angular momentum 1 generated by the spin flip nuclear transition. This is referred to as the Gamow–Teller transition, in which $\Delta J = 0, \pm 1$, $\Delta P = 0$.

Problem 15.2

Prove that if $F(Z, E) \simeq 1$, $E_e \gg m_e$, then $f = \Delta^5/30$ and the total decay rate becomes

$$\Gamma = \frac{G_\beta^2 \Delta^5}{60\pi^3} (C_F^2 |\langle \mathbf{1} \rangle|^2 + C_{GT}^2 |\langle \boldsymbol{\sigma} \rangle|^2) \quad (15.27)$$

Determination of the Interaction Type When the initial and final states of nuclear transitions are known, the mixing rate of the Fermi and Gamow–Teller transitions

$$F = \frac{C_F^2 |\langle \mathbf{1} \rangle|^2}{C_F^2 |\langle \mathbf{1} \rangle|^2 + C_{GT}^2 |\langle \boldsymbol{\sigma} \rangle|^2} \quad (15.28)$$

is a function of $|C_{GT}/C_F|$. If the ft value is measured as a function of F , $|C_{GT}/C_F|$ can be determined [233]. The experimental values are given by

$$\begin{aligned} G_\beta &= 1.14730 \pm 0.00064 \times 10^{-5} \text{ GeV}^2 \\ |C_{GT}/C_F| &= 1.2601 \pm 0.0025 \end{aligned} \quad (15.29)$$

When one interaction type dominates, the angular correlation coefficient λ is given by

$$\begin{aligned} \lambda &= -1 (S), \quad = +1 (V), \\ &= 1/3 (T), \quad = -1/3 (A) \end{aligned} \quad (15.30)$$

Measuring the angular coefficient λ for various nuclei, and plotting it as a function of mixing parameter F , one can determine which of the interactions are effective (Fig. 15.2). Historically, the interaction type was determined this way. The result is

$$\begin{aligned} C_S &= C_T = 0, \\ C_V &= 1, \quad C_A \simeq -1.260 C_V \end{aligned} \quad (15.31)$$

The sign of C_A was determined by angular correlation of $(e\nu)$ and of the neutrino with respect to the spin of free polarized neutrons [82, 83].

Beta Decay of the Neutron When the initial and final states belong to the same isospin multiplet, the values of $\langle \mathbf{1} \rangle$, $\langle \boldsymbol{\sigma} \rangle$ can be calculated without knowing the nuclear wave functions. For the neutron β decay, $|p\rangle$, $|n\rangle$ are eigenstates of isospin $I = 1/2$,

$$\langle \mathbf{1} \rangle^2 = 1, \quad \langle \boldsymbol{\sigma} \rangle^2 = 3 \quad (15.32)$$

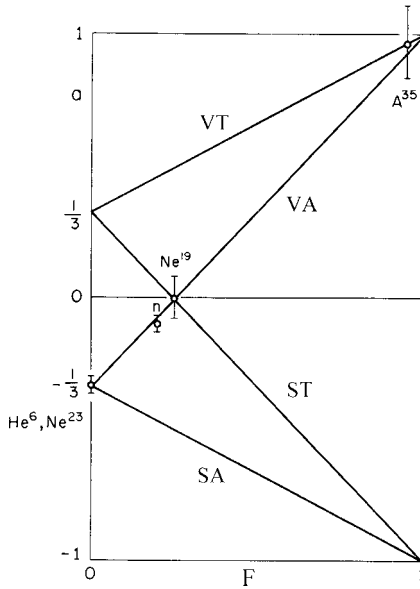


Figure 15.2 Electron–neutrino angular correlation coefficients plotted against the mixing parameter F of the decays in the allowed transitions [280].

As $\Delta = m_n - m_p = 1.29 \text{ MeV}$ is the same order as the electron mass $m_e = 0.511 \text{ MeV}$, the expression for Γ in Eq. (15.27) is not usable, but integration of f and consideration of the weak magnetism, which will be explained later, gives a value that agrees well with the experimental value ($\tau_n = 885.7 \pm 0.8 \text{ s}$).

15.2.2

Parity Violation

Probably the most shocking event in the history of particle physics was the discovery of parity violation in 1957. Today any conservation law is not sacred, considered as an entity to be checked by observations, but at that time it was beyond imagination that such a fundamental law as space-time symmetry could be broken. The first indication was the so-called τ – θ puzzle. The particle, known as the K meson today, decays to 3π (τ particle) at one time and to 2π (θ particle) at another. Both of them were known to have spin $s = 0$. Since the pion is known to have spin-parity $J^P = 0^-$, to make $J = 0$ in the 2π combination, the orbital angular momentum has also to be zero, resulting in positive parity. In the 3π system, taking l as the orbital angular momentum of 2π and L as that of the third π relative to the center of mass of the 2π , $l + L = J = 0$ must hold. This means $l = L$ and the parity is $P = (-1)^{3+l+L} = -1$. So τ and θ must be different particles. However, the mystery was that they had the same mass and the same lifetime, indicating they were one and the same particle.

Lee and Yang investigated past experiments and found that there was ample evidence of parity conservation in the strong and electromagnetic interactions but no evidence in the weak interaction. They proposed several experiments to test the validity [258].

How to Test Parity Violation The momentum $\mathbf{p}$ changes its sign by parity transformation but the angular momentum $\mathbf{J} \sim \mathbf{r} \times \mathbf{p}$ does not. Therefore if an observable such as $\mathbf{J} \cdot \mathbf{p}$ is found in a decay process, it is evidence of parity violation. Let $\mathbf{p}$ be the momentum of decay electrons from a polarized nucleus having angular momentum $\mathbf{J}$. Then $\mathbf{J} \cdot \mathbf{p}$ is proportional to the cosine of the electron emission angle θ with respect to the polarization axis. Therefore, by observing β decay electrons from polarized nuclei, one can test the existence of the term $\mathbf{J} \cdot \mathbf{p}$. It appears as up-down asymmetry with respect to a plane perpendicular to the polarization axis.

Wu's Experiment C.S. Wu, who was at the same university as Lee and Yang, made polarized ^{60}Co by cooling it down to 0.01 K and measured the angular distribution of the decay electrons. Figure 15.3a shows her apparatus [392]. By applying a weak magnetic field (a few hundred gauss⁶), electrons in the outer shell of the atom can be polarized. The nucleus is polarized dynamically by spin-spin in-

6) One gauss (1 Gs) is 10^{-4} tesla (10^{-4} T).

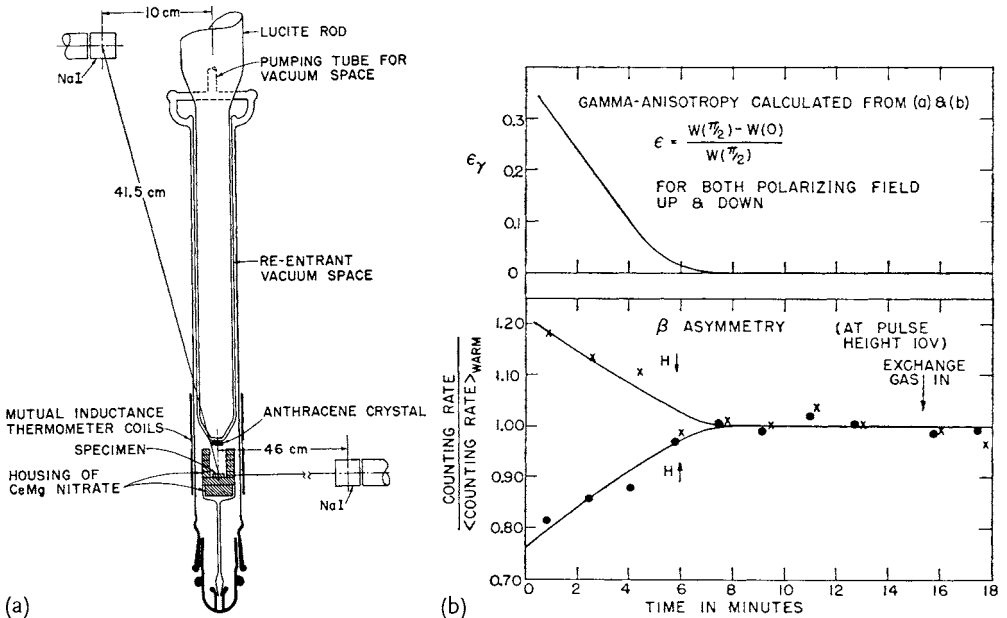


Figure 15.3 (a) Apparatus used to discover parity violation. (b) Observed data. $^{60}\text{Co}(5^+)$ β decays to $^{60}\text{Ni}(4^+)$ followed by gamma transitions $4^+ \rightarrow 2^+ \rightarrow 0^+$. The agreement

between the γ anisotropy and the asymmetry of the electron is evidence that the asymmetry exists and is proportional to the polarization [392].

teraction with the electron. The degree of nuclear polarization was determined by measuring simultaneously the anisotropy of its γ radiation in the transition $^{60}\text{Ni}(4^+) \rightarrow 2^+ \rightarrow 0^+$:

$$\epsilon = \frac{W(\pi/2) - W(0)}{W(\pi/2)} \quad (15.33)$$

Here, $W(\theta)$ is the relative intensity of the emitted gamma ray at θ with respect to the polarization axis. Since $^{60}\text{Co}(5^+) \rightarrow ^{60}\text{Ni}(4^+)$ is a Gamow–Teller transition, we obtain the angular distribution by setting $C_i = C'_i = 0$, $i \neq A$ in Eq. (15.23). The resultant angular distribution formula is

$$S(\theta) \sim 1 + APv \cos \theta$$

$$A = \frac{2\text{Re}(C_A C'_A{}^*)}{|C_A|^2 + |C'_A|^2} \quad (15.34)$$

where P is the polarization of the Co and v the electron velocity. The experimental result was $P = 0.65$, $v \sim 0.6$, $APv = -0.38$, which means

$$A = -1 \pm 0.1 \Rightarrow C'_A \sim -C_A \quad (15.35)$$

Therefore the parity violation is maximal. It also shows the maximum C violation, because C'_A is not pure imaginary with magnitude as big as C_A . It was later confirmed that for V type β decay also

$$C'_V \sim -C_V \quad (15.36)$$

Inserting Eqs. (15.31), (15.35) and (15.36) in Eq. (15.23), one can express the Lagrangian for β decay as

$$\mathcal{L}_\beta = -\frac{G_\beta}{\sqrt{2}} [\bar{p} \gamma^\mu (1 - 1.26\gamma^5) n] [\bar{e} \gamma_\mu (1 - \gamma^5) \nu] \quad (15.37)$$

The factor $1/\sqrt{2}$ is inserted in order to use the same G_β as before the discovery of parity violation. The Lagrangian (15.37) contains V and A type with approximately equal strength and opposite sign and is commonly referred to as $V-A$ type. We have now obtained an important characteristic of the weak interaction. Because of $1-\gamma^5$, as far as the lepton is concerned only components with left chirality contribute to the interaction. For the notion of chirality, see Sect. 4.3.5. We defer consideration of the hadrons (or quarks), where the axial vector coupling differs by as much as 26%, until we discuss the conserved vector current (CVC) hypothesis in Sect. 15.5.2.

15.2.3

π Meson Decay

We determined the coupling constants of S , V , T and A from β decays, but as the P type does not contribute in allowed transitions, we need to investigate pion

decay. Let us first calculate the decay rate using the interaction Lagrangian (15.37) as determined by β decay. Almost 100% of the pions decay as

$$\pi^- \rightarrow \mu^- + \bar{\nu}, \quad \pi^+ \rightarrow \mu^+ + \nu \quad (15.38)$$

with lifetime [311]

$$\tau = 2.603 \pm 0.0005 \times 10^{-8} \text{ s} \quad (15.39)$$

A conspicuous feature of pion decay is that the branching ratio of $\pi \rightarrow e + \nu$ is extremely small:

$$\text{BR}(\pi^\pm \rightarrow e^\pm \bar{\nu}) = \frac{\Gamma(\pi^\pm \rightarrow e^\pm \bar{\nu})}{\Gamma(\pi^\pm \rightarrow \text{all})} = 1.230 \pm 0.004 \times 10^{-4} \quad (15.40)$$

Considering its larger phase space volume, one would naively expect a larger decay rate than for $\pi \rightarrow \mu \nu$, an expectation not fulfilled by observation. To investigate the reason, let us write down the variables in the reaction as

$$\pi(k) \rightarrow l(p) + \nu(q) \quad (15.41)$$

Since the muon has the same properties as the electron except for its mass, we infer its emission can be treated in the same way as the β decay. Therefore, we simply make the replacement $e \rightarrow \mu$ in Eq. (15.37) and the interaction Lagrangian should have the form

$$\mathcal{L}_{\text{WI}} = -\frac{G_\beta}{\sqrt{2}} [\bar{p} \gamma^\mu (1 - 1.26\gamma^5) n] [\bar{\mu} \gamma_\mu (1 - \gamma^5) \nu] \quad (15.42)$$

The transition amplitude becomes

$$\begin{aligned} T_{fi} &= \langle \mu \nu | -\mathcal{L}_{\text{WI}} | \pi \rangle = (2\pi)^4 \delta^4(k - p - q) \\ &\times \frac{G_\beta}{\sqrt{2}} \langle 0 | \bar{p} \gamma^\mu (1 - 1.26\gamma^5) n | k \rangle \langle p, q | \bar{\mu} \gamma_\mu (1 - \gamma^5) \nu | 0 \rangle \end{aligned} \quad (15.43)$$

The hadronic part contains corrections due to strong interaction (wave functions made by quarks) and cannot be calculated easily from first principles, but using symmetry arguments some restrictions can be applied. From Lorentz invariance, the hadronic part should be able to be expressed as

$$\langle 0 | \bar{p} \gamma^\mu (1 - 1.26\gamma^5) n | \pi \rangle = -1.26 \langle 0 | \bar{p} \gamma^\mu \gamma^5 n | \pi \rangle \equiv -ik^\mu f_\pi \quad (15.44)$$

The first term in the parentheses does not contribute because the pion has negative parity. The restriction comes from the vector nature of the current and that the pion momentum is the only vector that appears in the process. f_π is referred to as the pion decay constant. From energy conservation and the Dirac equation

$$\begin{aligned} k_\mu \langle p, q | \bar{\mu} \gamma^\mu (1 - \gamma^5) \nu | 0 \rangle &= (p + q)_\mu \bar{u}(p) \gamma^\mu (1 - \gamma^5) v(q) \\ &= m_\mu \bar{u}(p) (1 - \gamma^5) v(q) \end{aligned} \quad (15.45)$$

Therefore the transition matrix is written as

$$T_{fi} = (2\pi)^4 \delta^4(k - p - q) \frac{G_\beta}{\sqrt{2}} i f_\pi m_\mu \bar{u}(p)(1 - \gamma^5)v(q) \quad (15.46)$$

The decay rate can be calculated using the γ -matrix algebra given in Appendix C, and becomes

$$\Gamma(\pi \rightarrow \mu \nu) = (G_\beta f_\pi)^2 \frac{m_\mu^2 (m_\pi^2 - m_\mu^2)^2}{8\pi m_\pi^3} \quad (15.47)$$

Since G_β is already given by β decay, the measurement of the decay rate gives

$$f_\pi = 131.74 \pm 0.15 \text{ MeV} \simeq 0.94 m_\pi \quad (15.48)$$

The decay rate $\Gamma(\pi \rightarrow e \nu)$ is obtained simply by replacing $m_\mu \rightarrow m_e$,

$$\therefore \frac{\Gamma(\pi \rightarrow e \nu)}{\Gamma(\pi \rightarrow \mu \nu)} = \frac{m_e^2 (m_\pi^2 - m_e^2)^2}{m_\mu^2 (m_\pi^2 - m_\mu^2)^2} = 1.284 \times 10^{-4} \quad (15.49)$$

Including the radiative corrections, the theoretical value becomes 1.233×10^{-4} [231] and agreement with the experimental value Eq. (15.40) is very good.

Why the decay rate is proportional to the mass squared of the charged lepton can be understood as follows. In the rest frame of the π^+ , the neutrino and the charged lepton fly away in opposite directions (Fig. 15.4). Since the neutrino has zero mass, the factor $(1 - \gamma^5)$ forces the neutrino helicity to be 100% negative. (Note, in π^- decay the antineutrino is produced, which has 100% positive helicity.) Since the pion spin is zero, angular momentum conservation forces the μ^+ to have negative helicity, too. As l^+ is an antiparticle, the negative helicity component is suppressed by a factor m_l/p in the negative chirality state (see arguments in Sect. 4.3.5). Therefore, if $m_l = 0$, the reaction is forbidden completely. It survives owing to its finite mass but with suppression factor $\sim m/p$. The origin of the helicity suppression is $1 - \gamma^5$.

That the lepton in the weak interaction has the factor $(1 - \gamma^5)$ can also be confirmed by the direct observation of the muon helicity in the $\pi \rightarrow \mu \nu$ decay. Experimental data [140] show

$$0.9959 < |h(\mu^+)| < 1.0 \quad (15.50)$$

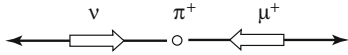


Figure 15.4 The decay rate $\pi^+ \rightarrow \mu^+ + \nu$ is proportional to m_μ^2 . The pion spin is 0, and the neutrino's helicity is negative. Therefore angular momentum conservation forces the muon helicity to be negative, too. But in the

V-A interaction, the component of the positive helicity in the muon wave function is proportional to m_μ/p , which gives the decay rate $\sim m_\mu^2/p^2$.

Exclusion of P-Type Interaction The pion decay rate also severely constrains the existence of the P-type interaction. If the Lagrangian is of P ($\Gamma = i\gamma^5$) type, the decay amplitude becomes

$$\begin{aligned} T_{fi} &\sim G_\beta C_P \langle 0 | \bar{p} i \gamma^5 n | \pi \rangle \bar{u}(p) i \gamma^5 (1 - \gamma^5) v(q) \\ &= G_\beta C_P f'_\pi \bar{u}(p) (1 - \gamma^5) v(q) \end{aligned} \quad (15.51)$$

Note, no lepton mass appears in the expression. This is because the P-type interaction flips the chirality (see next section) and the suppression factor disappears. Therefore, if the P-type dominates, the rate will be determined largely by the phase space volume and the electron decay rate will be larger than the muon decay rate. Quantitatively, the decay rate ratio when the Lagrangian contains an extra P-type interaction is given by

$$\frac{\Gamma(\pi \rightarrow e \nu)}{\Gamma(\pi \rightarrow \mu \nu)} = \left| \frac{1 - (m_e/m_\pi)^2}{1 - (m_\mu/m_\pi)^2} \right|^2 \left| \frac{C_P f'_\pi - (m_e/m_\pi)^2 C_{V-A} f_\pi}{C_P f'_\pi - (m_\mu/m_\pi)^2 C_{V-A} f_\pi} \right|^2 \quad (15.52)$$

Comparing the formula with the experimental value, one can easily understand that the P-type interaction is highly excluded. In order not to contradict the experimental value we need $|C_P|^2/|C_{V-A}|^2 \lesssim 10^{-4}$.

Summarizing the result, one may say that the characteristic of the V-A interaction manifests itself most conspicuously in the pion decay mode. Historically, the determination of the weak interaction type took a long and meandering road. At one time erroneous experiments puzzled people as to which of the solutions V, A or S, T, P should be adopted. Logically, however, it is much clearer to grasp various features if one simply looks at the helicity of the decay leptons. Simply from the fact that the decay lepton has negative helicity [$h(\nu) = -1$, $h(l) = -v_l$] while that of the antilepton is opposite [$h(\bar{\nu}) = +1$, $h(\bar{l}) = +v_l$], one can conclude that the parity is maximally violated and that

$$\begin{aligned} C_S = C_T = C_P = 0, \quad C_A = -C_V^* \\ C'_i = -C_i \quad i = V, A \end{aligned} \quad (15.53)$$

which will be discussed in the next section.

15.3

Chirality of the Leptons

15.3.1

Helicity and Angular Correlation

The significance of Eq. (15.37) cannot be overestimated. If the neutrino has vanishing mass, $(1 - \gamma^5)\nu$ implies a two-component neutrino and that the chirality of

7) The reason for setting $C_A = -C_V$ and not $C_A = -1.26C_V$ will be clarified in the discussion of the CVC hypothesis in Sect. 15.5.2.

the lepton that appears in the weak interaction is always negative. This means the helicity of the emitted e^- , $\bar{\nu}$ must be $-\nu$, $+1$ (for e^+ , ν , the opposite is true). We notice that the V, A-type interaction conserves the chirality, but S, T, P inverts it. To see this we note that the γ matrix has the following properties:

$$\gamma^5 \gamma^5 = 1, \quad \gamma^5 \gamma^0 = -\gamma^0 \gamma^5 \quad (15.54a)$$

$$\gamma^5 \Gamma = \begin{cases} -\Gamma \gamma^5 & \text{V, A} \\ +\Gamma \gamma^5 & \text{S, T, P} \end{cases} \quad (15.54b)$$

Then as $[(1 - \gamma^5)/2]^2 = (1 - \gamma^5)/2$

$$\begin{aligned} \frac{1}{2} \bar{e} \Gamma (1 - \gamma^5) \nu &= e^\dagger \gamma^0 \Gamma \left(\frac{1 - \gamma^5}{2} \right)^2 \nu \\ &= \left(\frac{1 \mp \gamma^5}{2} e \right)^\dagger \gamma^0 \Gamma \left(\frac{1 - \gamma^5}{2} \right) \nu \begin{cases} - & \text{V, A} \\ + & \text{S, T, P} \end{cases} \end{aligned} \quad (15.55)$$

This equation means that the measurement of the helicity of the emitted lepton can distinguish STP from VA. Once one accepts the fact that the neutrino has negative chirality, the fact that the angular correlation coefficient has value $\lambda = 1$ for V and $\lambda = -1/3$ for A [Eq. (15.25d)] can be understood as follows. In the V-type interaction $\Delta J = 0$ and the $e\bar{\nu}$ pair has total angular momentum $J = 0$. Since the helicity of the electron is negative and that of $\bar{\nu}$ is positive, they tend to fly in the same direction to have $J = 0$ (Fig. 15.5). Namely they want to have $\theta = 0$, hence

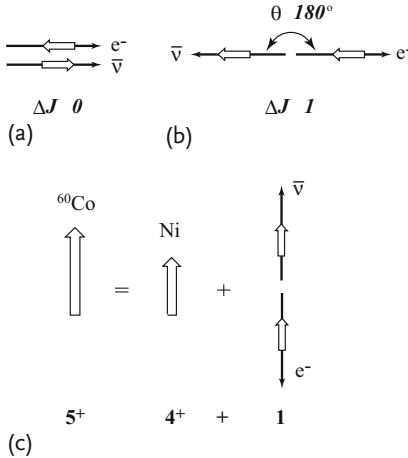


Figure 15.5 Explanation of angular correlation between e and ν . The negative chirality forces the electron to have $h(e^-) = -1$ and $h(\bar{\nu}) = +1$. In the Fermi transition (a), $\Delta J = 0$ and the two leptons have total spin

zero, therefore they align themselves in the same direction, while in the Gamow-Teller transition (b), $\Delta J = 1$ and the opposite is true. The asymmetry in Co decay can be explained in the same way.

$\lambda > 0$. For A-type interaction, $\Delta J = 1$ and, the total spin of $e\nu$ being 1, they want to fly in opposite directions ($\theta = \pi$), hence $\lambda < 0$ (Fig. 15.5). The factor $1/3$ reflects the spin degree of freedom.

In a similar manner, the reason for the asymmetry $A = -1$ in Wu's cobalt experiment can be explained. In the process $\text{Co}(5^+) \rightarrow \text{Ni}(4^+)$ (Fig. 15.5c), $\Delta J = 1$ and the total spin of $e\bar{\nu}$ has to be 1. They have to be emitted in opposite directions and their total spin has to be aligned along the same direction as the parent Co's spin. This means $A < 0$.

15.3.2

Electron Helicity

The helicity of the electron can be measured after converting the longitudinal polarization to transverse polarization. There are several methods. We list a few here.

1. The electrostatic deflector changes the momentum but does not change the spin direction (Fig. 15.6).
2. Applying both the electric and the magnetic field perpendicular to the electron motion in such a way as to keep the electron's path straight, the magnetic field can induce spin rotation according to

$$i \frac{d\sigma}{dt} = [\sigma, H] = [\sigma, -\mu \cdot B] = i|\mu|\sigma \times B \quad (15.56)$$

3. Use Coulomb scattering as described in Sect. 6.7.4.

The transverse polarization can be measured using Mott scattering with nuclei. The electrostatic interaction produces symmetric scattering but the magnetic (spin-orbit coupling $H = -gs \cdot l$) interaction produces left-right asymmetry in the scattering. The magnetic field is induced by the motion of the electron and acts on

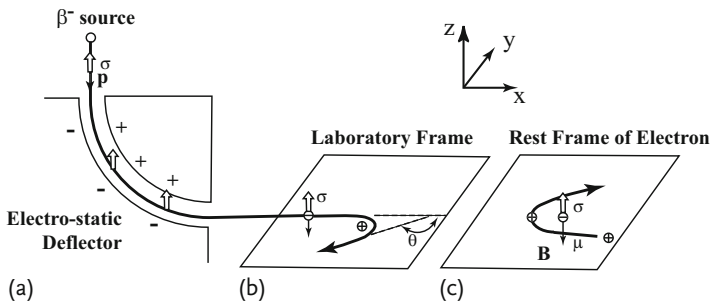


Figure 15.6 Measurement of the electron helicity. The helicity is determined by first converting the longitudinal polarization to transverse polarization and then measuring the left-right asymmetry of the Mott scattering.

Here, an electrostatic deflector (a) is used. The magnetic interaction of the spin induces asymmetry in the scattering and its magnitude can be calculated using QED.

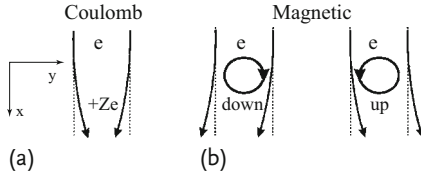


Figure 15.7 (a) Scattering by the Coulomb force is symmetric. (b) Scattering by the spin–orbit coupling force is asymmetric.

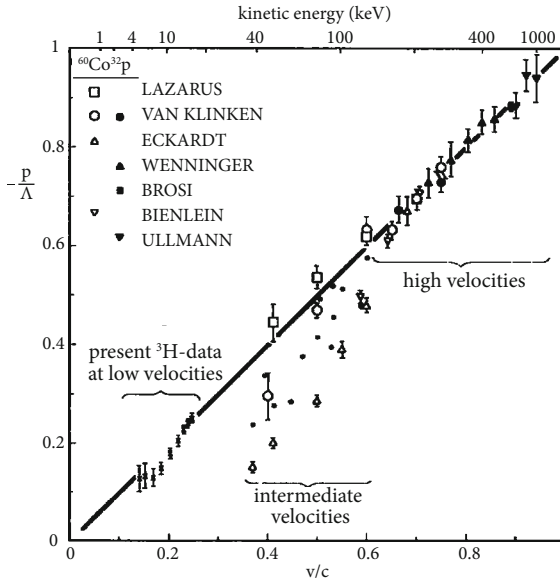


Figure 15.8 A compilation of measured electron helicity from allowed β decays. Λ is a small correction for Coulomb and screening effects [24].

the spin magnetic moment of the electron, because, in the electron rest frame, the target nucleus produces an electric current.

The reason for the asymmetry production can be understood classically as follows. As is shown in Fig. 15.7b, let us define the x -axis in the direction of the electron motion. Assume the electron is polarized in the z direction. When the electron is on a trajectory with $\gamma > 0$, $\mathbf{l} = \mathbf{r} \times \mathbf{p}$ is along the $-z$ direction, whereas when the electron is on a trajectory with $\gamma < 0$, $\mathbf{l} = \mathbf{r} \times \mathbf{p}$ is along the $+z$ direction, and the Hamiltonian changes its sign. In other words, the force (produced by a nonuniform magnetic field) is either attractive or repulsive depending on $\gamma > 0$ or $\gamma < 0$, and always bends the electron in the same direction, producing the asymmetry.

Measurements showed the electron's helicity to be $h(e^-) = -v$ and that of the positron $h(e^+) = +v$ and confirmed the negative chirality of the decay electron as shown in Fig. 15.8 [147, 280].

15.4

The Neutrino

15.4.1

Detection of the Neutrino

The neutrino is the only particle that does not interact strongly or electromagnetically. The weak interaction, literally, couples weakly to matter, meaning not only that its strength is weak but also that it occurs very seldom. It took nearly 30 years after Pauli's prediction before it was actually detected by Reines and Cowan in 1956 [105, 330]. Today the neutrino is mass produced in reactors as well as in accelerators and even a neutrino telescope has been set up to watch supernovae. However, Reines and Cowan's method of detecting the neutrino is still valid and is being used in a forefront experiment to confirm neutrino oscillation. Therefore, it is worthwhile describing it here. The scheme of their detector is shown in Fig. 15.9. The process used is an inverse β decay and the produced secondary particles were detected using the following reactions:

$$\bar{\nu}_e + p(u) \rightarrow (e^+ + W^-) + p(u) \rightarrow e^+ + n(d) \quad (15.57a)$$

$$\begin{cases} e^+ + e^- \rightarrow \gamma + \gamma \\ n + {}^{108}\text{Cd} \rightarrow {}^{109}\text{Cd}^* \rightarrow {}^{109}\text{Cd} + \gamma \end{cases} \quad (15.57b)$$

where we have denoted participating quarks in parentheses. Neutrinos produced in a reactor were used for the detection. They are actually antineutrinos produced in the β decays of fission products. Because of lepton number conservation, they can only be converted to positrons by (virtually) emitting W^- , and the target proton (actually a u quark in the proton) is converted to a neutron (d quark) by absorbing W^- . As the positron is an antielectron, it is annihilated immediately by an electron in matter, producing two gamma rays with total energy $2m_e$. The recoil neutron is

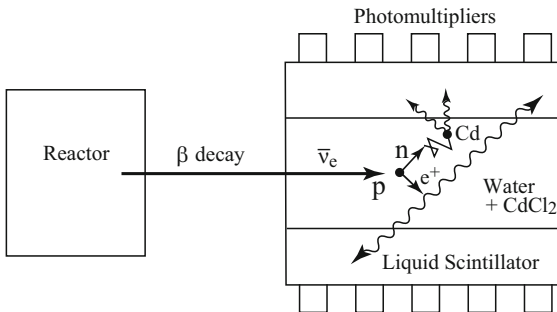


Figure 15.9 Layout of the apparatus used to detect the neutrino in 1956. The antineutrino produced in β decays of nuclear fission products reacts with a proton and emits a positron and a neutron. The positron immediately an-

ihilates an electron in matter, emitting two γ rays. The slow neutron loiters for a while, eventually being captured by a Cd nucleus to emit another γ [105, 330].

nonrelativistic and is scattered many times by matter before it is captured by a Cd nucleus, thereby emitting another gamma ray. A time as long as 10^{-4} s will have passed and the neutron “totters” ~ 1 m. The gamma rays kick electrons (Compton scattering) in matter and the recoil electrons, in turn, emit bremsstrahlung photons, creating pairs. These electrons are detected by scintillation counters. The signal of the neutrino reaction is thus the occurrence of a prompt signal produced by the positron followed by a delayed signal produced by neutron capture. The experiment was carried out at the Savannah River Plant in South Carolina, USA.

The detector consisted of a water tank sandwiched by liquid scintillators, which were viewed by photomultipliers. The tank contained about 200 liters of water with 45 kg of cadmium chloride (CdCl_2) dissolved in it. It was calculated that about 2×10^{14} neutrinos were emitted per 1 kW of the reactor power and the flux injected into the detector was $5 \times 10^{13} \text{ s}^{-1} \text{ cm}^{-2}$. The intensity was 100 times that of the solar neutrino, but the reaction rate was extremely small and suffered severely from false signals created by cosmic rays. To date, the cosmic ray background is a serious problem to be considered in designing an experiment. This was especially true for the first experiment. Eventually, taking the difference between measurements with the reactor in operation and not in operation, the experimenters confirmed an event rate of one per 20 min.

15.4.2

Mass of the Neutrino

The neutrino mass, i.e. whether it vanishes or not, has been a central issue in clarifying the properties of the neutrino. It still is, because a nonvanishing neutrino mass is a sign that the Standard Model is incomplete. There is evidence that it is finite because neutrino oscillation, in which a neutrino changes its flavor during flight, has been observed (see Sect. 19.2.1). However, the oscillation is sensitive only to the mass difference and determination of its absolute value is a forefront research goal as it offers a clue to new physics.

A most promising reaction that may be able to determine the neutrino mass directly is the nuclear β decay with a small Q value. The energy spectrum of the decay electron is obtained by integrating Eq. (15.18) over angles:

$$d\Gamma = \frac{G_\beta^2}{2\pi^3} |\mathcal{M}|^2 F(Z, E_e) p_e E_e p_\nu E_\nu dE_e \quad (15.58)$$

where we have used a more general expression $|\mathcal{M}|^2 F(Z, E)$, instead of $|\langle 1 \rangle|^2$, that includes the Coulomb correction factor $F(Z, E)$ and recovered the correct neutrino variables to discuss the mass of the neutrino. We have already shown that $|\mathcal{M}|^2$ does not contain the energy variable in an allowed transition. If $m_\nu \neq 0$

$$\begin{aligned} p_\nu E_\nu &= (E_0 - E_e) \sqrt{(E_0 - E_e)^2 - m_\nu^2}, \\ E_0 &= M(Z, A) - M(Z + 1, A) - m_\nu \end{aligned} \quad (15.59)$$

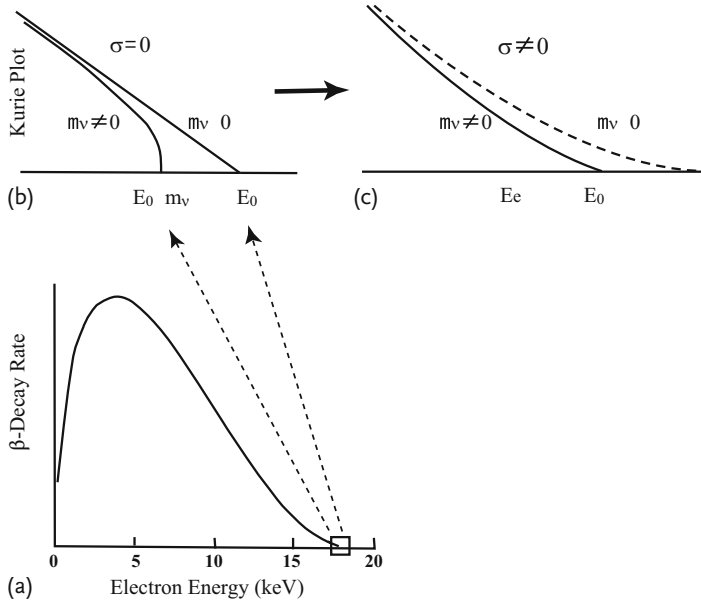


Figure 15.10 (a) Spectrum of the electron in β decays. (b) Theoretical Kurie plot near end point. (c) Kurie plot with smearing. Smearing is due to experimental resolution and contributions of final excited states.

We define a spectral function referred to as a Kurie plot by

$$K(E) = \left[\frac{d\Gamma/dE_e}{F(Z, E_e)p_e E_e} \right]^{1/2} \quad (15.60)$$

If $m_\nu = 0$, $K(E)$ is proportional to $E_0 - E_e$ and describes a straight line.

Figure 15.10a depicts a whole electron spectrum from β decays. For the m_ν determination, only electrons near the end point are important. If $m_\nu \neq 0$ the Kurie plot deviates from a straight line and is vertical at the end point (Fig. 15.10b), but after convolution with the experimental resolution (σ), the curve becomes blurred (Fig. 15.10c). As m_ν can be determined more accurately with smaller value of E_0 , the tritium isotope (T) with $E_0 = 18.6$ keV is usually chosen as the β source. Many experiments have been carried out, so far, in vain. The result has always been consistent with zero mass, albeit with progressively improved upper limit.

Troitzk–Mainz Experiment The previous arguments are valid for early experiments, in which the mass sensitivity was no more than 10–100 eV. Modern neutrino experiments have resolutions close to or better than a few eV, so that atomic or molecular binding energy (a few eV) of the final states have to be considered (see [308] for a review), because the β ray source used for the experiment is not a tritium nucleus but is prepared as a neutral tritium atom or molecule T_2 . In the following experiments the molecular gas of tritium T_2 , whose level structure is well understood, was used.

Taking into account various final states of daughter products $[(^3\text{He} - \text{T})^+ + e]$, the β ray energy spectrum is modified to

$$\begin{aligned} d\Gamma &\sim p_e E_e p_\nu E_\nu = p_e E_e (E_0 - E_e) \sqrt{(E_0 - E_e)^2 - m_\nu^2} \\ &\rightarrow p_e E_e \sum_i P_i (E_0 - V_i - E_e) \sqrt{(E_0 - V_i - E_e)^2 - m_\nu^2} \end{aligned} \quad (15.61)$$

where V_i is the excited energy of a final state and P_i is its probability. The recoil energy of the daughter is included in E_0 . Note that the above formula tells us that the experimental observable is m_ν^2 , not m_ν . The spectrum shape is shown in Fig. 15.10c, which includes smearing due to the excited energy of the daughters.

In the early measurements, a magnetic spectrometer was used to determine the momentum of the decay electrons, but it only measures the transverse component of the momentum relative to the magnetic field lines and smearing due to scattering cannot be avoided, thus making it hard to resolve the contributions of the final states. Modern detectors use an electrostatic spectrometer, which measures the total energy of the electron. It filters and passes only a small portion of the high-end spectrum, which is the region most sensitive to the mass value.

Here we describe the basic principles of the electrostatic filter of the Troitzk [54] and Mainz experiment [243, 379], which uses MAC-E (Electrostatic filter with Magnetic Adiabatic Collimator). A schematic layout is shown in Fig. 15.11. Two superconducting solenoids provide the guiding magnetic field. Beta particles, which are emitted from the tritium source in the strong field B_S inside the left solenoid are guided magnetically into the spectrometer along the magnetic field line and detected by a detector, which is located at the other end of the strong field B_D . HV (high voltage) electrodes generate a retarding (accelerating) graded electric field in the left (right) half of the spectrometer and act as a filter to pass only high-energy β particles determined by the voltage U_0 .

From the source, β particles are emitted isotopically. When a particle is emitted at angle θ , its transverse momentum relative to the magnetic field line causes it to describe a circular motion and proceeds forward in a spiral along the field line. The angular frequency of the circular motion is given by the cyclotron frequency $\omega_c = eB/\gamma m$, where m is the electron mass and γ the relativistic Lorentz factor. The orbital angular momentum is a constant of motion for a constant magnetic field. The magnetic field of the spectrometer is not uniform but strong at both ends and weakest ($B = B_{\min}$) in the central plane of the spectrometer. However, as long as the adiabatic condition

$$\frac{1}{B} \frac{dB}{dt} \ll \omega_c \quad (15.62)$$

is satisfied, the orbital angular momentum is still a constant of the motion. The orbital angular momentum L (times $e/2m$) is a product of the magnetic moment μ created by the circular motion and the relativistic γ factor and can be expressed as

$$\frac{e}{2m} L = \gamma \mu = \gamma I S = \gamma \frac{e \omega_c}{2\pi} \pi \rho^2 = \frac{p_T^2}{2m B} \equiv \frac{E_T}{B} \quad (15.63)$$

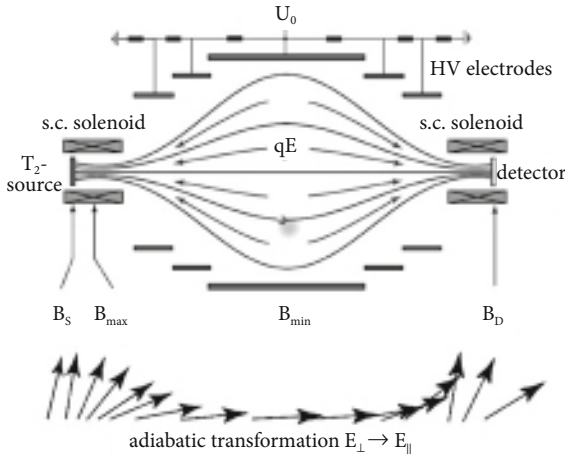


Figure 15.11 Principle of the MAC-E filter: Two superconducting solenoids produce a magnetic guiding field. The tritium source is placed in the left solenoid, while the detector sits in the right solenoid. Between the two solenoids there is an electrostatic retardation system consisting of several cylindrical electrodes. In the central plane perpendicular to

the field lines, the minimum of the magnetic field $B_{\min}$ coincides with the maximum of the electric retarding potential U_0 . The vectors in the lower part of the picture illustrate the aligning of the momentum vector along B in the region of the low magnetic field by the adiabatic transformation (plotted without electrostatic retardation) [308].

where I is the circular current, S the area of the circle, ρ its radius and E_T the transverse energy of the electron. In the nonrelativistic approximation, which is valid for the tritium β decay,

$$\mu = \frac{E_T}{B} = \text{constant}^8 \quad (15.64)$$

Equation (15.64) means the transverse energy diminishes in proportion to the strength of the magnetic field. Without the electric field, the total energy $E = E_{\parallel} + E_T$ is a constant of the motion, namely, the spectrometer transfers the transverse energy to the longitudinal energy, as is illustrated at the bottom of Fig. 15.11. At the central plane (analyzing plane), $B_{\min}/B_S = 3 \times 10^{-4}$ [Tesla]/6 [Tesla] and all the β particles are almost parallel to the axis of the two solenoidal coils. The electric field produces the retarding field and pushes all the particles back except those with energy higher than the maximum potential $e U_0$. The right half of the spectrometer, on the other hand, accelerates the particles that have survived the electric filter and delivers them to the detector.

8) This also means magnetic flux conservation in a circle, because

$$\Phi = \pi \rho^2 B = \pi \left(\frac{p_T}{e B} \right)^2 B = \frac{2\pi m}{e^2} \frac{E_T}{B} \quad (15.65)$$

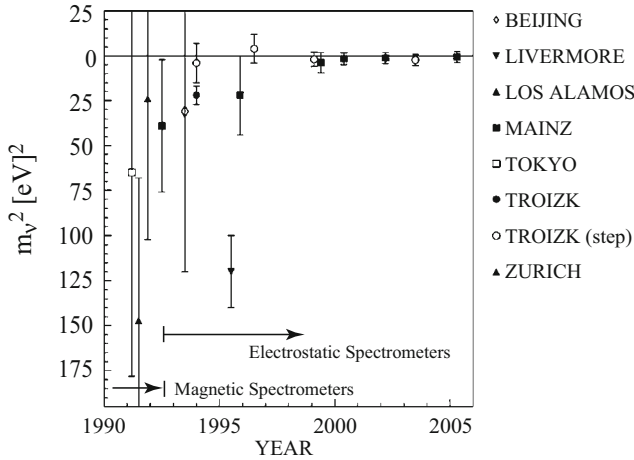


Figure 15.12 Squared neutrino mass values obtained from tritium β decay in the period 1990–2005 plotted against the year of publication [308].

Practically all the β particles that are directed toward the detector, except those with a tiny fraction ($B_{\min}/B_s$) of E_T , are accepted. Therefore the geometrical acceptance of the source is nearly 2π .

Figure 15.12 shows the history of neutrino mass measurements. The current best value is [311]

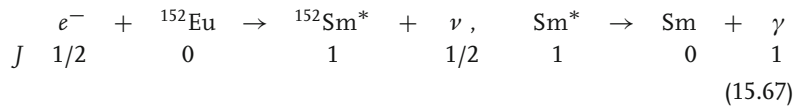
$$m_\nu = 1.1 \pm 2.4 \text{ eV} \quad (15.66)$$

Since from the oscillation data $m \lesssim 0.05 \text{ eV}$ is inferred (see Sect. 19.2.1), the absolute mass measurement should be aimed at least at this value when future experiments are being planned.

15.4.3

Helicity of the Electron Neutrino

The helicity of the neutrino cannot be measured directly, but it can be determined by observing the recoil nucleus as it emits a neutrino following electron capture [179]. The reaction used is described in Fig. 15.13a.



In the first reaction, Eu at rest absorbs an electron in an orbital angular momentum S state (K-capture), therefore the spin of the recoiled Sm^* on emitting a neutrino is aligned in the same direction as the electron spin and opposite to the neutrino spin. If the momentum of the emitted γ is along the spin direction of the



half-life of $^{152}\text{Sm}^*$ is $\sim 7 \times 10^{-14}$ s, with resultant broadening of the level, thus allowing the γ energy to reach the resonance energy. The helicity of the γ can be measured using the different absorption in passing through magnetized iron. (b) The apparatus contains ^{152}Eu in the magnetized iron. Emitted γ rays are scattered by Sm_2O_3 below and detected by a NaI scintillation counter [179].

parent Sm^* , the γ preserves the parent's helicity. Therefore by measuring the γ helicity, the helicity of the neutrino can be determined. The trick of this ingenious experiment is to pick out exclusively the forward-going γ by using resonance scattering. Resonance scattering occurs only when the incident γ has the right energy

to raise Sm to the resonant level 963 keV above the ground state



This means the energy of the incoming γ ray has to be slightly larger than 963 keV to give the recoil energy to the target. This extra energy is supplied by moving Sm^* , which emits the gamma ray in the forward direction. The helicity of the γ can be determined by passing it through magnetized iron as the absorption rate depends on the value of the helicity. When the electron spin in the iron atom is antiparallel to the γ spin, it can absorb the γ by flipping its spin, but when the spin is parallel it cannot flip its spin, hence there is less absorption. The experiment has demonstrated the negative helicity of the neutrino and confirmed the weak interaction type as represented by Eq. (15.37).

15.4.4

The Second Neutrino ν_μ

The decay electron from the muon has a continuous energy spectrum, indicating two neutral particles (neutrinos) are emitted.



The initial question was whether they are the same two neutrinos ($\bar{\nu}_e, \nu_e$) as appear in nuclear β decays or two different neutrinos ($\bar{\nu}_e, \nu_\mu$).

If ν_μ is identical to ν_e the muon should be able to decay to $e + \gamma$ through the process depicted in Fig. 15.14, because the second neutrino (ν_μ) couples to the muon and the first to the electron. The calculated decay rate was within experimental reach but the process has not been observed to this day [311]

$$\text{BR}(\mu \rightarrow e + \gamma) < 1.2 \times 10^{-11} \quad (15.70)$$

If $\nu_\mu \neq \nu_e$, the process in Fig. 15.14 is absent and its nonexistence can be explained. The hypothesis can be tested by making decay neutrinos from the pion

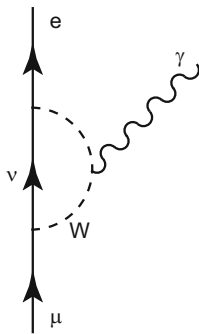


Figure 15.14 $\mu \rightarrow e + \gamma$ can occur through this process if the ν_μ is identical or mixes with ν_e , giving a much higher decay rate than is observed.

react with matter. If $\nu_\mu \neq \nu_e$, electrons will not be produced by $\nu_\mu N$ reactions.

$$\begin{aligned} \nu_\mu + n &\rightarrow \mu^- + p & \nu_\mu + n &\not\rightarrow e^- + p \\ \bar{\nu}_\mu + p &\rightarrow \mu^+ + n & \bar{\nu}_\mu + p &\not\rightarrow e^+ + n \end{aligned} \quad (15.72a)$$

$$\begin{aligned} \nu_e + n &\rightarrow e^- + p & \nu_e + n &\not\rightarrow \mu^- + p \\ \bar{\nu}_e + p &\rightarrow e^+ + n & \bar{\nu}_e + p &\not\rightarrow \mu^+ + n \end{aligned} \quad (15.72b)$$

An experiment carried out at BNL in 1962 confirmed the hypothesis [112]. Only muons are produced and no electrons were observed in the reaction using neutrinos from π decay. This was the first evidence that there exists a new quantum number (one might call it “e-ness” and “mu-ness” to distinguish the two kinds of leptons) that is conserved in the leptonic weak interaction. Later, the tau lepton τ and its associated neutrino were discovered in the $e^- + e^+$ reaction and “tau-ness” was added. Today, they are generically called lepton flavor.

15.5

The Universal V–A Interaction

15.5.1

Muon Decay

The muon was discovered in 1937 by Anderson and Neddermeyer [293] and was mistaken as the meson predicted by Yukawa because it had about the right mass ($m_\mu \simeq 200m_e$). The confusion lasted for ten years until finally the pion was discovered in 1947 by Powell [80] (see Fig. 13.5). But after the pion had been identified as the right Yukawa meson, the muon remained a mystery. Except for its mass, the muon had identical properties to the electron and may well be called a “heavy electron”. Except for the fact that it can decay into an electron and neutrinos, all the reaction formula for the muon may be reproduced from those of the electron simply by replacing its mass. Nobody could explain why it should exist. The perplexity was exemplified by Rabi’s words: “Who ordered it?” Although its existence itself was a mystery, its properties have provided essential clues in clarifying the characteristics of the weak interaction. Precision experiments on the muon properties provide a sensitive test of the Standard Model and remain a hot topic to this day.

Decay Rate The muon decays practically 100% of the time to an electron and two neutrinos. Here we investigate its decay rate and see what it tells us. Let us define the kinematical variables as

$$\mu^-(p_1) \rightarrow e^-(p_4) + \bar{\nu}_e(p_2) + \nu_\mu(p_3) \quad (15.73)$$

The interaction Lagrangian may be written as

$$\mathcal{L}_{\text{WI}} = -\frac{G_\mu}{\sqrt{2}} [\bar{\nu}_\mu \gamma^\mu (1 - \gamma^5) \mu] [\bar{e} \gamma_\mu (1 - \gamma^5) \nu_e] \quad (15.74)$$

Here, the Fermi coupling constant is denoted as G_μ for the sake of later discussions. The transition amplitude is given by

$$\begin{aligned} T_{fi} &= (2\pi)^4 \delta^4(p_1 - p_2 - p_3 - p_4) \\ &\times \frac{G_\mu}{\sqrt{2}} [\bar{u}(p_3) \gamma^\mu (1 - \gamma^5) u(p_1)] [\bar{u}(p_4) \gamma_\mu (1 - \gamma^5) v(p_2)] \end{aligned} \quad (15.75)$$

If the muon is not polarized, taking the spin average of the initial and the sum of all the final state spins, we can express the decay rate as

$$d\Gamma = \frac{1}{2m_\mu} |\overline{\mathcal{M}}|^2 dLIPS \quad (15.76a)$$

$$|\overline{\mathcal{M}}|^2 = \left(\frac{G_\mu}{\sqrt{2}} \right)^2 \frac{16}{2} K_L^{\mu\nu}(p_1, p_3) K_{L\mu\nu}(p_2, p_4) \quad (15.76b)$$

$$dLIPS = (2\pi)^4 \delta^4(p_1 - p_2 - p_3 - p_4) \frac{d^3 p_2}{(2\pi)^3 2E_2} \frac{d^3 p_3}{(2\pi)^3 2E_3} \frac{d^3 p_4}{(2\pi)^3 2E_4} \quad (15.76c)$$

where

$$\begin{aligned} K_L^{\mu\nu}(p_1, p_3) &= (1/4) \sum_{\text{spin}} [\bar{u}(p_3) \gamma^\mu (1 - \gamma^5) u(p_1)] \\ &\times [\bar{u}(p_3) \gamma^\nu (1 - \gamma^5) u(p_1)]^* \\ &= 2 [\{p_1^\mu p_3^\nu + p_1^\nu p_3^\mu - g^{\mu\nu} (p_1 \cdot p_3)\} + i\epsilon^{\mu\nu\rho\sigma} p_{1\rho} p_{3\sigma}] \\ &\equiv P^{\mu\nu}(p_1, p_3) + Q^{\mu\nu}(p_1, p_3) \end{aligned} \quad (15.77)$$

[see Appendix Eq. (C.10d)]. Using

$$\begin{aligned} P^{\mu\nu}(p_1, p_3) P_{\mu\nu}(p_2, p_4) &= 8 \{(p_1 \cdot p_2)(p_3 \cdot p_4) + (p_1 \cdot p_4)(p_2 \cdot p_3)\} \\ Q^{\mu\nu}(p_1, p_3) Q_{\mu\nu}(p_2, p_4) &= 8 \{(p_1 \cdot p_2)(p_3 \cdot p_4) - (p_1 \cdot p_4)(p_2 \cdot p_3)\} \\ P^{\mu\nu} Q_{\mu\nu} &= Q^{\mu\nu} P_{\mu\nu} = 0 \end{aligned} \quad (15.78)$$

we obtain

$$K_L^{\mu\nu}(p_1, p_3) K_{L\mu\nu}(p_2, p_4) = 16(p_1 \cdot p_2)(p_3 \cdot p_4) \quad (15.79)$$

If Eq. (15.79) is inserted in Eqs. (15.76), the expression for the decay rate becomes

$$|\overline{\mathcal{M}}|^2 = 64 G_\mu^2 (p_1 \cdot p_2)(p_3 \cdot p_4) \quad (15.80a)$$

$$d\Gamma = \frac{64 G_\mu^2 (p_1 \cdot p_2)(p_3 \cdot p_4)}{2m_\mu (2\pi)^5} \frac{d^3 p_2 d^3 p_3 d^3 p_4}{8E_2 E_3 E_4} \delta^4(p_1 - p_2 - p_3 - p_4) \quad (15.80b)$$

Problem 15.3

When $m_2 = m_3 = 0$, show

$$S^{\mu\nu} \equiv \int p_2^\mu p_3^\nu \delta^4(P^\mu - p_2 - p_3) \frac{d^3 p_2}{2E_2} \frac{d^3 p_3}{2E_3} = \frac{\pi}{24} (g^{\mu\nu} P^2 + 2P^\mu P^\nu) \quad (15.81)$$

Hint: From Lorentz invariance the equation can be expressed as $A g^{\mu\nu} + B P^\mu P^\nu$. Calculate A and B in the center of mass frame.

Using Eq. (15.81), we have

$$d\Gamma = \frac{G_\mu^2}{2\pi^5} (p_{1\mu} S^{\mu\nu} p_{4\nu}) \frac{d^3 p_4}{m_\mu E_4} \quad (15.82)$$

If we neglect the electron mass, energy takes its minimum value 0 in the muon rest frame, and maximum value $m_\mu/2$ when the momenta of the two neutrinos are both in the opposite direction to that of the electron. Putting $p_1 = (m_\mu, 0, 0, 0)$ and setting $m_e = 0$ we obtain the electron energy spectrum

$$\frac{d\Gamma}{dx} = \frac{G_\mu^2 m_\mu^5}{96\pi^3} x^2 (3 - 2x), \quad x = \frac{E_4}{m_\mu/2} \quad 0 \leq x \leq 1 \quad (15.83)$$

This reproduces the experimental data (Fig. 15.15) quite well.

Integrating over the energy, we obtain the total decay rate

$$\Gamma = \frac{G_\mu^2 m_\mu^5}{192\pi^3} \quad (15.84)$$

Problem 15.4

Derive Eq. (15.83).

Problem 15.5

Show for $m_e > 0$ Eq. (15.84) is modified to

$$\Gamma = \frac{G_\mu^2 m_\mu^5}{192\pi^3} f\left(\frac{m_e^2}{m_\mu^2}\right), \quad f(y) = 1 - 8y + 8y^3 - y^4 - 12y^2 \ln y \quad (15.85)$$

The precise expression including the radiative correction is given by [53, 333]

$$\Gamma = \frac{G_\mu^2 m_\mu^5}{192\pi^3} f\left(\frac{m_e^2}{m_\mu^2}\right) \left[1 + \frac{\alpha}{2\pi} \left(1 + \frac{2\alpha}{3\pi} \ln \frac{m_\mu}{m_e} \right) \left(\frac{25}{4} - \pi^2 \right) \right] \quad (15.86)$$

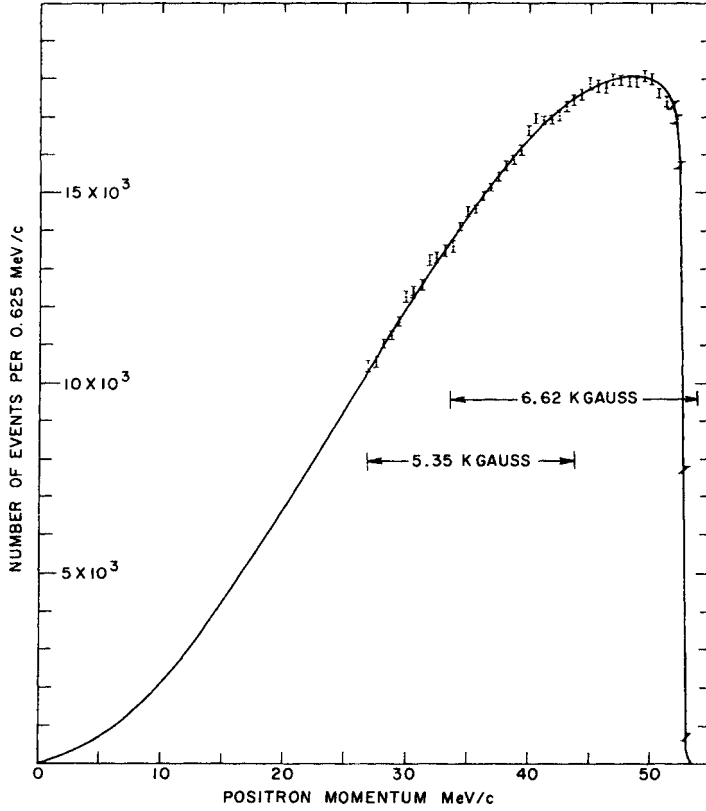


Figure 15.15 Energy spectrum of the decay electrons in $\mu \rightarrow e \nu \nu$. The line is a theoretical curve including the radiative correction with the Michel parameter $\rho = 3/4$ [44].

Inserting the observed muon lifetime [311]

$$\tau = \frac{1}{\Gamma} = 2.197019 \pm 0.000021 \times 10^{-8} \text{ s} \quad (15.87)$$

we find the coupling constant is determined as

$$G_\mu \simeq 10^{-5} \text{ GeV}^{-2} \quad (15.88)$$

which is almost identical with G_β [Eq. (15.29)] obtained from nuclear β decays. This was the first indication that the coupling constant of the weak interaction was universal. For later discussion we list the two coupling constants and compare them:

$$G_\mu = 1.16637 \pm 0.00001 \times 10^{-5} \text{ GeV}^{-2} \quad (15.89a)$$

$$G_\beta = 1.14730 \pm 0.00064 \times 10^{-5} \text{ GeV}^{-2} \quad (15.89b)$$

The small but significant difference between the two coupling constants will be discussed later.

Michel Parameters In calculating the muon decay rate, we assumed V–A interaction. Historically, however, the most general expression [Eq. (15.23)] was used to determine the interaction type. In the electron energy spectrum from polarized muons, angular correlations with the polarization axis vary according to the interaction type. Its general form is expressed in terms of Michel parameters [115, 141, 274]. The usefulness of the parameter lies in that it can be applied not only to the muon but also to the tau or any new lepton and determine its interaction type. Development of surface muons⁹⁾ has enabled precise experiments using polarized muons, providing a very effective test bench to check the validity of the Standard Model and beyond. The distribution of the decay electron from the polarized muon $\mu^\mp$ can be expressed in the approximation neglecting the electron mass as

$$\frac{1}{\Gamma} \frac{d\Gamma}{dx d\cos\theta} = 4x^2 \left[3(1-x) + \frac{2}{3}\rho(4x-3) \mp P_\mu \xi \cos\theta \right. \\ \left. \times \left\{ 1-x + \frac{2}{3}\delta(4x-3) \right\} \right] \quad (15.90)$$

where P_μ is the muon polarization, $x = 2E_e/m_\mu$ and θ is the angle between the muon spin and the electron momentum. Note that the term with the polarization is the parity-violating term. ρ, ξ, δ are the Michel parameters [141]. In order not to make the equation too complicated, we restrict the interaction type to V and A in the following. The general formula can be found in [311]. The Lagrangian can be written in the form

$$\mathcal{L}_{\text{WI}} = -\frac{G_\mu}{\sqrt{2}} \left[\bar{e}\gamma_\mu(1-\gamma^5)\nu_e \{ \bar{\nu}_\mu\gamma^\mu [g_{\text{LL}}(1-\gamma^5) + g_{\text{LR}}(1+\gamma^5)]\mu \} \right. \\ \left. + \bar{e}\gamma_\mu(1+\gamma^5)\nu_e \{ \bar{\nu}_\mu\gamma^\mu [g_{\text{RL}}(1-\gamma^5) + g_{\text{RR}}(1+\gamma^5)]\mu \} \right] + h.c. \quad (15.91)$$

The Michel parameters are expressed as

$$\rho = \frac{3}{4} \frac{|g_{\text{LL}}|^2 + |g_{\text{RR}}|^2}{D}, \quad \delta = \frac{3}{4} \frac{|g_{\text{LL}}|^2 - |g_{\text{RR}}|^2}{D}, \quad \xi = 1 \\ D = |g_{\text{LL}}|^2 + |g_{\text{LR}}|^2 + |g_{\text{RL}}|^2 + |g_{\text{RR}}|^2 \quad (15.92)$$

In the V–A or the Standard Model, $g_{\text{LL}} = 1, g_{\text{LR}} = g_{\text{RL}} = g_{\text{RR}} = 0$ and gives

$$\rho = \delta = \frac{3}{4}, \quad \xi = 1 \quad (15.93)$$

If we adopt V–A as proven for the electron, we can set $g_{\text{RL}} = g_{\text{RR}} = 0$ and test only the muon part, $g_{\text{LL}} = 0, g_{\text{LR}} = 1$ for the V+A type and $g_{\text{LL}} = g_{\text{LR}} (g_{\text{LL}} = -g_{\text{LR}})$ for

9) If matter is irradiated with pions, slow pions stopping in the matter, muons produced in $\pi \rightarrow \mu\nu$ have some definite energy. Especially, those from the surface of matter have a clean monochromatic spectrum.

pure V(A) type, giving $\rho = 0$ (V-A), $3/8$ (V), $3/8$ (A). The data show [311]

$$\rho = 0.7509 \pm 0.0010 \quad (15.94a)$$

$$\delta = 0.7495 \pm 0.0012 \quad (15.94b)$$

$$|\xi P_\mu| = 1.0007 \pm 0.0035 \quad (15.94c)$$

$|P_\mu|$ is the polarization of the muon and should be 100% for muons obtained from $\pi \rightarrow \mu \nu$ in the V-A interaction. An experimental analysis using the surface muon beam gives [40]

$$\begin{aligned} |g_{LL}| &> 0.888, & |g_{LR}| &< 0.060, & |g_{RL}| &< 0.110, \\ |g_{RR}| &< 0.033 & 90\%CL \end{aligned} \quad (15.95)$$

In summary, the data confirms the V-A-type Lagrangian with high accuracy.

15.5.2

CVC Hypothesis

Isospin Current The form of the interaction determined by the nuclear and muon decay can be combined to give

$$\begin{aligned} \mathcal{L}_{WI} &= \frac{G_\beta}{\sqrt{2}} [\bar{p} \gamma_\mu (1 - 1.26 \gamma^5) n] [\bar{e} \gamma^\mu (1 - \gamma^5) \nu_e + \bar{\mu} \gamma^\mu (1 - \gamma^5) \nu_\mu] \\ &+ \frac{G_\mu}{\sqrt{2}} [\bar{\nu}_\mu \gamma_\mu (1 - \gamma^5) \mu] [\bar{e} \gamma^\mu (1 - \gamma^5) \nu_e] + h.c. \end{aligned} \quad (15.96)$$

As far as the vector part of the weak current is concerned, the coupling constant of the muon decay G_μ and that of nuclear β decay (G_β) agree within 2%. $n \rightarrow p$ transformation is really $d \rightarrow u$ transformation and is influenced by quark wave functions to form a proton bound state. Namely, the coupling here includes the strong interaction correction. Based on near identity of the strength, it was suggested that the weak vector current is a conserved current, namely

$$\begin{aligned} \partial_\mu j_W^{+\mu} &\equiv \partial_\mu \bar{p} \gamma^\mu n = \partial_\mu \bar{\psi} \gamma^\mu \tau^+ \psi = 0 \\ \psi &= \begin{bmatrix} p \\ n \end{bmatrix}, \quad \tau^+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \end{aligned} \quad (15.97)$$

The inference is based on the analogy of the identical strength of the electric charge for the electron and the proton. The proton, interacting strongly, couples to the electromagnetic current not only as a single proton but also through pion clouds as depicted in Fig. 15.16a, and yet the strength of the charge is identical to that of the electron, which does not have such a complication. The identity is guaranteed by current conservation:

$$\partial_\mu j_{EM}^\mu = \partial_\mu \bar{p} \gamma^\mu p = \partial_\mu \bar{\psi} \gamma^\mu \left(\frac{1 + \tau_3}{2} \right) \psi = 0 \quad (15.98a)$$

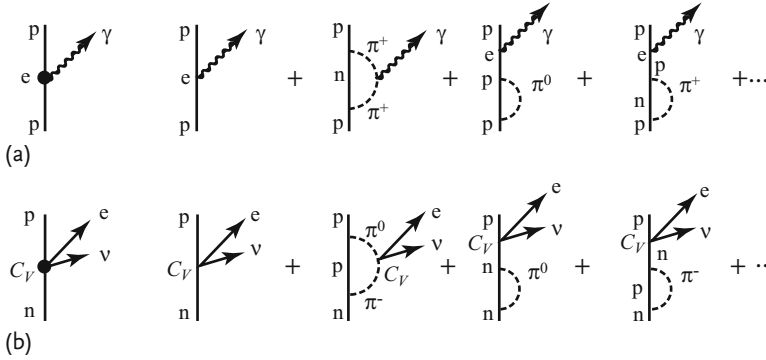


Figure 15.16 Diagrams to show the universality of the coupling constant. The strength of the electromagnetic current is invariant under the strong interaction effect as depicted in (a) because of the current conservation. Similarly,

the CVC hypothesis makes the weak charge invariant under the influence of strong interaction corrections. If $(e^- \nu)$ is replaced by W^- the correspondence is more apparent.

Here, in moving from the second equality to the third, we used

$$p = \frac{1 + \tau_3}{2} \psi, \quad \psi = \begin{bmatrix} p \\ n \end{bmatrix}, \quad \tau_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (15.98b)$$

Equations (15.98) have the form of Nishijima–Gell-Mann’s rule in a differential expression and $(1 + \tau_3)/2$ corresponds to $Y/2 + I_3$. Comparing (15.97) and (15.98), we see the τ_3 part of the electromagnetic current and the weak vector current ($\propto \tau^\pm$) have the same form. This suggests they are just different components of the same isospin current

$$j_I^\mu = \bar{\psi} \gamma^\mu \left(\frac{\boldsymbol{\tau}}{2} \right) \psi, \quad \partial_\mu j_I^\mu = 0 \quad (15.99)$$

This is referred to as the CVC (conserved vector current) hypothesis [167].

Chiral Invariance The coefficient of the axial vector current ($\gamma^5 \gamma^\mu$) of the nucleon differs from that of the lepton by 26%. If this is regarded as the result of the strong interaction correction, it is logical to believe that, at the quark level, the coupling must be equal. This advances to a stage one step beyond where the nuclear β decay $(Z, A) \rightarrow (Z + 1, A)$ was considered as $n \rightarrow p$, and β decay should be considered as a $d \rightarrow u$ process. At this most fundamental level, it is conceivable that the weak interaction is described by

$$\mathcal{L}_{\text{WI}} = -\frac{G_\beta}{\sqrt{2}} [\bar{u} \gamma^\mu (1 - \gamma^5) d] [\bar{e} \gamma_\mu (1 - \gamma^5) \nu_e] \quad (15.100)$$

This is referred to as the V–A interaction. Later Adler [5] and Weisberger [384] used the PCAC (partially conserved axial vector current) hypothesis¹⁰⁾ and succeeded in deriving $C_A/C_V = 1.26$ in terms of π – N scattering cross section, thereby confirming the above conjecture. Equation (15.100) has a simple and beautiful form and agrees with what can be derived theoretically from the chiral invariance [355, 356]. It requires that the equation of motion should be invariant under the chiral transformation

$$\psi \rightarrow \psi' = \exp(-i\gamma^5 \alpha) \psi, \quad \psi^\dagger \rightarrow \psi'^\dagger = \psi^\dagger \exp(i\gamma^5 \alpha) \quad (15.102)$$

Note, it is a gauge transformation but changes the phase of the left- and right-handed components in the opposite direction. The V–A type was also inferred from the two-component neutrino theory [144] and mass inversion requirement [342]. Namely, the V–A interaction has a certain theoretical foundation as a basic starting point of the equation of motion. Since

$$\begin{aligned} \psi &= \left(\frac{1 - \gamma^5}{2} + \frac{1 + \gamma^5}{2} \right) \psi = \psi_L + \psi_R \\ \gamma^5 \psi_L &= -\psi_L, \quad \gamma^5 \psi_R = \psi_R \end{aligned} \quad (15.103)$$

the Lagrangian for the free Dirac particle has the form

$$\begin{aligned} \mathcal{L}_{\text{Dirac}} &= \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \\ &= i\bar{\psi}_L \gamma^\mu \partial_\mu \psi_L + i\bar{\psi}_R \gamma^\mu \partial_\mu \psi_R + m(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L) \end{aligned} \quad (15.104)$$

The second equality can be derived using Eq. (15.103) and anticommutativity of γ^5 with all other γ^μ 's. It is easy to verify that the vector and axial vector current are invariant under chiral transformation, i.e. $\bar{\psi}' \gamma^\mu \psi' = \bar{\psi} \gamma^\mu \psi$, $\bar{\psi}' \gamma^5 \gamma^\mu \psi' = \bar{\psi} \gamma^5 \gamma^\mu \psi$. However, the mass term transforms as

$$\begin{aligned} \bar{\psi}' \psi' &= \psi^\dagger e^{i\alpha\gamma^5} \gamma^0 e^{-i\alpha\gamma^5} \psi \\ &= \bar{\psi} e^{-2i\alpha\gamma^5} \psi = \cos 2\alpha \bar{\psi} \psi - i \sin 2\alpha \bar{\psi} \gamma^5 \psi \end{aligned} \quad (15.105)$$

that is, it is not invariant. Namely, chiral invariance requires the vanishing of the fermion mass, which, in turn, makes the left- and right-handed components independent, as shown by Eq. (15.104), in the absence of interactions to couple left- and right-handed components. One notices that Yukawa-type interaction $\mathcal{L}_{\text{Yukawa}} \sim (g\bar{\psi}\psi + g'\bar{\psi}\gamma^5\psi)\phi$ mixes the L-R components but current-type interactions such as $\mathcal{L}_{\text{EM}} \sim e\bar{\psi}\gamma^\mu\psi A_\mu$ or four-fermion interaction do not. The chiral transformation can be considered as independent gauge transformations for left and right components and is referred to as the chiral gauge transformation:

$$\psi_L \rightarrow \psi'_L = e^{-i\alpha_L} \psi_L, \quad \psi_R \rightarrow \psi'_R = e^{-i\alpha_R} \psi_R \quad (15.106)$$

10) The axial current is not conserved because $C_A = -1.26$, but the deviation is small and one may say it is almost conserved. In a loose sense, this is what PCAC means. In a more rigorous definition, it is expressed as [162]

$$\partial_\mu j_A^\mu = f_\pi m_\pi^2 \phi_\pi \quad (15.101)$$

which is conserved in the limit $m_\pi \rightarrow 0$. Here, f_π is the decay constant defined in Eq. (15.44) and ϕ_π is the pion field.

Problem 15.6

Show that α in Eq. (15.105) and α_L, α_R in Eq. (15.106) are connected by $\alpha = (\alpha_R - \alpha_L)/2$.

The fact that the weak Lagrangian has the form that is required by chiral gauge invariance has a fundamental significance for an understanding of the weak interaction. It means that the fermion mass has to vanish and that the left- and right-handed components are independent, i.e. different particles. Indeed, if the weak force is generated by the weak charge, only the left-handed component carries it. The CVC hypothesis and chiral invariance provide an important clue for the unification of the electromagnetic and weak forces.

Experimental Test of CVC

(i) **Weak Magnetism:** Using Lorentz invariance, a general expression for the proton electromagnetic current can be written as

$$\begin{aligned} J_{EM}^\mu &= \langle p' | \bar{p} \gamma^\mu p | p \rangle \\ &= \bar{u}(p') \left[F_{p1}(q^2) \gamma^\mu + \frac{\chi_p}{2m_p} F_{p2}(q^2) i \sigma^{\mu\nu} q_\nu \right] u(p) \\ q^2 &= (p - p')^2, \quad F_{p1}(0) = F_{p2}(0) = 1 \end{aligned} \quad (15.107)$$

If the proton is a point particle, i.e. a Dirac particle without higher order corrections, $F_{p1}(q^2) = 1$, $F_{p2}(q^2) = 0$. χ_p is the anomalous magnetic moment of the proton in units of nuclear magneton $e/2m_p$. Note the term proportional to $q^\mu = p^\mu - p'^\mu$ vanishes by current conservation and that proportional to $p^\mu + p'^\mu$ can be expressed by the above two terms using the Dirac equation for $\bar{u}(p')$, $u(p)$ [see Eq. (8.5)]. The electromagnetic current of the neutron can be obtained by changing the subscript $p \rightarrow n$. Since the neutron has vanishing electric charge but finite anomalous magnetic moment, $F_{n1}(0) = 0$, $F_{n2}(0) = 1$. Comparing Eq. (15.107) with Eq. (15.98), we see the τ_3 part of Eq. (15.98) may be written as

$$J_{EM}^\mu = \bar{\psi} \left[F_{V1}(q^2) \gamma^\mu + \frac{\chi_V}{2m_p} F_{V2}(q^2) i \sigma^{\mu\nu} q_\nu \right] \frac{\tau_3}{2} \psi \quad (15.108a)$$

$$F_{V1} = F_{p1} - F_{n1} \xrightarrow{q^2 \rightarrow 0} 1 \quad (15.108b)$$

$$\begin{aligned} \chi_V F_{V2} &= \chi_p F_{p2} - \chi_n F_{n2} \xrightarrow{q^2 \rightarrow 0} \chi_p - \chi_n \\ \chi_p &= 2.792847351(28), \quad \chi_n = -1.91302473(45) \end{aligned} \quad (15.108c)$$

where the number in parenthesis of the experimental value denotes the error in the last two digits. The CVC hypothesis means the weak charged current Eq. (15.97) is the same as the terms in square brackets of Eq. (15.108a). Namely, the vector part of the weak current should be expressed as

$$J_W^{+\mu} = \bar{\psi} \left[F_{V1}(q^2) \gamma^\mu + \frac{\chi_V}{2m_p} F_{V2}(q^2) i \sigma^{\mu\nu} q_\nu \right] \tau^+ \psi \quad (15.109)$$

Therefore, if the second term (referred to as weak magnetism) in the square brackets of Eq. (15.109) can be shown to be the same as the anomalous magnetic moment of the nucleon, the CVC hypothesis is proved. As the term is proportional to q , it does not contribute to the allowed transition of β decay and one has to turn to the forbidden transition to discuss the matter. Since we do not want to go into too many details of nuclear theory, we do not discuss the CVC hypothesis in the β decay any more here. Historically, however, the weak magnetism was measured [259] and provided the first evidence of the CVC hypothesis.

Another test is to perform high-energy neutrino scattering by a proton ($\nu_l + p \rightarrow l + n$) and see if it has the same q^2 dependence as ep elastic scattering ($e + p \rightarrow e + p$) [9]. Here, we consider the effect of CVC in pion decay.

(ii) $\pi^+ \rightarrow \pi^0 + e^+ + \nu$: The electromagnetic current of $\pi^\pm$ is proportional to the third component of the isospin by the Nishijima–Gell-Mann rule. Considering Lorentz invariance and separating the coupling constant, one finds

$$\begin{aligned} J_{\text{EM}}^\mu &= \langle \pi^+(k') | j_{\text{EM}}^\mu(x) | \pi^+ \rangle \\ &= [F_1(q^2)(k'^\mu + k^\mu) + F_2(q^2)(k'^\mu - k^\mu)] \langle \pi^+ | I_3 | \pi^+ \rangle e^{-iq \cdot x} \\ q^2 &= (k' - k)^2 \end{aligned} \quad (15.110)$$

where k, k' are the initial and final momenta of the pion. The form factor F_1 is normalized to become 1 in the static limit ($q^2 = 0$). Current conservation means

$$\begin{aligned} \partial_\mu j_{\text{EM}}^\mu &= 0 \\ \therefore F_1(q^2)(k'^2 - k^2) + F_2(q^2)(k' - k)^2 &= 0 \end{aligned} \quad (15.111)$$

Substitution of $k'^2 = k^2 = m_\pi^2$ makes $F_2(q^2) = 0$. Therefore the electromagnetic current is expressed as

$$J_{\text{EM}}^\mu = (k'^\mu + k^\mu) F_1(q^2) \langle f | I_3 | i \rangle e^{-iq \cdot x} \quad (15.112)$$

The CVC hypothesis requires

$$J_W^\mu = (k'^\mu + k^\mu) F_1(q^2) \langle f | I_- | i \rangle e^{-iq \cdot x} \quad (15.113)$$

since $I_- |1, +1\rangle = \sqrt{2} |1, 0\rangle$, i.e. $\langle \pi^0 | I_- | \pi^+ \rangle = \sqrt{2}$ [Eqs. (E.3), (E.10)]. For the π^+ decay q^2 is small and the form factor can be approximated by $F_1(q^2) \simeq F_1(0) = 1$. We have the matrix element

$$\mathcal{M}(\pi^+ \rightarrow \pi^0 e^+ \nu) = \frac{G_\beta}{\sqrt{2}} \sqrt{2} (k'^\mu + k^\mu) \bar{u}(p_e) \gamma_\mu (1 - \gamma^5) v(p_\nu) \quad (15.114a)$$

Then the decay rate can be calculated as

$$\Gamma = (2\pi)^4 \delta^4(\Delta - p_e - p_\nu) \frac{1}{2m_\pi} \sum_{\text{spin}} |\mathcal{M}|^2 \frac{d^3 p_\pi}{(2\pi)^3 2\omega} \frac{d^3 p_e}{(2\pi)^3 2E_e} \frac{d^3 p_\nu}{(2\pi)^3 2E_\nu} \quad (15.114b)$$

The result is given in [104, 222]:

$$\Gamma = \frac{G_\beta^2 \Delta^5}{30\pi^3} \left(1 - \frac{1}{2} \frac{\Delta}{m_{\pi^+}}\right)^3 R\left(\frac{m_e}{\Delta}\right) \simeq 0.393 \text{ s}^{-1}$$

$$R(x) = \sqrt{1-x^2} \left(1 - \frac{9}{2}x^2 - 4x^4\right) + \frac{15}{2}x^4 \ln\left(\frac{1+\sqrt{1-x^2}}{x}\right) \quad (15.115)$$

The observed value is $0.394 \pm 0.01 \text{ s}^{-1}$ and agrees with the CVC hypothesis very well.

Problem 15.7

*

As $\Delta = m_{\pi^+} - m_{\pi^0} \ll m_{\pi^0}$, we may neglect the π^0 momentum, which gives $k'^\mu + k^\mu = (m_{\pi^+} + m_{\pi^0})\delta_0^\mu$. In this approximation, show Eq. (15.114b) becomes

$$\Gamma = \frac{G_\beta^2 \Delta^5}{30\pi^3} \frac{(1 - \Delta/2m_{\pi^+})^2}{1 - \Delta/m_{\pi^+}} R\left(\frac{m_e}{\Delta}\right) \simeq 0.414 \text{ s}^{-1} \quad (15.116)$$

which agrees with Eq. (15.115) within a few %.

15.6

Strange Particle Decays

15.6.1

$\Delta S = \Delta Q$ Rule

Particles carrying strangeness may not have the same interaction as the familiar hadrons composed of u and d quarks. So we shall see first if we can apply the same rules to phenomena associated with strange particles if we replace the d quark by the s quark. As the hadronic part of the Lagrangian of the β decay was expressed as

$$j_W^{+\mu} \sim \bar{u}\gamma^\mu(1 - \gamma^5)d \quad (15.117)$$

it is natural to think that the decay of strange particles is induced by

$$j_W^{+\mu} \sim \bar{u}\gamma^\mu(1 - \gamma^5)s \quad (15.118)$$

with leptonic part unchanged. Under this assumption, the relative ratio of $K^\pm \rightarrow e^\pm + \nu$ and $K^\pm \rightarrow \mu^\pm + \nu$ can be calculated in the same way as for the pion, leading to the formula corresponding to Eq. (15.49):

$$\frac{\Gamma(K \rightarrow e\nu)}{\Gamma(K \rightarrow \mu\nu)} = \frac{m_e^2(m_K^2 - m_e^2)^2}{m_\mu^2(m_K^2 - m_\mu^2)^2} \quad (15.119)$$

That the formula gives a value which agrees well with the experimental value $2.45 \pm 0.11 \times 10^{-5}$ supports the above assumption. If the hadronic weak current is expressed by Eq. (15.118) this is an operator to annihilate s and create u (and vice versa). Since u carries ($I = I_3 = 1/2$) and s ($S = -1, I = 0$), we expect that in the strangeness changing reaction, relations

$$s \rightarrow u \quad Q: -\frac{1}{3} \rightarrow \frac{2}{3}, \quad I: 0 \rightarrow \frac{1}{2}, \quad S: -1 \rightarrow 0 \quad (15.120)$$

hold which means that the quantum number of the hadronic part undergoes changes

$$\Delta S = \pm 1, \quad \Delta S = \Delta Q, \quad \Delta I = 1/2 \quad (15.121)$$

This is also obtained from Nishijima–Gell-Mann’s law:

$$Q = I_3 + \frac{1}{2}(B + S) \quad (15.122)$$

Consider the list of observed leptonic decays of the strange particles

$$K^\pm \rightarrow \pi^0 + e^\pm + \bar{\nu} \quad (15.123a)$$

$$K_L^0 \rightarrow \pi^\pm + e^\mp + \bar{\nu} \quad (15.123b)$$

$$\Lambda \rightarrow p + e^- + \bar{\nu} \quad (15.123c)$$

$$\Sigma^- \rightarrow n + e^- + \bar{\nu} \quad (15.123d)$$

$$\Xi^0 \rightarrow \Sigma^+ + e^- + \nu \quad (15.123e)$$

The regularity is obvious. They all satisfy $\Delta S = \Delta Q$. On the other hand, a reaction with $\Delta S = -\Delta Q$ such as

$$\Sigma^+ \rightarrow n + e^+ + \bar{\nu} \quad (15.124)$$

has not been observed and its decay rate relative to (15.123d) is below 0.009. It is easy to understand if we think in terms of the constituent quarks. As $\Sigma^- = uus$, s can be changed to u by emitting W^- , converting Σ^- to $n = uud$, but $\Sigma^+ = uus$ can only be changed to uuu and not to $n = udd$. K_L^0 can be obtained either from K^0 or $\bar{K}^0$ and it is a mixture of $S = 1$ ($d\bar{s}$) and $S = -1$ ($s\bar{d}$). Therefore, it should be able to decay to both, $\pi^\pm l^\mp \bar{\nu}$. However, once K_L ’s parent is identified as K^0 or $\bar{K}^0$, one of them violates the $\Delta S = \Delta Q$ rule and should be forbidden. One can measure an observable

$$x = \frac{(\Delta S = -\Delta Q) \text{ rate}}{(\Delta S = \Delta Q) \text{ rate}} = \frac{\langle \pi^- l^+ \nu | T | \bar{K}^0 \rangle}{\langle \pi^- l^+ \nu | T | K^0 \rangle} \quad (15.125)$$

where $K_L \rightarrow \pi^- l^+ \bar{\nu}$ in the numerator and denominator are separately obtained from $\bar{K}^0$ and K^0 . The measured value gives [311]

$$\text{Re}\{x\} = -0.0018 \pm 0.006, \quad \text{Im}\{x\} = -0.0012 \pm 0.002 \quad (15.126)$$

and is consistent with 0. A more comprehensive treatment is given in the next chapter.

15.6.2

$\Delta I = 1/2$ Rule

We turn now to the $\Delta I = 1/2$ rule. It is easy to understand it in leptonic or semileptonic decays in Eq. (15.123), but consider hadronic decays such as

$$A \rightarrow p + \pi^-, \quad A \rightarrow n + \pi^0 \quad (15.127)$$

The transition amplitude may be written as

$$T_{fi} = G_S \langle p | \bar{u} \gamma^\mu (1 - \gamma^5) s | A \rangle \langle \pi^- | \bar{d} \gamma_\mu (1 - \gamma^5) u | 0 \rangle \quad (15.128)$$

where G_S has been used instead of G_β , since it is to be determined by experiment. The above expression contains both a $\Delta I = 1/2$ ($\bar{u}s$) and a $\Delta I = 1$ ($\bar{d}u$) part, allowing both $\Delta I = 1/2$ and $\Delta I = 3/2$ transitions. Experimentally, however, the $\Delta I = 1/2$ rule seems to be respected in hadronic decays as well as leptonic decays. If one requires the rule, we should have

$$\frac{\Gamma(A \rightarrow n\pi^0)}{\Gamma(A \rightarrow n\pi^0) + \Gamma(A \rightarrow p\pi^-)} = \frac{1}{3} \quad (15.129)$$

Correcting for the phase space difference, the right-hand side becomes 0.345, which is to be compared with the experimental value 0.358 ± 0.005 .

The decay $K^+ \rightarrow \pi^+ \pi^0$ ($I = 2$) is a $\Delta I = 3/2$ process (Problem 15.8). It is much less common than the $\Delta I = 1/2$ process $K_S^0 \rightarrow \pi\pi$ ($I = 0$):

$$\frac{\Gamma(K^+ \rightarrow \pi^+ \pi^0 (I = 2))}{\Gamma(K_S^0 \rightarrow \pi\pi (I = 0))} = 4 \times 10^{-2} \quad (15.130)$$

The validity of the $\Delta I = 1/2$ rule in hadronic decay is believed to originate from the dynamics of the strong interaction, but it is not well understood yet and remains as a problem for QCD to solve.

Problem 15.8

Show that $\pi^+ \pi^0$ cannot have $I = 1$ when $J = 0$. Therefore, if $\Delta I = 1/2$ is strict, $K^+ \rightarrow \pi^+ \pi^0$ is forbidden.

Problem 15.9

(a) Using the $\Delta I = 1/2$ rule show that

$$\frac{\Gamma(K_S \rightarrow \pi^+ \pi^-)}{\Gamma(K_S \rightarrow \pi^0 \pi^0)} = \frac{\Gamma(K_L \rightarrow \pi^+ \pi^-)}{\Gamma(K_L \rightarrow \pi^0 \pi^0)} = 2 \quad (15.131a)$$

$$\frac{\Gamma(\Xi^- \rightarrow \Lambda \pi^-)}{\Gamma(\Xi^0 \rightarrow \Lambda \pi^0)} = 2 \quad (15.131b)$$

(b) Assuming a^+ , a^- , a^0 are decay amplitudes for $\Sigma^+ \rightarrow n\pi^+$, $\Sigma^- \rightarrow n\pi^-$, $\Sigma^+ \rightarrow p\pi^0$, show that

$$a^+ + \sqrt{2}a^0 = a^- \quad (15.131c)$$

Hint: Assume the existence of a hypothetical particle, the “spurion”, carrying $I = 1/2$ and consider the decay as the scattering with the spurion where the isospin is conserved (see Sect. 13.3.1 for isospin analysis).

15.6.3

$$K_{l3} : K^+ \rightarrow \pi^0 + l^+ + \nu$$

Decay Rate As a representative of decay processes satisfying $|\Delta S| = 1$, we pick out $K^+ \rightarrow \pi^0 l^+ \nu_l$ ($l = e, \mu$). The reaction is similar to $\pi^+ \rightarrow \pi^0 e \nu$, Eq. (15.114a), as far as the kinematics is concerned. Therefore, denoting the coupling constant as G_S , the decay amplitude is given by

$$\begin{aligned} T_{fi} &= \frac{G_S}{\sqrt{2}} \langle \pi^0(k_\pi) | \bar{s} \gamma^\mu (1 - \gamma^5) u | K^+(k_K) \rangle \langle l^+ \nu | \bar{\nu} \gamma_\mu (1 - \gamma^5) l | 0 \rangle \\ &\equiv \frac{G_S}{\sqrt{2}} S^\mu l_\mu \\ S^\mu &= \langle \pi^0(k_\pi) | \bar{s} \gamma^\mu (1 - \gamma^5) u | K^+(k_K) \rangle \\ &= f_+(q^2) (k_K^\mu + k_\pi^\mu) + f_-(q^2) (k_K^\mu - k_\pi^\mu) \\ l_\mu &= \bar{u}(p_\nu) \gamma_\mu (1 - \gamma^5) v(p_e) \end{aligned} \quad (15.132)$$

As both K and π have negative parity, the axial vector current does not contribute. $f_\pm$ are functions of $q^2 = (k_K - k_\pi)^2$ and real if T invariant. As

$$k_K - k_\pi = p_e + p_\nu, \quad k_K + k_\pi = 2k_K - (p_e + p_\nu) \quad (15.133)$$

and multiplication of the leptonic current by the lepton momenta produces m_l , f_- is multiplied by m_l^2 . Therefore, when $l = e$, the f_- term can be neglected. As an extension of the CVC hypothesis, let us assume that the vector part of the strangeness-changing weak current is conserved, too. Then just like the isospin current [$SU(2)$ symmetry] was induced from CVC, $SU(3)$ symmetry can be applied here and $f_+(0)$ becomes $1/\sqrt{2}$. As $SU(3)$ symmetry is not exact, the correction

changes the value slightly ($1 \rightarrow 0.982 \pm 0.008$ [263]). Since q^2 is a small number, the form factor is usually parametrized as

$$f_{\pm}(q^2) = f_{\pm}(0) \left[1 + \lambda_{\pm}(q^2/m_{\pi}^2) \right] \quad (15.134)$$

in the experimental analysis. In the $K_{\mu 3}$ decay, both f_+ and f_- can be determined, but in $K_{e 3}$ decay only f_+ is determined. We discuss $K_{e 3}$ only here. The electron energy spectrum and $(\pi^0 \nu)$ angular correlation can be calculated from Eq. (15.132)

$$d\Gamma = (2\pi)^4 \delta^4(k_K - k_{\pi} - p_e - p_{\nu}) \frac{|\overline{\mathcal{M}}|^2}{2m_K} \frac{d^3 k_{\pi}}{(2\pi)^3 2\omega_{\pi}} \frac{d^3 p_e}{(2\pi)^3 2E_e} \frac{d^3 p_{\nu}}{(2\pi)^3 2E_{\nu}} \quad (15.135a)$$

$$\begin{aligned} |\overline{\mathcal{M}}|^2 &= \sum_{\text{spin}} \left| \frac{G_S}{\sqrt{2}} S^{\mu} l_{\mu} \right|^2 = \frac{G_S^2}{2} \sum_{\text{spin}} |f_+ \bar{u}(p_{\nu}) \not{k}_K (1 - \gamma^5) \not{p}_e (p_e)|^2 \\ &= 2G_S^2 |f_+|^2 \text{Tr}[(\not{p}_e - m) \not{k}_K (1 - \gamma^5) \not{p}_{\nu} \not{k}_K (1 - \gamma^5)] \\ &= 16G_S^2 |f_+|^2 [2(k_K \cdot p_e)(k_K \cdot p_{\nu}) - m_K^2(p_e \cdot p_{\nu})] \end{aligned} \quad (15.135b)$$

Electron Spectrum To obtain the electron spectrum, we have to integrate over k_{π} and p_{ν} . We set $f_+(q^2) = f_+(0)$ in the following calculation. We need to calculate

$$\int p_{\nu}^{\mu} \delta^4(p_{\nu} + k_{\pi} - a) \frac{d^3 p_{\nu}}{E_{\nu}} \frac{d^3 k_{\pi}}{\omega_{\pi}} = A a^{\mu}, \quad \text{where } a = k_K - p_e \quad (15.136a)$$

The equality is deduced from Lorentz invariance. To find A , we multiply by a_{μ} on both sides and, making use of $a^2 - m_{\pi}^2 = 2(a \cdot p_{\nu})$, we obtain

$$A = \frac{a^2 - m_{\pi}^2}{2a^2} \int \frac{d^3 p_{\nu}}{E_{\nu}} \frac{d^3 k_{\pi}}{\omega_{\pi}} \delta^4(p_{\nu} + k_{\pi} - a) \quad (15.136b)$$

The integration can be performed easily in a frame where $\mathbf{a} = 0$:

$$\begin{aligned} \int \frac{d^3 p_{\nu}}{E_{\nu}} \frac{d^3 k_{\pi}}{\omega_{\pi}} \delta^4(p_{\nu} + k_{\pi} - a) &= \int \frac{d^3 p_{\nu}}{\omega_{\pi} E_{\nu}} \delta(E_{\nu} + \omega_{\pi} - E_a) \\ &= 4\pi \frac{p_{\nu} a_0}{a_0^2} = 2\pi \frac{a^2 - m_{\pi}^2}{a^2} \end{aligned} \quad (15.136c)$$

As a result we obtain

$$A a^{\mu} = \pi \frac{(a^2 - m_{\pi}^2)^2}{a^4} a^{\mu} \quad (15.136d)$$

Inserting Eq. (15.136d) in Eq. (15.135) gives

$$d\Gamma = \frac{G_S^2 |f_+|^2}{(2\pi)^5 E_K} [2(k_K \cdot p_e)(k_K \cdot a) - m_K^2(p_e \cdot a)] \pi \frac{(a^2 - m_{\pi}^2)^2}{a^4} \frac{d^3 p_e}{E_e} \quad (15.137)$$

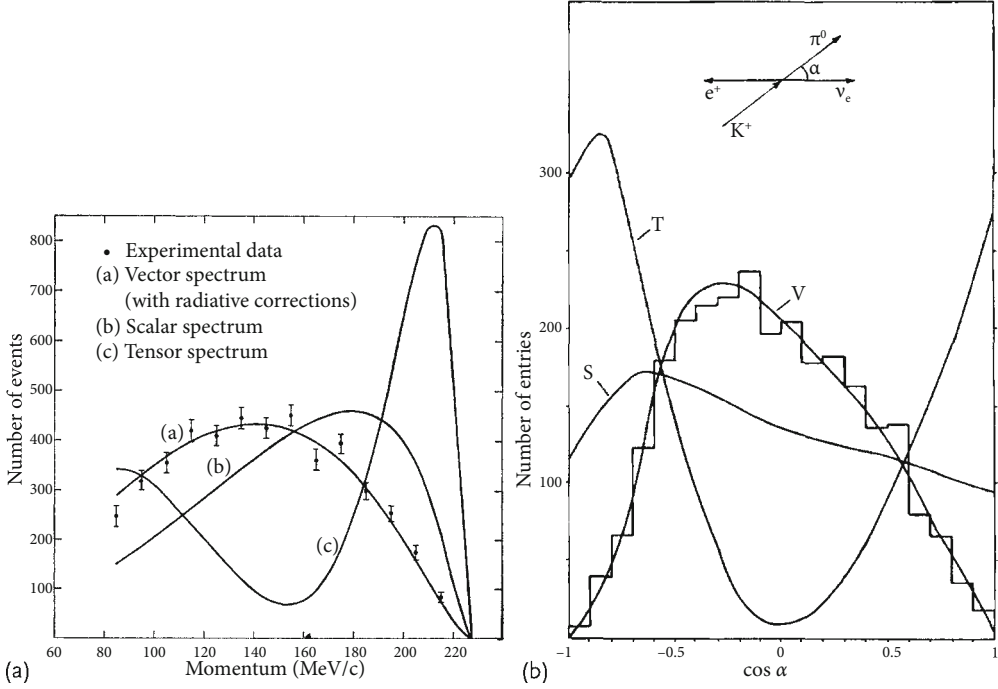


Figure 15.17 Comparison of V, S, T interaction type with $K^+ \rightarrow \pi^0 e^+ \nu_e$. (a) The energy spectrum of the decay electron [132]. (b) $\pi^0 \nu_e$ angular correlation. α is the angle between π^0 and ν in the dilepton center of mass system [75].

Using $a = k_K - p_e$, and evaluating variables in the kaon rest frame, we find

$$a^2 = m_K^2 - 2m_K E_e, \quad (15.138a)$$

$$a^2 - m_\pi^2 = 2m_K(E_{\max} - E_e), \quad E_{\max} = \frac{m_K^2 - m_\pi^2}{2m_K} \quad (15.138b)$$

$$2(k_K \cdot p_e)(k_K \cdot a) - m_K^2(p_e \cdot a) = m_K^2 E_e(m_K - 2E_e) \quad (15.138c)$$

where $E_{\max}$ is the maximum energy of the decay electron and we have set $m_e = 0$. Taking into account Eqs. (15.138), we finally obtain the energy spectrum of the decay electron:

$$\frac{d\Gamma}{dE_e} = \frac{G_S^2 |f_+|^2}{2\pi^3} m_K E_e^2 \frac{(E_{\max} - E_e)^2}{(m_K - 2E_e)} \quad (15.139)$$

This formula agrees with experimental data (Fig. 15.17a).

We have confirmed that the vector-type interaction (with maximal parity-violating component, of course) reproduces the experimental data very well. In deriving Eq. (15.139), we assumed $f_+(q^2) = \text{constant}$.

Form Factors of K_{l3} How can we derive information on $f_+(q^2)$? Since $q^2 = (k_K - k_\pi)^2 = m_K^2 + m_\pi^2 - 2m_K\omega_\pi$ is a function of only ω_π , information on $f_+(q^2)$ can be obtained by observing the energy spectrum of the pion. It can be calculated by performing a similar calculation as for the electron spectrum. The formula is

$$\frac{d\Gamma}{d\omega_\pi} = \frac{G_S^2 |f_+|^2}{12\pi^3} m_K k_\pi^3 \quad (15.140)$$

Then from K_{e3} data one can obtain λ_+ and from $K_{\mu 3}$ data, λ_- . However, recently, it has become more conventional to use [311]

$$f_0(q^2) = f_+(q^2) + f_-(q^2) \frac{q^2}{m_{K^+}^2 - m_{\pi^0}^2} \quad (15.141a)$$

$$f_0(q^2) = f_0(0) \left(1 + \lambda_0 \frac{q^2}{m_{\pi^+}^2} \right) \quad (15.141b)$$

The present value of observed $\lambda_\pm$ [311] is given by

$$\begin{aligned} \lambda_+ &= 0.0296 \pm 0.0006 \\ \lambda_0 &= 0.0196 \pm 0.0012 \\ \text{Im } \xi &\equiv \text{Im} \left(\frac{f_-}{f_+} \right) = -0.006 \pm 0.008 \end{aligned} \quad (15.142)$$

The information on $\text{Im } \xi$ comes from T (time reversal) violating transverse polarization P_T of μ [226], where

$$P_T = \text{Im } \xi \frac{m_\mu}{m_K} \frac{|\mathbf{p}_\mu|}{E_\mu + |\mathbf{p}_\mu| \cos \theta_{\mu\nu} - m_\mu^2/m_K} \quad (15.143)$$

Interaction Type In Fig. 15.17, theoretical spectra for the interaction type S (scalar) and T (tensor) are also given. They can be calculated starting from

$$T_{fi} = \frac{G_S}{\sqrt{2}} \left[2m_K f_S \{e(1 - \gamma^5) \bar{\nu}_l\} + \frac{2f_T}{m_K} k_K^\mu k_\pi^\nu \{e \sigma_{\mu\nu} (1 - \gamma^5) \bar{\nu}_l\} \right] \quad (15.144)$$

instead of Eqs. (15.132). The angular correlation between π^0 and ν in the dilepton center of mass frame is given by [75]

$$\rho(E_\pi, \cos \alpha) = \kappa \left[|f_S + f_T \cos \alpha|^2 m_K \omega_\pi + \frac{1}{2} |f_+|^2 k_\pi \sin^2 \alpha \right] \quad (15.145)$$

Figure 15.17b shows $\pi^0 \nu_e$ angular correlation. It is clear from the figure that V-A interaction is equally valid for strangeness-changing weak interactions.

Strength of Strangeness-Changing Weak Decays The total decay rate can be obtained from Eq. (15.139) by integrating over the electron energy:

$$\Gamma_{K_{e3}} = \frac{G_S^2 f_+^2 m_K^5}{768\pi^3} f\left(\frac{m_\pi^2}{m_K^2}\right) \quad (15.146)$$

$$f(\gamma) = 1 - 8\gamma + 8\gamma^3 - \gamma^4 - 12\gamma^2 \ln \gamma$$

$f(m_\pi^2/m_K^2) = 0.579$ for $m_\pi/m_K = 0.273$. Comparing Eq. (15.146) with the experimental value $\tau_K = 1.2380 \pm 0.0021 \times 10^{-8}$ s, $\text{BR}(K^+ \rightarrow \pi^0 e^+ \nu_e) = 0.0508 \pm 0.0005$, gives

$$G_S/G_\beta \sim 0.22 \quad (15.147)$$

Problem 15.10

Derive Eqs. (15.140) and (15.146).

15.6.4

Cabibbo Rotation

We have learned that the V–A interaction derived from the muon and nuclear β decays are equally valid for strangeness-changing decays. However, the measured coupling constant G_S^2 is approximately 1/20 of G_β^2 and is very small. This is also true for other strangeness-changing decays. In other words, the coupling is universal in the sense that the strength is equal for all strangeness-changing processes, yet the universality does not apply between reactions with $\Delta S = 0$ and $\Delta S \neq 0$. This very annoying situation was saved by introduction of the Cabibbo rotation [84].

It frequently happens in quantum mechanics that a good quantum number changes depending on what kind of interaction is active. For example, the orbital angular momentum is a good quantum number when the central force is active and spherical harmonic wave functions are suitable for the treatment of dynamics. However, when spin–orbit coupling is active, the total angular momentum is the good quantum number and the wave functions are mixtures of differing orbital angular momentum. Cabibbo thought that when the weak force is active, a state suitable for dynamic treatment may not necessarily be the mass eigenstate that is observable as a particle. The state suitable for describing the dynamical behavior of the particle is actually a mixture of mass eigenstates

$$d' = \cos \theta_c d + \sin \theta_c s \quad (15.148)$$

and the particle that participates in the weak interaction is not d but d' , which is referred to as the “weak eigenstate”. Cabibbo proposed that the weak current should be expressed as

$$j_W^{+\mu} = \bar{u} \gamma^\mu (1 - \gamma^5) d' \quad (15.149)$$

This is referred to as Cabibbo rotation. According to this idea, the Fermi coupling constant should be universal, but in experiments only the mass eigenstates are picked up. Consequently, only a component of the particle that is contained in the weak eigenstate is measured. The fraction is given by the (square of) of the Cabibbo rotation angle. In the strangeness-changing process the fraction is given by $\sin^2 \theta_c$ changing $G_\mu \rightarrow G_S = G_\mu \sin \theta_c$. From K_{e3} data [75]

$$\sin \theta_c = 0.220 \pm 0.002 \quad (15.150)$$

From $A \beta$ decay [72], it was determined as

$$\sin \theta_c = 0.231 \pm 0.003 \quad (15.151)$$

and on the average,

$$\sin \theta_c = 0.221 \pm 0.002^{11)} \quad (15.152)$$

Strictly speaking, the above value should be restricted to vector coupling only, but other measurements including the axial current also confirmed the same value. In summary, the V–A interaction can be applied equally well to all the strangeness-changing weak interactions by introducing a single parameter, the “Cabibbo angle” θ_c .

Cabibbo’s proposal means that the coupling constant G_β of the nuclear β decay is actually $G_\beta = G_\mu \cos \theta_c$. No Cabibbo rotation is necessary for the leptonic current. Since there was a small but significant difference between G_β and G_μ , as was shown in Eq. (15.89), identifying the difference as $\cos \theta_c$ gives the same value of the Cabibbo angle and

$$\cos \theta_c = 0.973 \pm 0.002^{12)} \quad (15.153)$$

From now on we use G_F to denote the strength of the universal four-Fermi interactions instead of G_μ in honor of Fermi.

11) When the quark family is extended to three generations including (c, s) , (t, b) , the Cabibbo rotation becomes the Kobayashi–Maskawa (KM) matrix. In this case $\sin \theta_c$ should be replaced by V_{12} or V_{us} . Numerically, however, the difference is very small.

12) For the same reason as in footnote 11, $\cos \theta_c$ should be replaced by V_{11} or V_{ud} . From the

unitarity of the Kobayashi–Maskawa matrix

$$|V_{11}|^2 + |V_{12}|^2 + |V_{13}|^2 = 1 \quad (15.154)$$

Hence $|V_{11}|^2 + |V_{12}|^2$ need not necessarily become unity, but it accidentally happened that $|V_{13}|^2 \ll |V_{11}|^2, |V_{12}|^2$, and the Cabibbo rotation, which closes within the two generations, worked well.

15.7

Flavor Conservation

15.7.1

GIM Mechanism

A careful reader might have already noticed that so far we have talked only about the charged current weak interaction or that mediated by charged $W^\pm$, where the participating members change their electric charge. It is natural to imagine that there is a neutral current interaction and it is compulsory if the weak force carrier W constitutes an $SU(2)$ triplet, as is the case in the Standard Model. Indeed, the existence of Z^0 and its associated reactions was among the major predictions of the Glashow–Weinberg–Salam model, and subsequent experimental confirmation established its validity. Historically and phenomenologically, the need for the neutral current was not strong. The effect of the ordinary (strangeness conserving) neutral current (i.e. $j_{\text{NC}}^\mu \sim \bar{u}\gamma^\mu u, \bar{d}\gamma^\mu d$) is hidden by much stronger electromagnetic currents, hence it is not surprising that their effect was not observed. Moreover, there was strong evidence of the absence (or suppression to be exact) of the strangeness-changing neutral current. For instance, consider the kaon decays in Table 15.1 that are induced by the charged current (CC) and the corresponding neutral current interaction (NC). Their Feynman diagrams are depicted in Fig. 15.18. Table 15.1 suggests the absence of Z^0 for the diagrams (b) and (d) in Fig. 15.18. However, there was no reason to believe in the nonexistence of the strangeness-changing neutral current and it was a mystery at that time. In 1970 Glashow, Iliopoulos and Maiani [174] proposed the existence of the fourth quark, and that it should make a doublet with s' , an orthogonal state to the Cabibbo-rotated d' :

$$\begin{aligned} d' &= d \cos \theta_c + s \sin \theta_c \\ s' &= -d \sin \theta_c + s \cos \theta_c \end{aligned} \quad (15.155)$$

The Cabibbo-rotated d' makes a neutral current

$$\bar{d}'d' = \bar{d}d \cos^2 \theta_c + \bar{s}s \sin^2 \theta_c + \sin \theta_c \cos \theta_c (\bar{s}d + \bar{d}s) \quad (15.156)$$

where space-time structure was omitted in the above equation. Equation (15.156) contains the strangeness-changing current $\bar{s}d + \bar{d}s$, which should induce diagrams

Table 15.1 Suppression of strangeness-changing neutral current decays.

CC	BR	NC	BR
$K^+ \rightarrow \mu^+ \nu_\mu$	$63.54 \pm 0.14\%$	$K^0 \rightarrow \mu^- \mu^+$	$6.84 \pm 0.11 \times 10^{-9}$
$K^+ \rightarrow \pi^0 e^+ \nu_e$	$5.08 \pm 0.05\%$	$K^+ \rightarrow \pi^+ \bar{\nu}_e \nu_e$	$1.5^{+1.3}_{-0.9} \times 10^{-10}$ ^a

^a The experimental value is for $K^+ \rightarrow \pi^+ + \sum_i \bar{\nu}_i \nu_i$.

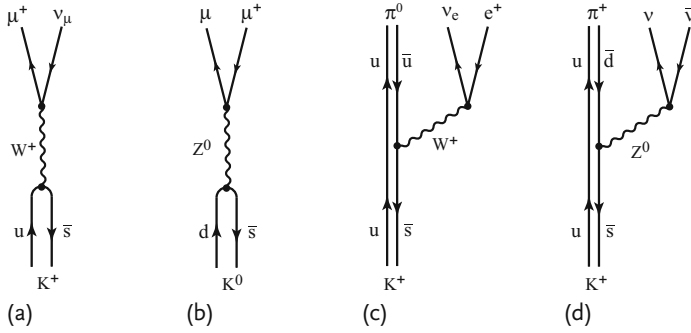


Figure 15.18 The charged current decays (a) $K^+ \rightarrow \mu^+ \nu$ and (c) $K^+ \rightarrow \pi^0 e^+ \nu_e$ and the strangeness-changing neutral current decays (b) $K^0 \rightarrow \mu^- \mu^+$ and (d) $K^+ \rightarrow \pi^+ \nu \bar{\nu}$.

(b) and (d) of Fig. 15.18. But if a state expressed as s' in Eq. (15.155) exists, the neutral current that $\bar{s}'s'$ makes

$$\bar{s}'s' = \bar{d}d \sin^2 \theta_c + \bar{s}s \cos^2 \theta_c - \sin \theta_c \cos \theta_c (\bar{s}d + \bar{d}s) \quad (15.157)$$

cancels the strangeness-changing part of $\bar{d}'d'$ and together with $\bar{d}'d'$ the neutral current becomes

$$\bar{d}'d' + \bar{s}'s' = \bar{d}d + \bar{s}s \quad (15.158)$$

The absence of the strangeness-changing current is explained. This is called the GIM mechanism. There had been other predictions of the fourth quark as early as 1964, but its existence was seriously considered only after the GIM proposal. The charm quark was subsequently discovered in 1974.

The reason why the strangeness-changing neutral current exists, albeit very suppressed, is due to the higher order charged current processes (Fig. 15.19). In the Standard Model, $G_F \simeq g_W^2/m_W^2$, and the value of g_W is comparable with that of e .¹³⁾ Therefore the contribution of higher order diagrams (Fig. 15.19) would be on the order of α , but the observed values are much smaller. This is because the GIM mechanism is at work in the higher order contributions. There are processes depicted in Fig. 15.19(b) or (d) corresponding to Fig. 15.19(a) or (c) that are obtained by replacing c with u . Since the coupling of W to ud or cs is $\sim \cos \theta_c$ and that to us or $cd \sim \sin \theta_c$ with differing sign, the contributions of (a) and (b) or (c) and (d) are equal but with opposite sign. If the mass of the charm quark were the same as that of the up quark, the cancellation would be perfect and the decays would not happen. Assuming $m_c \gg m_u$ the decay amplitude can be calculated and gives [154]

$$\mathcal{M}(K^0 \rightarrow \mu \bar{\mu}) \sim \alpha G_F \left(\frac{m_u^2}{m_c^2} \right)^2 \quad (15.159)$$

13) To be exact, $G_F/\sqrt{2} = g_W^2/(8m_W^2)$, $e = g_W/\sin \theta_W$, where θ_W is the Weinberg mixing angle of the electromagnetic and the weak interactions.

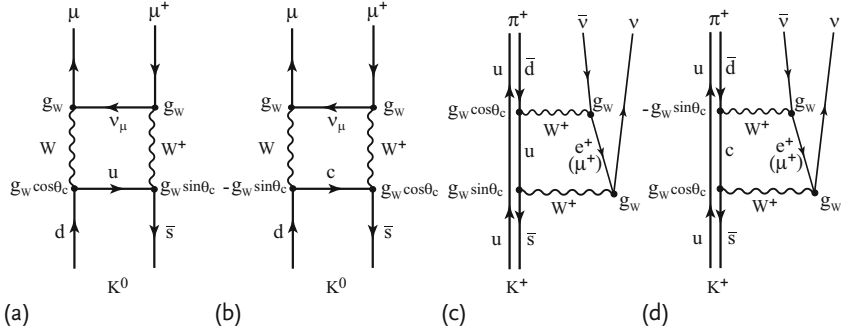


Figure 15.19 GIM mechanism at work in the higher order corrections. (a) and (b) are diagrams for $K^- \rightarrow \mu^- \mu^+$, (c) and (d) for $K^+ \rightarrow \pi^+ \nu \bar{\nu}$. If $m_c = m_d$, cancellation due

to the GIM mechanism should be complete and the above processes would be forbidden. The mass difference makes a finite contribution to (a) + (b) or (c) + (d).

Adjusting the mass value to fit the experimental value, one can obtain $m_c \sim 1.5$ GeV. The mass prediction was done before the charm quark was discovered. The significance of the c quark discovery was not just the confirmation of the fourth quark, it also established the doublet structure of the weak interaction and at the same time explained the absence of the strangeness-changing neutral current. Unlike the s quark, the c quark was predicted and anticipated. The discovery was an important milestone in the construction of the Standard Model. In summary, there is strong evidence for the absence of the strangeness-changing neutral current (SCNC) and there is no compelling reason to believe in the existence of the strangeness-conserving neutral current, at least from a phenomenological point of view.

15.7.2

Kobayashi–Maskawa Matrix

We know today that there are six quarks in three doublets with hierarchical structure in the sense that the mass of the quarks increases with the generation (see Table 19.1). The Cabibbo rotation is generalized to a 3×3 unitary transformation matrix, which is referred to as the Cabibbo–Kobayashi–Maskawa (or CKM) matrix [239]. Namely, the W boson couples to the three doublets

$$\begin{bmatrix} u \\ d' \end{bmatrix} \quad \begin{bmatrix} c \\ s' \end{bmatrix} \quad \begin{bmatrix} t \\ b' \end{bmatrix} \quad (15.160a)$$

with equal strength, and they are mixtures of the mass eigenstate $D = (d, s, b)$:

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = U_{\text{CKM}} \begin{bmatrix} d \\ s \\ b \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix} \quad (15.160b)$$

The unitarity of the CKM matrix is a GIM mechanism extended to the three generations:

$$\bar{d}'d' + \bar{s}'s' + \bar{b}'b = \bar{D}'D' = \bar{D}U^\dagger UD = \bar{D}D = \bar{d}d + \bar{s}s + \bar{b}b \quad (15.161)$$

There is no flavor-changing neutral current of any kind; this is not limited to strangeness only. Though there are nine independent variables in the 3×3 unitary matrix, considering the quark field's freedom to change its phase, the number is reduced to four: three rotation angles $\theta_{12}, \theta_{23}, \theta_{13}$ and one phase δ . It is conventional to express the CKM matrix in the form

$$\begin{aligned} U_{\text{CKM}} &= R_{23} I_{\delta_D} R_{13} I_{\delta_D}^\dagger R_{12} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \\ I_{\delta_D} &= \text{diag}(1, 1, e^{i\delta}), \quad s_{12} \simeq 0.220, \quad s_{23} \simeq 0.04, \quad s_{13} \simeq 0.004 \end{aligned} \quad (15.162)$$

Here, R_{ij} is a rotation matrix on the ij plane and c_{ij}, s_{ij} stand for $\cos \theta_{ij}, \sin \theta_{ij}$. The Cabibbo angle $\theta_c = \theta_{12}$, but the 2×2 Cabibbo rotation matrix worked well because $c_{13} \simeq 1, s_{13} \simeq 0$. However, the $b \rightarrow u$ transition [actually $B^-(b\bar{u}) \rightarrow \pi^0(u\bar{u}) + e^+ + \nu_e$ etc.] has been observed, showing $s_{13} \neq 0$.

It is worth noting that the motivation for introducing the CKM matrix was not merely an extension from two to three generations. Indeed, Kobayashi and Maskawa proposed it in 1973 before the discovery of the charm quark. Its motivation was to introduce a complex phase to explain CP violation, which will be explained in the next chapter in more detail.

15.7.3

Tau Lepton

The τ lepton was discovered in 1975 [318] and is one of the few elementary particles that can be observed in its bare form. The decay of the τ has many features in common with muon decay, and is suitable for investigating the universality of the weak interaction. We shall describe a few properties of τ and how to determine them, then we discuss general features of the weak interaction in the lepton sector.

Production If there is enough energy ($\sqrt{s} > 2m_\tau$), a τ pair can be produced in the reaction

$$e^- + e^+ \rightarrow \tau^- + \tau^+ \quad (15.163)$$

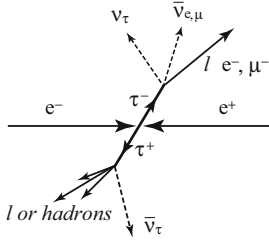


Figure 15.20 Topology of $e^- + e^+ \rightarrow \tau^- + \tau^+$ production. τ decays to $l + \nu_l + \nu_\tau$ or ν_τ + hadrons. As τ 's are short lived, only decay products of the pair are observed. A typical topology is an acoplanar lepton pair.

Decay patterns of the τ lepton are written as

$$\tau^\pm \rightarrow \overset{c \rightarrow}{\nu}_\tau + (W^\pm) \quad (15.164)$$

$$(W^\pm) \rightarrow \begin{cases} e^\pm + \overset{c \rightarrow}{\nu}_e \\ \mu^\pm + \overset{c \rightarrow}{\nu}_\mu \\ q_i \bar{q}_j \rightarrow \text{hadrons} : q_i \bar{q}_j = u\bar{d}, u\bar{s}, \text{ etc.} \end{cases}$$

where $(W^\pm)$ is virtual. As the τ lepton has very short decay lifetime ($2.910 \pm 0.015 \times 10^{-13}$ s), what are observed are decay products of $\tau^- \tau^+$ pairs (Fig. 15.20). If both tau's decay leptonically, the topology consists of an acoplanar¹⁴⁾ lepton pair ($e^- e^+$, $e^- \mu^+$, $\mu^- \mu^+$) and it is easily recognized. If only one τ decays leptonically, the topology is a single lepton in one hemisphere and a few hadrons in the other hemisphere, another combination easy to recognize.

Mass and Spin The production cross section of a spin 1/2 particle is given by [see Eq. (7.38)]

$$\sigma_{1/2} = \frac{4\pi\alpha^2}{3} \frac{1}{s} \quad (15.165)$$

The equation is common to $ee \rightarrow \mu\mu, \tau\tau$ if $s \gg (2m_l)^2$. Immediately above the threshold, however, the effect of the produced mass or nonrelativistic correction is necessary. If we write the Lorentz factor of the particle $\beta = p_\tau/E_\tau$, $\gamma = (1-\beta^2)^{-1/2}$, the cross section is modified as (see Problem 7.2)

$$\sigma_{\text{TOT}} = \sigma_{1/2} \beta \left(\frac{3-\beta^2}{2} \right) \quad (15.166)$$

In order to determine the spin of τ , theoretical cross sections for different spins

¹⁴⁾ Acoplanar means that one of the pair is not on a plane made by the incoming $e^- e^+$ and the other particle.

were considered [370, 371]:

$$\sigma_{\tau\tau} = \sigma_{1/2} F(\beta) \quad (15.167a)$$

$$F(\beta) = \begin{cases} \beta^3/4 & s = 0 \\ \beta(3 - \beta^2)/2 & s = 1/2 \\ \beta^3(3 + 4\gamma^2)/4 & s = 1, \kappa = 0 \\ \beta^3(3 + 2\gamma^2 + 4\gamma^4)/4 & s = 1, \kappa = 1 \\ \beta(15\gamma^{-2} + 30\beta^2 + 40\beta^4\gamma^2 + 16\beta^6\gamma^4)/9 & s = 3/2, A = 1, \\ & B = C = D = 0 \end{cases} \quad (15.167b)$$

where the magnetic moment of the spin 1 particle is defined as $\mu = (1 + \kappa)e/2M$. Note that a fermion–antifermion pair is produced in S-wave ($\sigma \propto |\mathbf{p}|$) and the cross section rises sharply at the threshold as a function of energy, while that of boson pairs is in P-wave ($\sigma \propto |\mathbf{p}|^3$). This is due to parity conservation in e^-e^+ reaction, where the dominant contribution comes from one-photon exchange, which has negative parity. The boson pairs have positive intrinsic parity, hence S-wave is forbidden.

Figure 15.21 compares the measured cross section with theoretical predictions for various spins [317]. It is clear that τ has $s = 1/2$. Fitting the cross section with the theoretical curve immediately above the threshold gave the mass as

$$m_\tau = 1777.0 \pm 0.3 \text{ MeV} \quad (15.168)$$

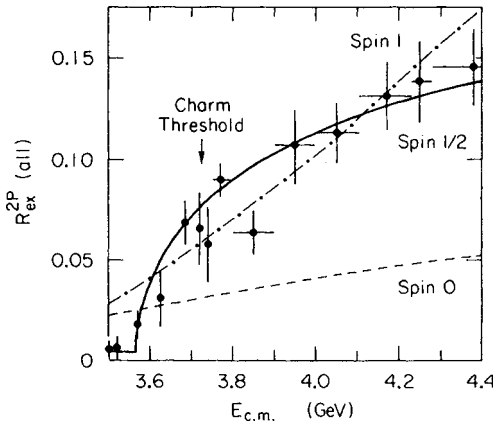


Figure 15.21 Comparison of the total production cross section $e\bar{e} \rightarrow \tau\bar{\tau}$ with theoretical predictions for various spins. $R_{\tau\tau} = \sigma(e\bar{e} \rightarrow \tau\bar{\tau})/\sigma(e\bar{e} \rightarrow \mu\bar{\mu})$ [35, 317].

Coupling Type As the τ has a large mass, it has many decay modes. Among them, purely leptonic modes

$$\begin{aligned}\tau^- &\rightarrow e^- + \bar{\nu}_e + \nu_\tau \\ \tau^- &\rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau\end{aligned}\quad (15.169)$$

have the same kinematical structure as muon decay. Therefore, investigation of the energy spectrum of the decay leptons allows the Michel parameters to be determined (Fig. 15.22a,b), [28, 311]):

$$\rho = 0.745 \pm 0.008 \quad (15.170)$$

The value agrees with that of V–A interaction.

Universality The leptonic decay rate can be calculated using the identical formula as for muon decay [Eq. (15.84)] with the mass replaced by m_τ :

$$\Gamma(\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau) = \frac{G_\tau^2 m_\tau^5}{192\pi^3} \quad (15.171)$$

Then the lifetime of τ is given by

$$\tau_\tau = \tau_\mu \left(\frac{G_\mu}{G_\tau} \right)^2 \left(\frac{m_\mu}{m_\tau} \right)^5 \text{BR}(\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau) \quad (15.172)$$

Setting $G_\mu = G_\tau$ and inserting the observed values for $m_\tau = 1777.0 \pm 0.3$ MeV, $m_\mu = 105.658389 \pm 0.000034$ MeV, $\tau_\mu = 2.19703 \pm 0.00004 \times 10^{-6}$ s, $\text{BR} = 0.1783 \pm 0.0008$, one finds the calculated value agrees with the experimental value of $\tau_\tau(\text{exp}) = 2.910 \pm 0.015 \times 10^{-13}$ s within the error. Conversely, using the experimental value to test the universality of the coupling constant gives [164, 169]

$$\frac{G_\tau}{G_\mu} = 0.999 \pm 0.006 \quad (15.173)$$

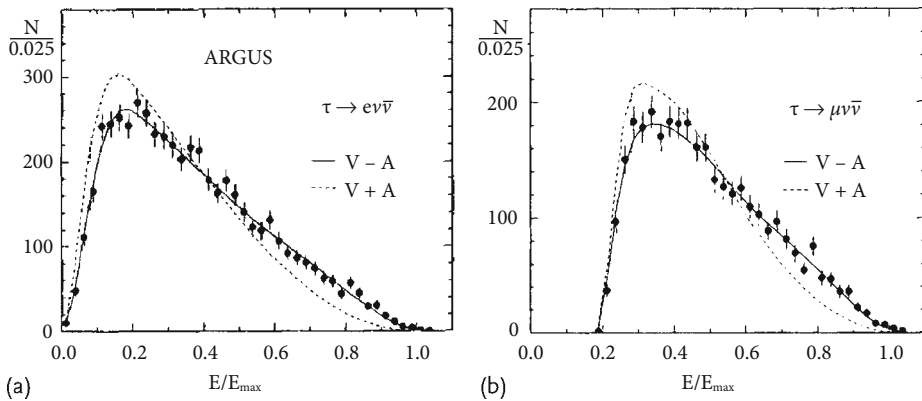


Figure 15.22 Comparison of the τ energy spectrum with $V \pm A$ predictions. (a) $\tau \rightarrow e \nu \bar{\nu}$, (b) $\tau \rightarrow \mu \nu \bar{\nu}$ [28].

If the weak interaction actually occurs through W exchange, then $G_\mu \sim g_\mu^2 g_e^2$ and $G_\tau \sim g_\tau^2 g_e^2$. So the above universality applies to $g_\mu - g_\tau$, giving $g_\mu/g_\tau = 1.001 \pm 0.003$. By investigating the ratio of $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$ and $\tau \rightarrow e \bar{\nu}_e \nu_\tau$, we can also test $g_\mu - g_e$ universality. Using

$$\frac{\Gamma(\tau^- \rightarrow \mu \bar{\nu}_\mu \nu_\tau)}{\Gamma(\tau^- \rightarrow e \bar{\nu}_e \nu_\tau)} = \frac{\text{BR}(\tau^- \rightarrow \mu \bar{\nu}_\mu \nu_\tau)}{\text{BR}(\tau^- \rightarrow e \bar{\nu}_e \nu_\tau)} = \frac{g_\tau^2 g_\mu^2}{g_\tau^2 g_e^2} \frac{f(m_\mu^2/m_\tau^2)}{f(m_e^2/m_\tau^2)} \quad (15.174)$$

where $f(y)$ is given in Eq. (15.85) and comparing with the experimental value

$$\frac{\text{BR}(\tau^- \rightarrow \mu \bar{\nu}_\mu \nu_\tau)}{\text{BR}(\tau^- \rightarrow e \bar{\nu}_e \nu_\tau)} = \frac{(17.36 \pm 0.05)\%}{(17.85 \pm 0.05)\%} = 0.974 \pm 0.004 \quad (15.175)$$

gives $g_\mu/g_e = 1.001 \pm 0.002$.

We may say the universality is very good.

Branching Ratio and Quark Counting As the mass of τ is large, there are many hadronic decay modes. But as $m_c > m_\tau$, the hadronic decay modes are limited to

$$\tau^- \rightarrow \nu_\tau + W^- \rightarrow \nu_\tau + \bar{u} + d' \quad (15.176a)$$

$$= \begin{cases} \tau^- \rightarrow \nu_\tau + d + \bar{u} & \text{probability} = \cos^2 \theta_c \\ \tau^- \rightarrow \nu_\tau + s + \bar{u} & \text{probability} = \sin^2 \theta_c \end{cases} \quad (15.176b)$$

Owing to the Cabibbo rotation, τ decays to $d\bar{u}$ with probability $\cos^2 \theta_c$ and to $s\bar{u}$ with $\sin^2 \theta_c$; their combined decay rate can be considered as one to a single channel d' . At the quark level, the decay rate $W^- \rightarrow \bar{u}d'$ is the same as the leptonic decay ($W^- \rightarrow e^- \bar{\nu}_e$ or $\mu^- \bar{\nu}_\mu$) provided the mass of the quarks and leptons is neglected. It is a good approximation since $m_e \sim m_u \sim m_d \sim$ a few MeV, $m_\mu \sim 100$ MeV. Even the heaviest μ is much lighter than $m_\tau (\simeq 2 \text{ GeV})$. Since the quark has an extra color degree of freedom, the decay rate to the quarks has to be tripled. Therefore a total of five channels are open for τ decay. Assuming all the quarks hadronize, the electronic and muonic decay modes are predicted each to be $1/5 = 20\%$. This is very close to the observed value of 17.7%. It demonstrates a practical use of the quark model, that simple quark counting gives fairly good agreement with experiment. A more precise estimate considering the QCD correction gives much better agreement [169].

15.7.4

The Generation Puzzle

From what we have analyzed so far, the three leptons e, μ, τ have the same properties except for their mass. Various kinds of reaction rates can be accurately reproduced by simply changing the mass. Despite their similarity, they have their own quantum number, referred to as lepton flavor. (ν_e, e^-) has e -quantum number, (ν_μ, μ^-) μ -number and (ν_τ, τ^-) τ -number, which are conserved in the weak interaction, and the three pairs do not mix. The fact that the three neutrinos are also distinct, each having separate flavor, can be deduced from the following observations.

From the energy spectrum of the pure leptonic decay of the muon ($\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$) and the τ lepton ($\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau, \mu^- \bar{\nu}_\mu \nu_\tau$), we know there are at least two kinds of neutral particles in each decay, which we call neutrinos. That ν_e and ν_μ have different flavor can be deduced from the absence of $\mu \rightarrow e \gamma$, as discussed in Sect. 15.4.4. Using the same arguments and from the absence of $\tau \rightarrow \mu \gamma, e \gamma$, we can deduce $\nu_\tau \neq \nu_\mu, \nu_e$. Other evidence comes from the fact that ν_e, ν_μ cannot produce τ directly. Precision tests of the leptonic flavors come from the absence of the following decay modes:

$$\begin{array}{ll}
 \text{BR}(\mu \rightarrow e \gamma) & < 4.9 \times 10^{-11} \\
 \text{BR}(\mu \rightarrow e e e) & < 1.0 \times 10^{-12} \\
 \text{BR}(\tau \rightarrow e \gamma) & < 1.1 \times 10^{-4} \\
 \text{BR}(\tau \rightarrow \mu \gamma) & < 4.2 \times 10^{-6} \\
 \text{BR}(\tau \rightarrow 3l) & < 3.3 \times 10^{-6} \\
 \text{BR}(Z \rightarrow l_1 l_2) (l_1 \neq l_2) & < 1.7 \times 10^{-6} \quad (l_1 = e, l_2 = \mu)
 \end{array} \tag{15.177}$$

Today, lepton flavor conservation is accepted and is built in to the Standard Model. The flavor does not mix in the Standard Model. If one adds a small mass¹⁵⁾ to the neutrino, which is a natural extension of, but nonetheless one step beyond, the Standard Model, the mixing is proportional to $(m_\nu/m_W)^2$ [97], which is extremely hard, to say the least, to detect in decay reactions. Why do leptons with the same properties appear repeatedly? This is a question not resolved by the Standard Model. It is the same old problem of Rabi's "who ordered it?" extended to the e - μ - τ puzzle. It is also referred to as the family or generation problem, because the same question exists in the quark sector. Namely, (c, s) and (t, b) seem to be replicas of (u, d) except for their mass. Combining the quark and lepton doublets, (u, d, ν_e, e^-) , (c, s, ν_μ, μ^-) and (t, b, ν_τ, τ^-) constitute three generations.

One idea to resolve the puzzle is to consider that they belong to the same symmetry group, called horizontal (or family) symmetry. Another is to consider that leptons and quarks are composites of more fundamental particles. In this model, μ, τ are excited states of the electron. In modern string theory the family structure arises from vacuum or space-time structure of extra dimensions. At present, there are no hints as to how to differentiate between the models.

How Many Generations? So far we have discovered three generations of the quark/lepton family. The Standard Model assumes there are no more. The assertion is not supported by compelling evidence, however, there is an indication that the number of generations is indeed limited to three. It came later, after the Standard Model was established, but we will jump ahead. It comes from the absence of an extra neutrino in the decay of Z^0 . Z^0 is produced and decays in the following

15) If there are mass differences between the neutrinos and if they mix, neutrino oscillation occurs (see Sect. 19.2.1), as has been observed.

reactions:

$$e^- + e^+ \rightarrow Z \rightarrow \begin{cases} l\bar{l} & l = e, \mu, \tau \\ \nu_l \bar{\nu}_l & l = e, \mu, \tau \\ u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}, b\bar{b} \end{cases} \quad (15.178a)$$

The Z boson ($m_Z = 91.2$ GeV) cannot decay to a top quark pair because the top quark is too heavy ($m_{\text{top}} = 171.2 \pm 1.2$ GeV). Production of quark pairs is observed as hadrons. The neutrino pairs cannot be observed, but their decay branching ratios can be accurately estimated from

$$\Gamma \left(Z \rightarrow \sum_i \nu_i \bar{\nu}_i \right) = \Gamma_{\text{TOT}} - \Gamma(Z \rightarrow \text{hadrons}) - \Gamma \left(Z \rightarrow \sum_i l_i \bar{l}_i \right) \quad (15.179)$$

where Γ_{TOT} is the total decay width of Z , which can be determined from the cross-section shape (Fig. 15.23). If there is an extra neutrino, the total decay rate Γ_{TOT} changes, hence an observable branching ratio, for instance that to hadrons $\text{BR} = \Gamma(Z \rightarrow \text{hadrons})/\Gamma_{\text{TOT}}$, does, too.

The data agree quite well with the assumption that Z decays to three neutrinos, three charged leptons and five quarks without top but with three color degrees of freedom. Figure 15.23 shows hadronic total cross sections. From the data the number of types of neutrinos is determined to be

$$N_\nu = 2.92 \pm 0.05 \quad (15.180)$$

The above argument does not exclude the possibility that a neutrino with mass greater than $m_Z/2 \simeq 45$ GeV or an exotic neutrino that does not couple to Z (referred to as a “sterile neutrino”) exists. However, the fourth generation, if it exists,

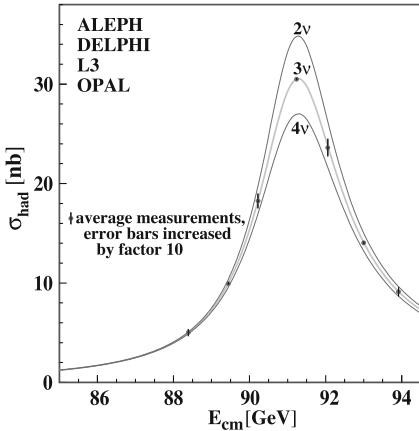


Figure 15.23 Hadronic total cross section vs number of neutrino types. $e^- + e^+ \rightarrow Z \rightarrow \sum q_i \bar{q}_i$ where $q_i = u, d, s, c, b$. The data agree with the Standard Model and the number of neutrino types was determined to be 2.92 ± 0.05 [262].

must include a neutrino with large mass ($> 45 \text{ GeV}$). Considering all the known neutrinos have nearly vanishing mass, it is considered an unlikely situation.

15.8

A Step Toward a Unified Theory

15.8.1

Organizing the Weak Phenomena

In the Standard Model, the weak force is mediated by W bosons, but that was not necessarily taken for granted from the beginning. The notion of the force carrier in the weak interaction was seriously considered only when the near universality of the coupling constant in the muon and the nuclear β decay was realized [256]. With the establishment of the V–A interaction, it became a main current. Let us summarize what we have learned so far and see where the W boson fits in.

The phenomenological weak interaction can be reproduced by the Lagrangian

$$\mathcal{L}_{\text{WI}} = -\frac{4G_F}{\sqrt{2}}(J_L^+)^\mu (J_L^-)_\mu \quad (15.181a)$$

$$(J_L^+)^\mu = (J_L^-)^\mu{}^\dagger = \sum_i \bar{\psi}_{iL} \gamma^\mu \tau^+ \psi_{iL} \quad (15.181b)$$

where τ^+ is the charge-raising operator and not to be confused with the τ lepton.

$$\psi_i = \begin{bmatrix} u \\ d' \end{bmatrix}, \quad \begin{bmatrix} c \\ s' \end{bmatrix}, \quad \begin{bmatrix} t \\ b' \end{bmatrix}, \quad \begin{bmatrix} \nu_e \\ e^- \end{bmatrix}, \quad \begin{bmatrix} \nu_\mu \\ \mu^- \end{bmatrix}, \quad \begin{bmatrix} \nu_\tau \\ \tau^- \end{bmatrix} \quad (15.181c)$$

$$\psi_L = \frac{1}{2}(1 - \gamma^5)\psi \quad (15.181d)$$

Namely, all the weak interactions of the elementary particles have the same common coupling constant G_F and are of V–A type. We can rephrase the term V–A interaction to mean that only left-handed (or more rigorously negative chirality) particles carry the weak charge that generates the force. Then it can be simply stated that the weak force is a universal vector-type interaction. Fermi's intuition was right after all.

At about the same time as Fermi proposed the weak theory, Yukawa proposed the π meson as the nuclear force carrier. He tried to explain the weak force using the same π but was unsuccessful. However, the existence of a force carrier is a far-sighted concept that is carried over to the modern Standard Model. If one considers the weak force to be mediated by a vector boson (W) having a large mass, the Fermi interaction can be reproduced as a second-order process, as we discussed at the very beginning (Sect. 15.2). The characteristics of the weak force having a vector force carrier with universal coupling constant is reminiscent of QED. One naturally wonders whether the weak interaction has the same structure as the electromagnetic force.

The Fermi coupling constant has dimensions of “mass²”. In the old days, the proton mass was used, making the value of $G_F \sim 10^{-5}/m_p^2$ very small. This is the origin of the name “weak interaction”. However, if m_W (~ 80 GeV) is used, the strength of the coupling constant is about the same order as the electromagnetic force. This was one of the first clues to the unified theory. In other words, in the high-energy region, where the typical energy scale is comparable to or larger than m_W , the effect of the weak force would be just as strong as that of the electromagnetic force. The weak force is weak simply because the energy scale E being considered is much smaller than the mass of the force carrier, reducing the effective strength of the force in proportion to E^2/m_W^2 .

Once we accept the existence of W 's, not only can all the weak interaction dynamics be reproduced in an organized way but also the resemblance to QED is quite apparent:

$$\mathcal{L}_{WI} = -\frac{g_W}{\sqrt{2}} \sum_i \left[\bar{\psi}_{iL} \gamma^\mu \tau^+ \psi_{iL} W^+_\mu + h.c. \right] \quad (15.182a)$$

$$\frac{g_W^2}{8m_W^2} = \frac{G_F}{\sqrt{2}} \quad (15.182b)$$

The factor $1/\sqrt{2}$ is conveniently added for later discussions and has no special meaning. In summary, all the weak reactions can be described in a simple and beautiful form by taking into account

1. the quark model
2. V–A interaction
3. $W^\pm$ as the force carrier
4. the Cabibbo rotation

Figure 15.24 describes how various weak reactions can be interpreted by the W mediating processes. The V–A interaction type was determined around 1957. Until then interpretations of β decays were in great confusion, but the discovery of parity violation prompted a rapid convergence within a year. 1957 was an epoch in the history of the weak interaction. A big tide of aspiration for the discovery of the W boson emerged, and from then on many accelerator construction proposals tried to justify themselves by including the potential capability of W production. However, the mass being unknown at that time, whether a proposed accelerator could really produce it remained uncertain. Then finally came the Weinberg–Salam model, which predicted the mass accurately.

The universal V–A theory could be said to be a summary of the pre-Weinberg–Salam era, and it is very powerful phenomenologically. Even today, unless one wants to discuss the direct production of W and related phenomena, it can give satisfactory explanations for not only decay reactions but also neutrino scatterings. The phenomenological success was a big factor in causing the Fermi theory to be taken seriously despite its theoretical shortcomings, which will be described in the following. Similarity with QED, including the universality and vector-type interaction and the fact that the weak force shares part of it with the electromagnetic force

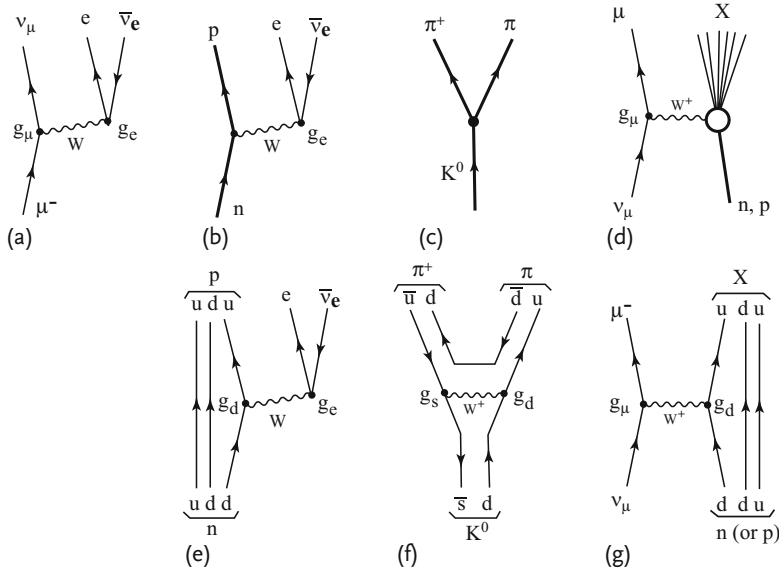


Figure 15.24 Reorganizing hadronic weak interactions in terms of quarks and W . All the interactions are reproduced by a current made of (u, d') (where $d' = d \sin \theta_c + s \sin \theta_c$), (ν_e, e^-) and (ν_μ, μ^-) coupled to $W^\pm$. (a) The muon decay is elementary as it is.

When (b) β decay, (c) $K^0 \rightarrow \pi^+ \pi^-$ and (d) $\nu_\mu n \rightarrow \mu^- X$ (hadrons) are rewritten using the quarks, they are represented by (e), (f) and (g), respectively. The coupling constants are written as $g_d = g_W \cos \theta_c$, $g_s = g_W \sin \theta_c$ and $g_e = g_\mu = g_W$.

as exemplified in weak magnetism, provided crucial clues and a springboard for a unified theory.

15.8.2

Limitations of the Fermi Theory

Although phenomenologically powerful, it is immediately apparent that the four-Fermi V-A theory is not satisfactory from the theoretical point of view. For instance, the scattering cross section of a neutrino by an electron ($\nu_e e^- \rightarrow \nu_e e^-$) can be easily calculated using Eq. (15.181) to give

$$\frac{d\sigma}{d\Omega}(\nu_e e^- \rightarrow \nu_e e^-) = \frac{G_F^2}{4\pi^2} s, \quad \sigma_{\text{TOT}} = \frac{G_F^2}{\pi} s \quad (15.183)$$

The cross section increases proportional to the square of the center of mass energy s . On the other hand, the unitarity restricts the cross section with angular momentum J (unitarity limit)

$$\sigma_{\text{TOT}} \leq (2J + 1)\pi \lambda^2 = (2J + 1)\pi/p^2 = 4\pi(2J + 1)/s \quad (15.184)$$

As the four-fermion Lagrangian is a contact interaction, only the S wave contributes, hence the theory is valid only if

$$\sigma_{\text{TOT}} < \pi \lambda^2 \rightarrow s \lesssim 2\pi/G_F \rightarrow p \lesssim 300 \text{ GeV} \quad (15.185)$$

One might think the introduction of W may save the situation. However, in the calculation of higher order corrections, the W boson contributions to the intermediate state have to be added or are integrated over the energy. In the limit $p \rightarrow \infty$, all the masses can be neglected and the most divergent term in the correction generally has the form (see for instance Eq. (8.2))

$$\sim G_F \int f(p) \frac{d^4 p}{p^4} \quad (15.186)$$

In the case of the weak force, the coupling constant G_F has dimension E^{-2} , which makes $f(p) \propto p^2$ and the integral diverges badly. The divergence at $p \rightarrow \infty$ is the inevitable fate of quantum field theory, but there are two kinds of divergences, good and bad. The four-Fermi interaction is a typical example of bad divergence. Good divergence means that it can be removed by introducing a finite number of correction terms (renormalizable theory) and gives correct results if treated with the right recipe, as was clarified in QED. We saw that the anomalous magnetic moment of the electron and muon was reproduced to an accuracy of 10^{-10} and that the applicability of QED extends to the energy region as high as a few TeV. This fact provides a solid foundation for the renormalizable theory that, despite its shortcomings, is very close to the true theory. This is why QED is always considered as a model in reaching the right theory. In QED, the coupling constant corresponding to G_F is dimensionless. If one chooses QED as a model, the coupling constant has to be dimensionless, too. Therefore introduction of W seems the right step for remedying the problem.

Introduction of W Let us introduce W and see that g_W in the Lagrangian Eq. (15.182) is indeed dimensionless. Let us consider $\nu_e + e^- \rightarrow \nu_e + e^-$ again. The interaction Lagrangian has a similar form to the QED Lagrangian $\mathcal{L}_{\text{QED}} = -j^\mu A_\mu$:

$$\mathcal{L}_{\text{WI}} = -\frac{g_W}{\sqrt{2}} (\bar{\nu}_L \gamma^\mu e_L W_\mu^+ + \bar{e}_L \gamma^\mu \nu_L W_\mu^-) \quad (15.187)$$

That g_W is dimensionless can be seen as follows. Pick up the mass term in the Lagrangian of the vector field [see Eqs. (5.20d) and (5.22a)] and consider its physical dimension:

$$\left[\int d^3 x \mathcal{L}_W \right] = \left[\int d^3 x m \bar{\psi} \psi \right] = \left[\int d^3 x m_W^2 W_\mu W^\mu \right] = [E] \quad (15.188)$$

This gives $[\mathcal{L}] = [E^4]$, $[\psi] = [E]^{3/2}$, $[W_\mu] = [E]$, and comparing the dimensions of the two sides of Eq. (15.187) proves g_W is dimensionless.

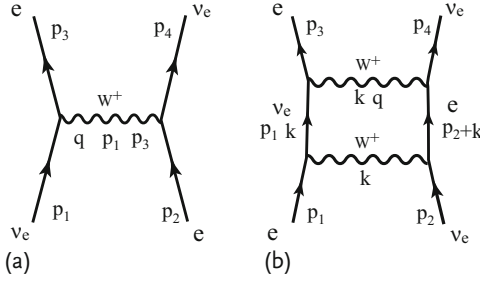


Figure 15.25 $\nu_e e^-$ scattering diagram. (a) The lowest order. (b) Higher order W exchange.

Now we can investigate the effect of W exchange in ν_e scattering as shown in Fig. 15.25. The matrix element of Fig. 15.25a is given by

$$\begin{aligned}
 -i T_{fi} = & \left[\bar{e} \left(-i \frac{g_W}{\sqrt{2}} \gamma^\mu \right) \frac{(1 - \gamma^5)}{2} \nu_e \right] \frac{-i (g_{\mu\nu} - q_\mu q_\nu / m_W^2)}{q^2 - m_W^2} \\
 & \times \left[\bar{\nu}_e \left(-i \frac{g_W}{\sqrt{2}} \gamma^\nu \right) \frac{(1 - \gamma^5)}{2} e \right]
 \end{aligned} \quad (15.189)$$

The differential cross section becomes

$$\frac{d\sigma}{d\Omega} = \frac{G_F^2 s}{4\pi^2} \left(\frac{m_W^2}{m_W^2 - q^2} \right)^2 = \frac{(G_F^2 / 4\pi^2) s}{[1 + (s/m_W^2)(1 - \cos \theta)]^2} \quad (15.190)$$

Integration over the solid angle gives

$$\sigma_{\text{TOT}}(\nu_e e^- \rightarrow \nu_e e^-) = \frac{G_F^2 s}{\pi} \frac{1}{1 + s/m_W^2} \quad (15.191)$$

$\sigma_{\text{TOT}} \rightarrow \text{constant}$, which is a great improvement compared to linear divergence in Eq. (15.183).

However, if we calculate the S-wave scattering amplitude f_0

$$\begin{aligned}
 f_0 &= \frac{G_F s}{4\pi} \int d\cos \theta \frac{1}{1 + (s/m_W^2)(1 - \cos \theta)} \\
 &= \frac{G_F m_W^2}{2\pi} \ln \left(1 + \frac{s}{m_W^2} \right)
 \end{aligned} \quad (15.192)$$

we find this diverges logarithmically and the problem is not solved, at least not in the lowest order. What about higher order contributions? If we consider a diagram that contains two W 's in the intermediate states like the one in Fig. 15.25b, we need to add those for all polarization states of W as well as the internal momentum k . As the Feynman propagator of the W boson is given by

$$\Delta_F(q) \sim \frac{\sum_\lambda \epsilon^\mu(\lambda) \epsilon^\nu(\lambda)^*}{q^2 - m_W^2} = \frac{-g^{\mu\nu} + q^\mu q^\nu / m_W^2}{q^2 - m_W^2} \quad (15.193)$$

the contribution of Fig. 15.25b contains an integral of the form

$$\begin{aligned} & \sim g_W^4 \int d^4 k \Delta_{F\rho\sigma}(k-q) \Delta_{F\mu\nu}(k) \\ & \quad \times [\bar{u}(p_3) \gamma^\rho S_F(p_1-k) \gamma^\mu u(p_1)] [\bar{u}(p_4) \gamma^\sigma S_F(p_2+k) \gamma^\nu u(p_2)] \quad (15.194) \\ S_F(p) &= \frac{1}{(2\pi)^4} \int d^4 p e^{-ip \cdot x} \frac{\not{p} + m}{p^2 - m^2 + i\varepsilon} \end{aligned}$$

If we pick up the dominant contribution at high energy ($k \rightarrow \infty$), it becomes

$$\begin{aligned} & \sim g_W^4 \int d^4 k \frac{k^\rho k^\sigma / m_W^2}{k^2} \frac{k^\mu k^\nu / m_W^2}{k^2} \frac{E \not{k}}{k^2} \frac{E \not{k}}{k^2} \\ & \sim \left(\frac{g_W^2 E^2}{m_W^2} \right) \frac{g_W^2}{m_W^2} \int \frac{d^4 k}{k^4} O(k^2) \quad (15.195) \end{aligned}$$

where E^2 arises from the normalization of the wave function and $(g_W^2 E^2 / m_W^2)$ can be taken away from the divergence consideration. The last expression corresponds to $f(p) \sim p^2$ in Eq. (15.186). One notices that the divergent term is generated by $q^\mu q^\nu$ in the numerator of the W propagator.

Origin of Divergences The origin of the divergent term is easy to understand. A massive vector boson has three degrees of freedom in the spin components. The polarization vector $\epsilon^\mu(\lambda)$, ($\lambda = 1, \dots, 3$) in its rest frame can be taken as

$$\epsilon^\mu(1) = (0, 1, 0, 0) \quad (15.196a)$$

$$\epsilon^\mu(2) = (0, 0, 1, 0) \quad (15.196b)$$

$$\epsilon^\mu(3) = (0, 0, 0, 1) \quad (15.196c)$$

Setting the z -axis along the momentum direction of the W , and Lorentz boosting the particle to where W has $p_W^\mu = (\omega, 0, 0, q)$, we find the transverse polarizations do not change but the longitudinal polarization is transformed to

$$\epsilon^\mu(3) = (q/m_W, 0, 0, \omega/m_W) \quad (15.197)$$

We realize that the second term in the numerator of the propagator (15.193) comes from the longitudinal polarization. This is equivalent to introducing the dimension $[E]^{-2}$ in the coupling constant and induces a divergent integral. When the mass of the force carrier is zero, it has only transverse polarization and the divergent term is not generated.

Furthermore, gauge invariance, which forces the mass of the force carrier to vanish, also guarantees the reduction of the divergent terms [see Eq. (8.51) and associated arguments]. It was QED's lesson that a gauge-invariant theory has, after all, at most logarithmic divergences and they can be subtracted by renormalizing the coupling constant and the mass.

15.8.3

Introduction of $SU(2)$

If one wants to apply gauge theory to W 's, one has to deal with the difference between the W boson and the photon. This is twofold. First, the W boson is charged, and second, it has a finite and large mass. In order to solve the problems one at a time, we forget the finite mass for the moment and see if gauge theory can be applied to the charged force carrier. Let us consider what kind of symmetry it can have. Defining a column vector

$$\Psi_L = \begin{bmatrix} \nu_L \\ e_L^- \end{bmatrix} \quad (15.198)$$

and using

$$\tau_{\pm} = \frac{1}{2}(\tau_1 \pm i\tau_2) \quad (15.199)$$

where $\tau_i, i = 1, \dots, 3$ are the Pauli matrices, we can reformulate the Lagrangian (15.187) as

$$\begin{aligned} \mathcal{L}_{W1} &= -\frac{g_w}{\sqrt{2}} \bar{\Psi}_L \gamma^\mu (\tau_+ W_\mu^+ + \tau_- W_\mu^-) \Psi_L \\ &= -\frac{g_w}{2} \bar{\Psi}_L \gamma^\mu (\tau_1 W_{1\mu} + \tau_2 W_{2\mu}) \Psi_L \\ W^\pm &= \frac{W_1 \mp iW_2}{\sqrt{2}} \end{aligned} \quad (15.200)$$

The form reminds one of the isospin-independent nuclear force:

$$\begin{aligned} \mathcal{L}_N &= -i \frac{g_N}{2} \bar{\Psi}_N \gamma^5 (\boldsymbol{\tau} \cdot \boldsymbol{\pi}) \Psi_N \\ &= -i g_N \bar{\Psi}_N \left[\frac{\tau^+ \pi^+ + \tau^- \pi^-}{\sqrt{2}} + \frac{1}{2} \tau_3 \pi^0 \right] \Psi_N \\ \Psi_N &= \begin{bmatrix} p \\ n \end{bmatrix}, \quad \boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3) \end{aligned} \quad (15.201)$$

Apart from the space-time structure (not vector but pseudoscalar), it has almost identical structure to Eq. (15.200) if we replace

$$\Psi_N \rightarrow \Psi_L, \quad \pi \rightarrow W \quad (15.202)$$

Here, we need to introduce a neutral W^0 to make them symmetrical and add $(\tau_3/2) W_3^\mu$ to terms inside the parentheses on the second line of Eq. (15.200). The nuclear force was isospin independent or invariant under exchange of p and n . This is $SU(2)$ symmetry, to which the gauge theory of QED was generalized by Yang and Mills in 1954 [396]. Therefore, requiring the weak interaction to be invariant under

exchange of $(\nu_e \leftrightarrow e^-)$ or the “weak isospin independence”, we postulate the weak Lagrangian has the form

$$\mathcal{L}_{\text{WI}} = -\frac{g_W}{2} \bar{\Psi}_L \gamma^\mu (\boldsymbol{\tau} \cdot \mathbf{W}_\mu) \Psi_L \quad (15.203)$$

The $SU(2)$ symmetry assumption is to require the existence of the third member, a neutral W^0 boson. For comparison’s sake let us write down the electromagnetic Lagrangian for the (ν_e, e^-) pair:

$$\mathcal{L}_{\text{EM}} = e(\bar{e} \gamma^\mu e) = \frac{e}{2} \bar{\Psi} \gamma^\mu (1 - \tau_3) \Psi A_\mu, \quad \Psi = \begin{bmatrix} \nu_e \\ e \end{bmatrix} \quad (15.204)$$

Comparing Eq. (15.204) with the weak current we see immediately that the isospin current

$$j_i^\mu \sim \bar{\Psi} \gamma^\mu \tau_i \Psi \quad (15.205)$$

appears in both the weak and the electromagnetic current. In fact, we have already seen that the vector part of the weak current and the electromagnetic current is identical apart from the coupling constant. It is a consequence of the CVC hypothesis and has been experimentally confirmed. Thus, the weak current not only resembles the electromagnetic current but also owns a part of it jointly. This strongly suggests that the weak interaction is also governed by a gauge theory and is a part of a larger symmetry group containing both the weak force and electromagnetism. Conjecture had reached this point when the V–A interaction was determined in 1957. Glashow even conceived a model to unify the weak and electromagnetic interactions [173]. However, the finite mass of W , which apparently violates gauge invariance, was the obstacle to going further. To reach the electroweak unified model, deep insight to the theory of gauge invariance is necessary. We shall discuss gauge theory and the electroweak theory in Chap. 18.

16

Neutral Kaons and CP Violation*

Symmetry is a main theme of modern particle physics. The discovery of parity violation was an epoch-making event in its history and we have learned in the previous chapter that it stems from the different action of the weak force on left- and right-handed particles. Pursuit of its origin has led us to chiral gauge symmetry, which is the foundation of modern particle theory. The discovery of CP violation in 1964 was also unexpected, and despite 50 years' effort there has been little progress in understanding its true nature. Nevertheless, we have ample reason to believe that elucidating its origin would establish yet another paradigm of particle physics, because recently it was confirmed that the framework in which CP violation is at work requires (at least) three generations of quarks and leptons, whose existence remains a mystery referred to as the flavor problem (or family or generation problem). It is also known that CP violation played a crucial role in the early universe in developing from the hot big bang chaos to the present form of a matter-dominant universe. There is a good chance that pursuit of CP violation will lead to true understanding of the origin of the universe and uncover the mystery of flavors, about which the Standard Model has little to offer.

CP violation is built into the Standard Model within the framework of the Kobayashi–Maskawa (KM) matrix, but its dynamical origin remains to be clarified. It is also suggested that although CP violation as observed experimentally can be fully explained by the KM framework, it is not enough to explain the matter–antimatter asymmetry of the universe. In this sense “CP violation” can be considered as an exceptional topic in the Standard Model. Needless to say, concentrating on a small anomaly has often provided a chance to open a door to a new view, and, in the author's personal opinion, understanding the vacuum issues (also known as the Higgs mechanism and dark energy) and CP violation are the two key factors to a new physics beyond the Standard Model. In this chapter, we discuss neutral kaon processes, in which the CP violation was discovered and which are the most investigated. Although we do not touch phenomena involving B decays, the neutral kaon provides sensitive tests for other fundamental questions, including T, CPT invariance and quantum gravity, and deserves a special chapter of its own.

16.1

Introduction

16.1.1

Strangeness Eigenstates and CP Eigenstates

There are two kinds of neutral kaons, K^0 and $\bar{K}^0$, which have strangeness $S = +1$ and $S = -1$, respectively. S is a conserved quantum number in the strong interaction, as can be seen in the reactions with particles that have known strangeness ($S_\pi = S_p = S_n = 0$, $S_\Lambda = -1$):

$$\begin{array}{rclcl}
 \pi^- (d\bar{u}) & + & p (uud) & \rightarrow & K^0 (d\bar{s}) + \Lambda (uds) & (a) \\
 S \quad 0 & & 0 & & +1 \quad -1 \\
 \pi^- (d\bar{u}) & + & p (uud) & \rightarrow & K^0 (d\bar{s}) + \bar{K}^0 (s\bar{d}) + n (udd) & (b) \\
 S \quad 0 & & 0 & & +1 \quad -1 \quad 0
 \end{array} \tag{16.1a}$$

$$\begin{array}{rclcl}
 K^0 (d\bar{s}) & + & p (uud) & \rightarrow & K^+ (u\bar{s}) + n (udd) & (c) \\
 S \quad +1 & & 0 & & +1 \quad 0 \\
 \bar{K}^0 (s\bar{d}) & + & n (udd) & \rightarrow & K^- (s\bar{u}) + p (uud) & (d) \\
 S \quad -1 & & 0 & & -1 \quad 0 \\
 \bar{K}^0 (s\bar{d}) & + & p (uud) & \rightarrow & \pi^+ (u\bar{d}) + \Lambda (uds) & (e) \\
 S \quad -1 & & 0 & & 0 \quad -1 \\
 \bar{K}^0 (s\bar{d}) & + & n (udd) & \rightarrow & \pi^- (d\bar{u}) + \Sigma^+ (uus) & (f) \\
 S \quad -1 & & 0 & & 0 \quad -1
 \end{array} \tag{16.1b}$$

$K^0 (d\bar{s}, S = +1)$ pairs with $K^+ (u\bar{s})$ and $\bar{K}^0 (s\bar{d}, S = -1)$ pairs with $K^- (s\bar{u})$ to make isospin doublets that are antiparticles of each other:

$$\begin{bmatrix} K^+ \\ K^0 \end{bmatrix}, \quad \begin{bmatrix} -\bar{K}^0 \\ K^- \end{bmatrix} \tag{16.2}$$

The above strangeness-conserving reactions and their doublet structure are obvious if one analyzes them in terms of the quark model, taking note of the fact that the strong interaction does not change a quark's flavor.

As to decay processes, they have common decay modes including $2\pi, 3\pi$ and $\pi^\pm l^\mp \nu$ and cannot be distinguished by strangeness. This reflects the simple fact that it is not a good quantum number in the weak interaction. On the other hand, since the discovery of P as well as C violation, the combined operation CP has been considered to be conserved in the weak interaction. We shall come back to CP violation later, but let us discuss processes first assuming it is conserved. We define a CP operator by

$$\begin{aligned}
 \text{CP} |K^0\rangle &= e^{i\xi_{\text{CP}}} |\bar{K}^0\rangle = \eta_{\text{CP}} |\bar{K}^0\rangle, \\
 \text{CP} |\bar{K}^0\rangle &= e^{-i\xi_{\text{CP}}} |K^0\rangle = \eta_{\text{CP}}^* |K^0\rangle,
 \end{aligned} \tag{16.3}$$

It is conventional to set the phase $\xi_{\text{CP}} = 0$ and define CP eigenstates by

$$|K_1\rangle = (|K^0\rangle + |\bar{K}^0\rangle)/\sqrt{2} \quad (16.4a)$$

$$|K_2\rangle = (|K^0\rangle - |\bar{K}^0\rangle)/\sqrt{2} \quad (16.4b)$$

Then $\text{CP}|K_1\rangle = |K_1\rangle$, $\text{CP}|K_2\rangle = -|K_2\rangle$ and K_1, K_2 can have different decay modes. The 2π and 3π systems to which the K meson can decay have CP values +1 and -1, respectively, hence the decay modes are separated as $K_1 \rightarrow 2\pi$, $K_2 \rightarrow 3\pi$.

(1) 2π : Remembering $P\pi = -\pi$, $P(2\pi) = (-1)^l$ and the C operation exchanges $\pi^+ \leftrightarrow \pi^-$, we find Bose statistics dictates that C and P operations give the same sign. Therefore $\text{CP}(\pi^+\pi^-) = (-1)^{2l} = +1$, where l is the orbital angular momentum. As $\text{CP}\pi^0 = \pi^0$, Bose statistics forces l of the $2\pi^0$ system to be even and $\text{CP}(\pi^0\pi^0) = +1$.

(2) 3π : For $\pi^+\pi^-\pi^0$, denoting the orbital angular momentum of $\pi^+\pi^-$ as l and that of π^0 relative to the center of mass of $\pi^+\pi^-$ as L , we have $\text{CP}(\pi^+\pi^-\pi^0) = (-1)^{2l+L+3}$. As the spin of the K meson is 0, $l = L$. Because of the angular momentum barrier (see Sect. 13.5.3), the dominant component is $\text{CP} = -1$.¹⁾ For $3\pi^0$, the same argument as for $2\pi^0$ makes $l = \text{even}$ and $\text{CP}(3\pi^0) = -1$.

The above argument shows that K_1 decays to 2π and K_2 to 3π . As the Q -value ($m_K - \sum m_\pi$) of $K_1 \rightarrow 2\pi$ is much larger than that of $K_2 \rightarrow 3\pi$, the $K_1 \rightarrow 2\pi$ decay rate is much larger than that of $K_2 \rightarrow 3\pi$. Namely, K_1 is short lived compared to K_2 .

16.1.2

Schrödinger Equation for $K^0 - \bar{K}^0$ States

Both K^0 and $\bar{K}^0$ share common decay modes 2π and 3π . This means a transition $K^0 \leftrightarrow (2\pi, 3\pi) \leftrightarrow \bar{K}^0$ or mixing is possible and $\langle K^0 | H | \bar{K}^0 \rangle \neq 0$. The existence of the mixing means that K^0 can change to $\bar{K}^0$ in vacuum. Time development of the wave function of the $K^0, \bar{K}^0$ system can be written as

$$i \frac{\partial}{\partial t} \psi(t) = H \psi(t) \quad (16.5a)$$

$$\psi(t) = a(t)|K^0\rangle + b(t)|\bar{K}^0\rangle + \sum_n c_n(t)|m\rangle \quad (16.5b)$$

Here, $|m\rangle$ is any other state than K^0 and $\bar{K}^0$ into which $K^0, \bar{K}^0$ can decay. The Hamiltonian H is hermitian and contains the strong interaction as well as the weak interaction, which induces mixing. However, it is very difficult to solve the equation in general. It is convenient to consider only two states $K^0, \bar{K}^0$ and introduce a mass operator $\mathcal{A}$ such that the two-component wave function obeys the Schrödinger-type equation

$$i \frac{\partial \psi(t)}{\partial t} = \mathcal{A} \psi(t), \quad \psi(t) = \alpha(t)|K^0\rangle + \beta(t)|\bar{K}^0\rangle \quad (16.6)$$

1) Because of this $K_1 \rightarrow 3\pi$ can occur (albeit rarely), but $K_2 \not\rightarrow 2\pi$.

Any transition to other states ($m \neq K^0, \bar{K}^0$) is considered as K^0 or $\bar{K}^0$ lost from this two-component system. Then the probability amplitude in the kaon rest frame must decrease with time. The conservation of the probability (or unitarity) means

$$\frac{\partial \rho}{\partial t} = \frac{\partial(\psi^\dagger \psi)}{\partial t} = -i\psi^\dagger (\mathcal{A} - \mathcal{A}^\dagger) \psi \equiv -\psi^\dagger(t) \Gamma \psi(t) \quad (16.7)$$

This means $\mathcal{A}$ can be expressed in terms of two hermitian operators M and Γ

$$\mathcal{A} = \frac{\mathcal{A} + \mathcal{A}^\dagger}{2} + \frac{\mathcal{A} - \mathcal{A}^\dagger}{2} \equiv M - \frac{i}{2} \Gamma, \quad M^\dagger = M, \quad \Gamma^\dagger = \Gamma \quad (16.8)$$

Then the time evolution of the eigenstate of $\mathcal{A}$ can be expressed as

$$e^{-i\mathcal{A}t} |K\rangle \sim \exp \left[-im_K t - \frac{\Gamma}{2} t \right] |K\rangle \quad (16.9)$$

Multiplying $\langle K^0|, \langle \bar{K}^0|$ from left and noting that $|K^0\rangle$ and $|\bar{K}^0\rangle$ are orthogonal, Eq. (16.6) can be cast in matrix form

$$i \frac{\partial}{\partial t} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} = \begin{bmatrix} \langle K^0 | \mathcal{A} | K^0 \rangle & \langle K^0 | \mathcal{A} | \bar{K}^0 \rangle \\ \langle \bar{K}^0 | \mathcal{A} | K^0 \rangle & \langle \bar{K}^0 | \mathcal{A} | \bar{K}^0 \rangle \end{bmatrix} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} \quad (16.10)$$

Note the mass matrix $\mathcal{A}$ has complex eigenvalues and is not a hermitian operator. The mass matrix can be expressed as (see Appendix H)

$$\mathcal{A}_{ij} = M_{ij} - i \frac{\Gamma_{ij}}{2} \quad (16.11a)$$

$$M_{ij} = m_K \delta_{ij} + \langle i | H_W | j \rangle + \sum_{n \neq K^0, \bar{K}^0} P \frac{\langle i | H_W | n \rangle \langle n | H_W | j \rangle}{m_j - E_n} \quad (16.11b)$$

$$\Gamma_{ij} = 2\pi \sum_{n \neq K^0, \bar{K}^0} \delta(m_j - E_n) \langle i | H_W | n \rangle \langle n | H_W | j \rangle \quad (16.11c)$$

where i, j denote K^0 or $\bar{K}^0$ and P means principal value. The intermediate states n is any state that K^0 or $\bar{K}^0$ can transition through the weak interaction. The above formula is referred to as the Weisskopf–Wigner approximation. In this approximation weak interactions between final states are neglected and it is valid for time t large compared with the strong interaction scale $t_{\text{st}} \sim 1/m_K \sim 10^{-23 \sim 24}$ s. Equations (16.11), except for the first term in (b), look like the second-order transition matrix Eq. (6.37) introduced in Chap. 6. However, there are some differences. Here, the sum over intermediate states is taken only over continuum states to which $K^0, \bar{K}^0$ can decay and does not include $K^0, \bar{K}^0$ themselves. In the transition matrix Eq. (6.37), the intermediate states constitute a complete set and the formula is valid for time t much greater than t_{st} .

Diagonalization of Eq. (16.10) changes (α, β) to (α', β') , which is a representation in the bases $|K_1\rangle, |K_2\rangle$

$$\psi = \alpha'(t)|K_1\rangle + \beta'(t)|K_2\rangle = \alpha(t)|K^0\rangle + \beta(t)|\bar{K}^0\rangle \quad (16.12a)$$

$$i \frac{\partial}{\partial t} \begin{bmatrix} \alpha' \\ \beta' \end{bmatrix} = A_{\text{dia}} \begin{bmatrix} \alpha' \\ \beta' \end{bmatrix} = \begin{bmatrix} \lambda_S & 0 \\ 0 & \lambda_L \end{bmatrix} \begin{bmatrix} \alpha' \\ \beta' \end{bmatrix}, \quad (16.12b)$$

$$\lambda_S = m_S - \frac{i}{2}\Gamma_S \quad \lambda_L = m_L - \frac{i}{2}\Gamma_L \quad (16.12c)$$

We use $K_{1,2}$ to denote CP = ± 1 eigenstates, but have attached subscripts S, L to denote the eigenvalues. The nomenclature $K_{S,L}$ is saved for the weak decay eigenstates including CP violation effects that appear later. Note λ is a complex number but $m_{S,L}, \Gamma_{S,L}$ are real numbers. Representations in the $K^0, \bar{K}^0$ and K_1, K_2 bases are related by

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = U \begin{bmatrix} \alpha' \\ \beta' \end{bmatrix}, \quad U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (16.13a)$$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = U A_{\text{dia}} U^{-1} = \begin{bmatrix} M & -\Delta_\lambda/2 \\ -\Delta_\lambda/2 & M \end{bmatrix} \quad (16.13b)$$

$$M = (\lambda_L + \lambda_S)/2, \quad \Delta_\lambda = \lambda_L - \lambda_S \quad (16.13c)$$

The eigenvalues are expressed in terms of the original mass matrices as

$$\lambda_L = \frac{A_{11} + A_{22}}{2} \pm \sqrt{\left(\frac{A_{11} - A_{22}}{2}\right)^2 + A_{12}A_{21}} \equiv M \mp \Delta_\lambda/2 \quad (16.14a)$$

$$\text{Tr}[A] = \lambda_S + \lambda_L = A_{11} + A_{22} \quad (16.14b)$$

$$\det[A] = \lambda_S \lambda_L = A_{11}A_{22} - A_{12}A_{21} \quad (16.14c)$$

The symmetry requirements constrain the matrix elements, which will be discussed in detail in Sect. 16.2.1:

$$A_{11} = A_{22} = M, \quad \Delta_\lambda = -2\sqrt{A_{12}A_{21}} \quad \text{CPT} \quad (16.15a)$$

$$|A_{12}| = |A_{21}| \quad \text{T} \quad (16.15b)$$

$$A_{11} = A_{22} = M \quad \text{and} \quad |A_{12}| = |A_{21}| \quad \text{CP} \quad (16.15c)$$

Note that for off-diagonal elements (A_{12}, A_{21}), only the absolute magnitude can be constrained. This is because we are free to change their phase,

$$|K^0\rangle \rightarrow e^{-i\xi}|K^0\rangle, \quad |\bar{K}^0\rangle \rightarrow e^{i\xi}|\bar{K}^0\rangle \quad (16.16)$$

which originates from strangeness conservation in the strong interaction. Or rather, the freedom to change the phase is the origin of the strangeness con-

servation (Noether's theorem).²⁾ By rephasing the elements $|\bar{K}^0\rangle \rightarrow e^{i\xi}|\bar{K}^0\rangle$, $|K^0\rangle \rightarrow e^{-i\xi}|K^0\rangle$,

$$A_{12} \rightarrow A_{12}e^{2i\xi}, \quad A_{21} \rightarrow A_{21}e^{-2i\xi} \quad (16.17)$$

Observable quantities do not depend on the phase convention. Note, however, that once one has fixed it one has to stick to it, otherwise contradictions will ensue. Referring to Eq. (16.11), we see that Eq. (16.15a) means the mass and lifetime of a particle and its antiparticle are the same.

Problem 16.1

Derive Eq. (16.13) and Eq. (16.14).

16.1.3

Strangeness Oscillation

Strangeness oscillation is a phenomenon that K^0 and $\bar{K}^0$ states transition back and forth. Suppose we have a pure $|K^0\rangle$ beam at $t = 0$. It can be made in several ways:

1. By using the difference of the energy threshold between reactions (a) ($E_{\text{th}} = 0.91$ GeV) and (b) ($E_{\text{th}} = 1.5$ GeV) in Eq. (16.1a) and setting the π energy between 0.91 and 1.5 GeV.
2. By stopping a high-flux low-energy antiproton beam in a pressurized hydrogen gas target and selecting the following reactions:

$$\bar{p} + p \rightarrow \begin{cases} K^- \pi^+ K^0 \\ K^+ \pi^- \bar{K}^0 \end{cases} \quad (16.18)$$

The CPLEAR experiment at CERN [109] used a beam with momentum set at $p = 200$ MeV/c and obtained pure K^0 ($\bar{K}^0$) by tagging $K^- \pi^+$ ($K^+ \pi^-$).

3. By using a ϕ factory. An electron-positron collider with total center of mass energy $\sqrt{s} = m_\phi = 1019.4$ MeV called DAPHNE of INFN (Istituto Nazionale di Fisica Nucleare) at Frascati, Italy [237] can produce a coherent $K^0 - \bar{K}^0$ state with $J^{PC} = 1^{--}$ by

$$e^- + e^+ \rightarrow \phi, \quad \phi \rightarrow K^0 \bar{K}^0 \quad (16.19)$$

If both CPT and CP invariance hold, K_1, K_2 are the mass eigenstates. Setting $|K^0(t=0)\rangle = |K^0\rangle$, its time development can be described as

$$\begin{aligned} |K^0(t)\rangle &= [|K_1(t)\rangle + |K_2(t)\rangle]/\sqrt{2} = [|K_1\rangle e^{-\lambda_S t} + |K_2\rangle e^{-\lambda_L t}]/\sqrt{2} \\ &= g_+(t)|K^0\rangle - g_-(t)|\bar{K}^0\rangle \end{aligned} \quad (16.20a)$$

$$g_\pm(t) = \frac{1}{2} \left[e^{-\lambda_L t} \pm e^{-\lambda_S t} \right] \quad (16.20b)$$

- 2) Rephasing would change the CP phase $\eta_{CP} \rightarrow \eta_{CP} e^{2i\xi}$ defined in Eq. (16.3) for consistency, but it would not change any conclusion related to CP phase.

which shows that a $\bar{K}^0$ component appears in the originally pure K^0 beam as time passes. The probability of detecting K^0 (or $\bar{K}^0$) at time t is given by

$$P(K^0 \rightarrow K^0) = |\langle K^0 | K^0(t) \rangle|^2 = |g_+(t)|^2 \quad (16.21a)$$

$$P(K^0 \rightarrow \bar{K}^0) = |\langle \bar{K}^0 | K^0(t) \rangle|^2 = |g_-(t)|^2 \quad (16.21b)$$

$$\begin{aligned} |g_{\pm}|^2 &= \frac{1}{4} \left[e^{-\Gamma_S t} + e^{-\Gamma_L t} \pm 2e^{-\bar{\Gamma} t} \cos \Delta m t \right] \\ &= \frac{e^{-\bar{\Gamma} t}}{2} \left[\cosh \frac{\Delta \Gamma}{2} t \pm \cos \Delta m t \right] \end{aligned} \quad (16.21c)$$

$$\bar{\Gamma} = \frac{1}{2} (\Gamma_S + \Gamma_L) \simeq \frac{\Gamma_S}{2}, \quad \Delta \Gamma = \Gamma_S - \Gamma_L \simeq \Gamma_S, \quad \Delta m = m_L - m_S, \quad (16.21d)$$

Note, if CP invariance is assumed, $P(K^0 \rightarrow K^0) = P(\bar{K}^0 \rightarrow \bar{K}^0)$ and $P(K^0 \rightarrow \bar{K}^0) = P(\bar{K}^0 \rightarrow K^0)$. Equations (16.21) mean components of K^0 and $\bar{K}^0$ oscillate as a function of time. If $\Delta m \gg \bar{\Gamma} \simeq \Gamma_S$, the oscillation is too fast to detect K^0 and $\bar{K}^0$ separately. However, if $\Delta m \simeq \Gamma_S$, the oscillation can be observed. Observations, show $\Delta m \simeq \Gamma_S/2$. In Fig. 16.1, however, an oscillation with $\Delta m = \Delta \Gamma$ is illustrated to emphasize the oscillation feature.

The right-hand panel of Fig. 16.1 shows asymmetry data of CPLEAR [107] defined as

$$\begin{aligned} A_{\Delta m}(\tau) &= \frac{\{P(K^0 \rightarrow K^0) + P(\bar{K}^0 \rightarrow \bar{K}^0)\} - \{P(K^0 \rightarrow \bar{K}^0) + P(\bar{K}^0 \rightarrow K^0)\}}{\{P(K^0 \rightarrow K^0) + P(\bar{K}^0 \rightarrow \bar{K}^0)\} + \{P(K^0 \rightarrow \bar{K}^0) + P(\bar{K}^0 \rightarrow K^0)\}} \\ &= \frac{|g_+(t)|^2 - |g_-(t)|^2}{|g_+(t)|^2 + |g_-(t)|^2} = \frac{\cos \Delta m \tau}{\cosh \Delta \Gamma \tau / 2} \end{aligned} \quad (16.22)$$

where τ is the proper time of the kaon. From the observed curve, Δm can be determined and is given by

$$\begin{aligned} \Delta m = m_L - m_S &= 3.485 \pm 0.01 \times 10^{-6} \text{ eV} \\ &= 0.4738 \pm 0.0013 \Gamma_S \end{aligned} \quad (16.23)$$

The sign of Δm was determined using the regeneration process described in the next section.

In order to observe the strangeness oscillation, K^0 and $\bar{K}^0$ have to be identified as a function of time. Since they are eigenstates of the strong interaction, the observation must be done in principle not by their decay mode but by their collision mode in the strong interaction. Notice, however, that the subsequent decay mode has to be identified to confirm what reaction has occurred. For example, in a hydrogen

- 3) Their measurement is sufficiently accurate that CPT- and T-violating terms have to be considered. Then $P(K^0 \rightarrow K^0) \neq P(\bar{K}^0 \rightarrow \bar{K}^0)$, $P(K^0 \rightarrow \bar{K}^0) \neq P(\bar{K}^0 \rightarrow K^0)$. However, by taking the sum, both CPT- and T-violating terms are cancelled and give the same expressions as Eqs. (16.21).

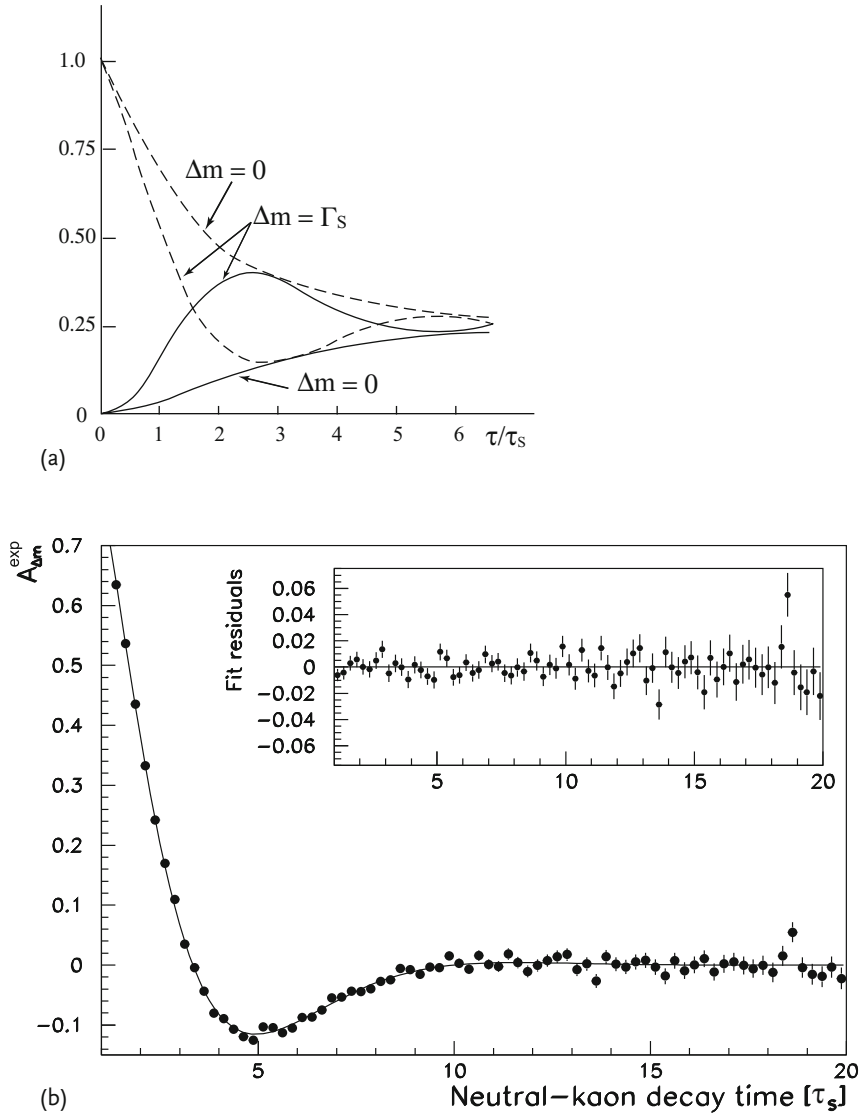


Figure 16.1 (a) Illustration of the strangeness oscillation. The intensity of a pure K^0 beam at $t = 0$ decreases as time passes (dashed line) and that of $\bar{K}^0$, which did not exist at $t = 0$, appears. If $\Delta m \neq 0$, the two curves oscillate. Here, Δm is set equal to Γ to emphasize the

oscillation behavior. (b) Asymmetry $A_{\Delta m}(\tau)$ data of CPLEAR of strangeness-conserving and strangeness-changing components. The strangeness was tagged by the sign of the decay leptons [107].

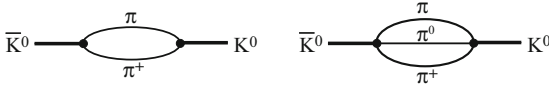


Figure 16.2 $K^0 - \bar{K}^0$ mixing through 2π , 3π intermediate states.

bubble chamber, that a K^0 was produced by the reaction $\pi^- p \rightarrow K^0 \Lambda$ at $t = 0$ can be confirmed by detecting subsequent decay of $\Lambda \rightarrow \pi^- p$ ($\tau \sim 2.63 \times 10^{-10}$ s). $\bar{K}^0$ at time t can be detected, for example, by observing the reaction $\bar{K}^0 p \rightarrow \pi^+ \Lambda$, which leaves a visible π^+ track, but must be confirmed by its associated Λ decay. As the momentum and direction of K^0 can be reconstructed from π^+ and Λ , the flight distance of $\bar{K}$ can be converted to its proper time. Other usable reactions are

$$\begin{aligned} K^0 + p &\rightarrow K^+ + n, \quad \bar{K}^0 + n \rightarrow K^- + n, \\ \bar{K}^0 + n &\rightarrow \pi^0 + \Lambda (\Lambda \rightarrow \pi^- + p) \end{aligned} \quad (16.24)$$

Having said that, we mention another and more practical method which is a special feature to $K^0 - \bar{K}^0$ system, that is, to use the $\Delta S = \Delta Q$ selection rule. It is a rule that hadronic component in the decay process has to respect because at the quark level it is $s \rightarrow u + W^-$ (see Sect. 15.6). As $m_{K^\pm} > m_{K^0}$, no strangeness-conserving decay mode $d \rightarrow u e^- \bar{\nu}_e$ ($\bar{d} \rightarrow \bar{u} e^+ \nu_e$) exists, only the strangeness-changing mode $s \rightarrow u e^- \bar{\nu}_e$ ($\bar{s} \rightarrow \bar{u} e^+ \nu_e$) can occur, and the sign of the decay leptons can tell us which of the parents it comes from. The strangeness oscillation curve of CPLEAR in Fig. 16.1 was tagged in this way. Note, the $\Delta S = \Delta Q$ rule is obvious in the quark model, but experimentally one has to be careful about the possible existence of $\Delta S \neq \Delta Q$ processes.

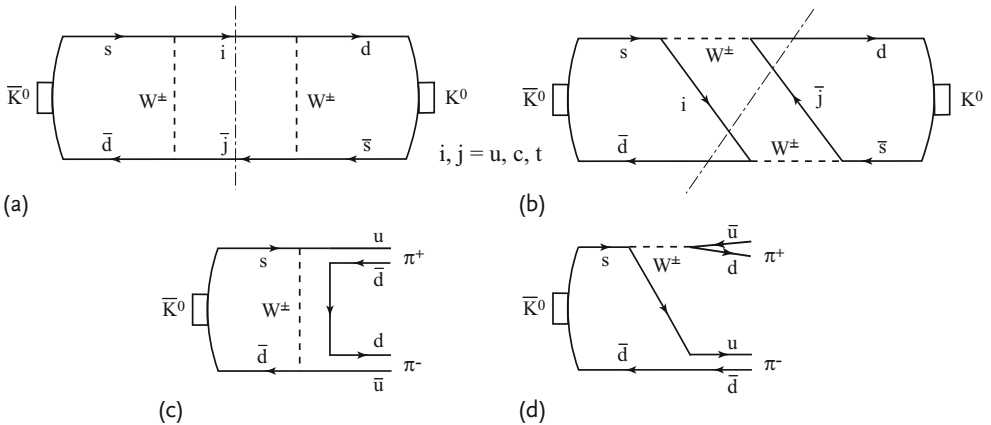


Figure 16.3 Box diagram. (a,b) $K^0 - \bar{K}^0$ mixing in the Standard Model. (c,d) Description of how 2π states appear in the intermediate states by cutting the mass matrix diagram across the dash-dotted lines. 3π or π/ν states can be constructed similarly.

The mass difference Δm arises because of mixing between K^0 and $\bar{K}^0$ through intermediate states such as those depicted in Fig. 16.2. Namely, it is an effect that arises from the nondiagonal elements of the mass matrix. In the language of the Standard Model, the intermediate states are given by $W^\pm$ and/or up-type quarks (u, c, t) shown in Fig. 16.3a,b.

Figure 16.3c,d shows how to associate the modern interpretation in terms of quarks with the old picture ($2\pi, l\pi\nu$, etc.). The mixing is a $\Delta S = 2$ effect and second-order in the weak interaction. Therefore Δm should be proportional to the square of the coupling constant $G_F^2 \sin^2 \theta_c$. From dimensional analysis

$$\Delta m \sim G_F^2 \sin^2 \theta_c m_K^5 \sim 10^{-4} \text{ eV} \quad (16.25)$$

which is not too far from the observed value.

16.1.4

Regeneration of K_1

A pure beam, whether K^0 or $\bar{K}^0$ at $t = 0$, becomes a pure K_2 beam after a long enough time ($t \gg \tau_S = 1/\Gamma_S$) where τ_S is K_1 's mean lifetime. This can be confirmed by the fact that two-prong decay (a topology that splits to two branches) modes appear in the neighborhood of production ($l \lesssim l_S = \beta\gamma\tau_S c$), where $\gamma = E_K/m_K$ is the Lorentz factor, but at distances $l \gg l_S$ only three-prong decay modes are observed (Fig. 16.4).

Let us consider making this beam pass through matter. K^0 having $S = +1$ is elastically ($K^0 N \rightarrow K^0 N$) or quasi-elastically ($K^0 p \rightarrow K^+ n$) scattered at most, but $\bar{K}^0$, having $S = -1$, can also produce hyperons [Eqs. (16.1b) (e)(f)] and thus have larger absorption rate. If we write the states after matter passage as

$$|K^0\rangle \rightarrow f|K^0\rangle, \quad |\bar{K}^0\rangle \rightarrow \bar{f}|\bar{K}^0\rangle \quad (16.26)$$

then

$$\begin{aligned} |K_2\rangle &= (|K^0\rangle - |\bar{K}^0\rangle)/\sqrt{2} \rightarrow (f|K^0\rangle - \bar{f}|\bar{K}^0\rangle)/\sqrt{2} \\ &= \frac{1}{2}(f + \bar{f})|K_2\rangle + \frac{1}{2}(f - \bar{f})|K_1\rangle \end{aligned} \quad (16.27)$$

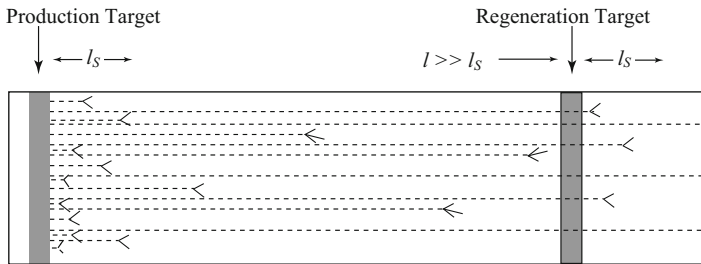


Figure 16.4 Regeneration of K_1 . Two-prong decay modes ($K \rightarrow 2\pi$) are limited to the neighborhood ($l \lesssim l_S$) of the production and regeneration targets. Three-prong decay modes ($K \rightarrow 3\pi$) are observed at long distances ($l \gg l_S$) as well as at short distances.

Since $f \neq \bar{f}$, the K_1 component, which did not exist before, appears after passage through the matter, which is referred to as regeneration.⁴⁾ The regeneration occurs from the difference in the reaction rate and does not depend on which reaction. However, depending on the target size, it can be classified as inelastic scattering, in which the kaon reacts with individual nucleons in the nucleus; diffraction scattering, in which it is scattered elastically by the nucleus as a whole; and coherent scattering, in which it reacts coherently with many atoms. Among them, coherent scattering plays a special role in the study of CP violation and deserves a more detailed explanation. As there is a small difference between the masses of K_1 and K_2 , the momentum k_2 of the parent K_2 and k_1 of the daughter K_1 are related by

$$\begin{aligned} k_2^2 + m_L^2 &= k_1^2 + m_S^2 \\ \Delta k &= k_1 - k_2 = \frac{m_L + m_S}{k_1 + k_2} (m_L - m_S) = \frac{m}{k} \Delta m \\ m &= (m_L + m_S)/2, \quad k = (k_1 + k_2)/2 \end{aligned} \quad (16.28)$$

Because of the Heisenberg uncertainty principle, scattering from targets within the range $d \lesssim 1/\Delta k = k/m\Delta m$ can interfere coherently. If we take $k = 10$ GeV, insertion of $m = m_K \simeq 0.5$ GeV, $\Delta m \simeq 3.5 \times 10^{-6}$ eV gives $d \sim 100$ cm. The scattering angle by coherent reaction becomes

$$\begin{aligned} d(1 - \cos \theta) &\lesssim \lambda \sim 1/k \\ \theta^2 &\sim 2\Delta k/k \sim 10^{-17} \end{aligned} \quad (16.29)$$

The momentum transfer to the generated k_1 is $p_T \lesssim 100$ eV, which is much smaller than the diffraction scattering ($p_T \lesssim 10$ MeV) and inelastic scattering ($p_T \gtrsim 100$ MeV) and it is almost a perfect forward scattering. Figure 16.5 shows an example of p_T^2 distribution after regeneration [244].

Writing the regenerated state after passage of matter with thickness L as

$$|\psi\rangle = |K_2\rangle + \rho(L)|K_1\rangle \quad (16.30)$$

we see $\rho(L)$ is given by [181, 257]

$$\rho(L) = \frac{i\pi N}{k} \{f(0) - \bar{f}(0)\} l_S \frac{1 - \exp[\{i(\Delta m/\Gamma_S) - 1/2\}L/l_S]}{1/2 - i(\Delta m/\Gamma_S)} \quad (16.31)$$

N	:	Number density of the scattering center
$k = m_K \beta \gamma$	:	Momentum of the kaon
$f(0), \bar{f}(0)$	:	Forward scattering amplitude of $K^0, \bar{K}^0$
$l_S = \beta \gamma / \Gamma_S$	:	Mean flight distance of K_1
$\beta \gamma$	:	Lorentz factor
Γ_S	:	Total decay rate of K_1
$\Delta m = m_L - m_S$	:	Mass difference of K_2 and K_1

4) This is like circularly polarized light traversing an optically active medium. It is also a typical phenomenon characteristic of quantum mechanics and is a manifestation of the principle of superposition.

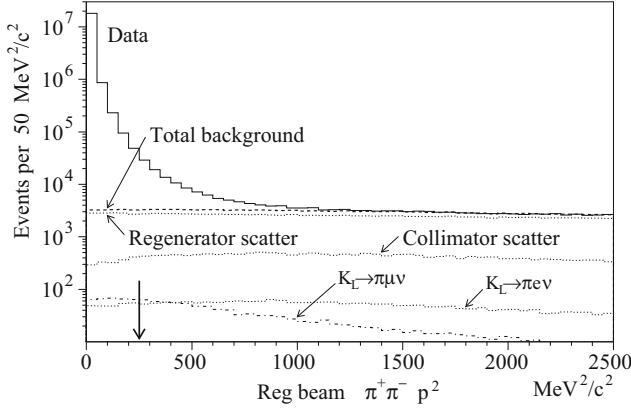


Figure 16.5 Data p_T^2 of $K \rightarrow \pi^+ \pi^-$ of the regenerator beam. The Monte Carlo predictions for the background components are overlaid. The arrow indicates cut to select sam-

ples [244]. Notice that the width of forward peak is determined not by intrinsic scattering width but by experimental resolution.

Proof: As is well known from optical theory, a combined wave of incoming and forward-scattered waves added coherently over all scattering centers is well described by a plane wave with wave number nk , where n is the index of refraction [176]. The expression for n is

$$n^2 = 1 + \frac{4\pi N f(0)}{k^2}, \quad n^2 - 1 \ll 1 \quad (16.32)$$

□

The eigenmass M of the wave is given by

$$k^2 + m_K^2 = (nk)^2 + M^2$$

$$M = \sqrt{m_K^2 - (n^2 - 1)k^2} \simeq m_K - \frac{(n^2 - 1)k^2}{2m_K} = m_K - \frac{2\pi N f(0)}{m_K} \quad (16.33)$$

Since K^0 and $\bar{K}^0$ give different indexes of refraction, the Schrödinger equation Eq. (16.10) is modified to

$$i \frac{\partial}{\partial t} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} - \frac{2\pi N}{m_K} \begin{bmatrix} f & 0 \\ 0 & \bar{f} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (16.34)$$

To derive the solution, we first note that the mass matrix can be expressed in an analogous form to that given in Eq. (16.13b):

$$\begin{bmatrix} M' - \delta\Delta_\lambda & -\Delta_\lambda/2 \\ -\Delta_\lambda/2 & M' + \delta\Delta_\lambda \end{bmatrix} \quad (16.35)$$

$$M' = \frac{1}{2}(\lambda_L + \lambda_S) - \frac{\pi N}{m_K}(f + \bar{f}), \quad \delta\Delta_\lambda = \frac{\pi N}{m_K}(f - \bar{f})$$

As δ is a small number, the solution of Eq. (16.35) is a small perturbation to $|K_1\rangle$, $|K_2\rangle$ and is expressed as

$$\begin{aligned} |K_{SM}\rangle &\simeq |K_1\rangle + \delta|K_2\rangle \\ |K_{LM}\rangle &\simeq |K_2\rangle - \delta|K_1\rangle \end{aligned} \quad (16.36)$$

where only up to $O(\delta)$ was considered. The eigenmass is given by the formula

$$\lambda_{LM} = \lambda_S - \frac{\pi N}{m_K} (f + \bar{f}) \quad (16.37)$$

Problem 16.2

Show that Eqs. (16.36) are the solutions to Eq. (16.34).

Setting the initial condition as the beam being pure K_2 before it enters the matter, we find the amplitude on exiting the regenerator of thickness L is given by

$$\begin{aligned} |K_2(t)\rangle &\simeq |K_{LM}(t)\rangle + \delta|K_1(t)\rangle \simeq |K_{LM}\rangle e^{-i\lambda_{LM}t} + \delta|K_{SM}\rangle e^{-i\lambda_{SM}t} \\ &\simeq \{|K_2\rangle - \delta|K_1\rangle\} e^{-i\lambda_{LM}t} + \delta|K_1\rangle e^{-i\lambda_{SM}t} \\ &= e^{-i\lambda_{LM}t} \left[|K_2\rangle + (e^{i\Delta\lambda m t} - 1)\delta|K_1\rangle \right] \\ \Delta\lambda m &= \lambda_{LM} - \lambda_{SM} = \lambda_L - \lambda_S = \Delta\lambda \end{aligned} \quad (16.38)$$

Inserting time $t = l/\beta\gamma$ to travel distance l and comparing with Eq. (16.30), we obtain

$$\begin{aligned} \rho(L) &= (e^{i\Delta\lambda L/\beta\gamma} - 1)\delta = -\frac{\pi N}{m_K} \frac{(f - \bar{f})}{\Delta\lambda} (1 - e^{i\Delta\lambda L/\beta\gamma}) \\ &= \frac{i\pi N}{k} (f - \bar{f}) l_S \frac{1 - \exp[i(\Delta m/\Delta\Gamma_S) - 1/2]l/l_S}{1/2 - i(\Delta m/\Gamma_S)} \\ |\rho(L)|^2 &= \pi^2 N^2 \frac{|f - \bar{f}|^2}{m_K^2} \frac{1 + e^{-l/l_S} - 2e^{-l/(2l_S)} \cos(\Delta m/\Gamma_S) l/l_S}{|\Delta m|^2 + \Gamma_S^2/4} \end{aligned} \quad (16.39)$$

In deriving the above expression, we have neglected Γ_L compared to Γ_S as $\Gamma_S \gg \Gamma_L$.

Sign of Δm The sign of Δm can be determined if there is another K_1 component whose intensity is comparable to that of the regenerated K_1 .

1. One method is to use a pure $K^0(\bar{K}^0)$ beam at $t = 0$. Suppose the regenerator of thickness $L = \beta\gamma t_L$ is set at distance $D = \beta\gamma t_D$ from the position of beam generation. The amplitude of the beam on exiting the regenerator can be described as

$$\begin{aligned} \psi(t_D + t_L) &\sim |K_1\rangle e^{-i\lambda_S(t_D + t_L)} + e^{-i\lambda_L(t_D + t_L)} \{|K_2\rangle + \rho(L)|K_1\rangle\} \\ &= e^{-i\lambda_L(t_D + t_L)} \left[|K_2\rangle + \left\{ e^{i\Delta\lambda(t_D + t_L)} + \rho(L) \right\} |K_1\rangle \right] \end{aligned} \quad (16.40)$$

The intensity of the K_1 component is proportional to

$$e^{-(D+L)/l_s} + |\rho(L)|^2 + 2e^{-(D+L)/2l_s} |\rho(L)| \cos \left\{ \arg[\rho(L)] - \frac{\Delta m}{\Gamma_s} \frac{D+L}{l_s} \right\} \quad (16.41)$$

where $l_s = \beta\gamma/\Gamma_s$ is the mean attenuation length of K_1 . By changing L , the phase of the cos term can be measured as a function of L and hence the sign of Δm can be determined. Using this method, the sign of $\Delta m = m_L - m_S$ was determined to be positive [272].

2. The second method is to use a CP-violating K_1 component (to be described soon) in the pure K_L beam. In this case, the K_1 component of the beam after passing the regenerator is given by

$$\langle K_1 | \psi(t) \rangle \sim \eta e^{-i\lambda_L t} + \rho(L) e^{-i\lambda_S t} \quad (16.42)$$

where η is the CP-violating amplitude (K_1 in K_L). The K_1 intensity is proportional to

$$|\eta|^2 e^{-\Gamma_L t} + |\rho|^2 e^{-\Gamma_S t} + 2|\eta||\rho| e^{-\bar{\Gamma} t} \cos[\Delta m + \arg(\rho) - \arg(\eta)] \quad (16.43)$$

Then by varying the length of the regenerator and measuring the ensuing modulation, the sign and magnitude of Δm can be determined. Note, however, this can be done only after the CP-violating amplitude has been determined without the regenerator.

16.1.5

Discovery of CP Violation

CP violation was discovered in 1964, seven years after the discovery of parity violation. As long-lived K_2 has $CP = -1$, it was thought that it can decay only to 3π and $\pi^\pm l^\mp \nu_l$, and not to 2π . Fitch and Cronin [99] observed for the first time that K_2 decays to 2π . Figure 16.6 shows their apparatus.

The experiments was carried out using the 30 GeV AGS (Alternating Gradient Synchrotron) proton accelerator at Brookhaven National Laboratory (BNL). Neutral kaons were extracted from a secondary beam created by bombarding the primary proton beam on a Be target through a collimator at 30° . Charged particles were eliminated by a sweeping magnet. The beam has average momentum $\sim 1100 \text{ MeV}/c$, and a helium bag⁵⁾ was placed in the decay region at $l \sim 20l_s$ (decay length of K_1) to minimize spurious interactions. To measure decay points and momenta of decay $\pi^\pm$, a pair of spectrometers consisting of a magnet, scintillation counters and spark chambers were placed in a symmetrical configuration.

5) This is a poor man's vacuum chamber. One would also have used proportional tubes or drift chambers with much less material and fast response instead of spark chambers if modern techniques had been available.

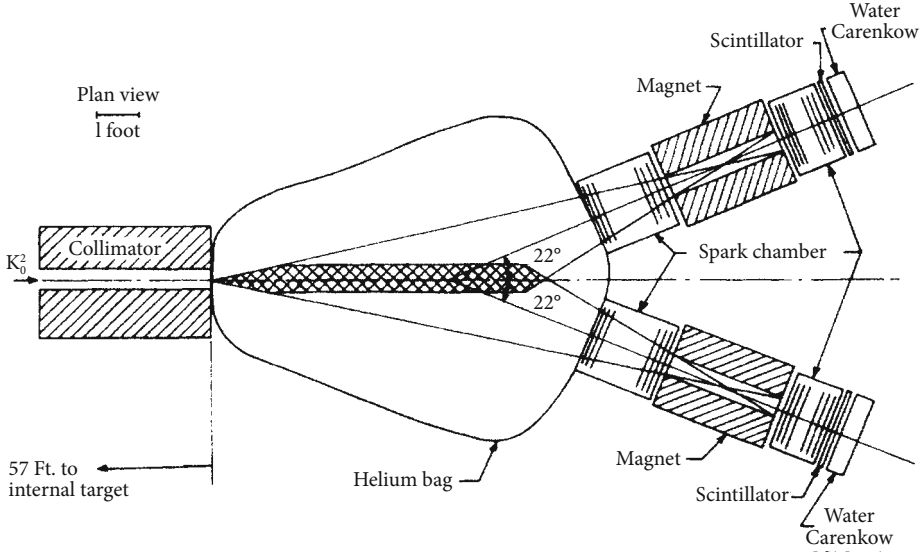


Figure 16.6 Discovery of CP violation. Layout of apparatus. K_2 beam comes in from the left and decays in a helium bag. 2π are detected by a pair of telescopes, which consist

of a magnet, scintillation counters and spark chambers, and determines the pions' momentum. The helium bag contains regenerators to produce 2π for calibration [99].

Furthermore, to calibrate the apparatus, a total of 43 g of tungsten was placed at every 28 cm in the decay region. Regenerated K_1 has the same momentum as its parent K_2 ; it is an ideal calibrator for $K_2 \rightarrow 2\pi$ events. Denoting the momenta of decay pions as $\mathbf{p}_1$, $\mathbf{p}_2$ and assuming the observed particles are pions, we can reconstruct an effective mass m^* of the parent particle from the equation

$$m^{*2} = (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 \simeq 2m_\pi^2 + 2p_1p_2(1 - \cos \theta) \quad (16.44)$$

Events that gives $m^* \simeq m_K$ are identified as decay products from kaons. Backgrounds from other kaon decays etc. are eliminated by requiring that the reconstructed kaon momentum $\mathbf{p}_K = \mathbf{p}_1 + \mathbf{p}_2$ is directed in the beam direction. Figure 16.7 shows events with $m^* \sim m_K$ and their momentum in the same direction as the beam. A clear signal above the background can be seen in the middle graph on the right-hand side. From the data

$$\text{BR}(K_2 \rightarrow \pi^+\pi^-) = \frac{\Gamma(K_2 \rightarrow \pi^+\pi^-)}{\Gamma(K_2 \rightarrow \text{all})} = 2.0 \pm 0.4 \times 10^{-3} \quad (16.45)$$

was obtained. In 1967, $K_2 \rightarrow \pi^0\pi^0$ was observed, too [110, 153].

Since observed kaons with short and long lives were proved not to be eigenstates of CP, we rename them K_S and K_L instead of K_1 and K_2 .

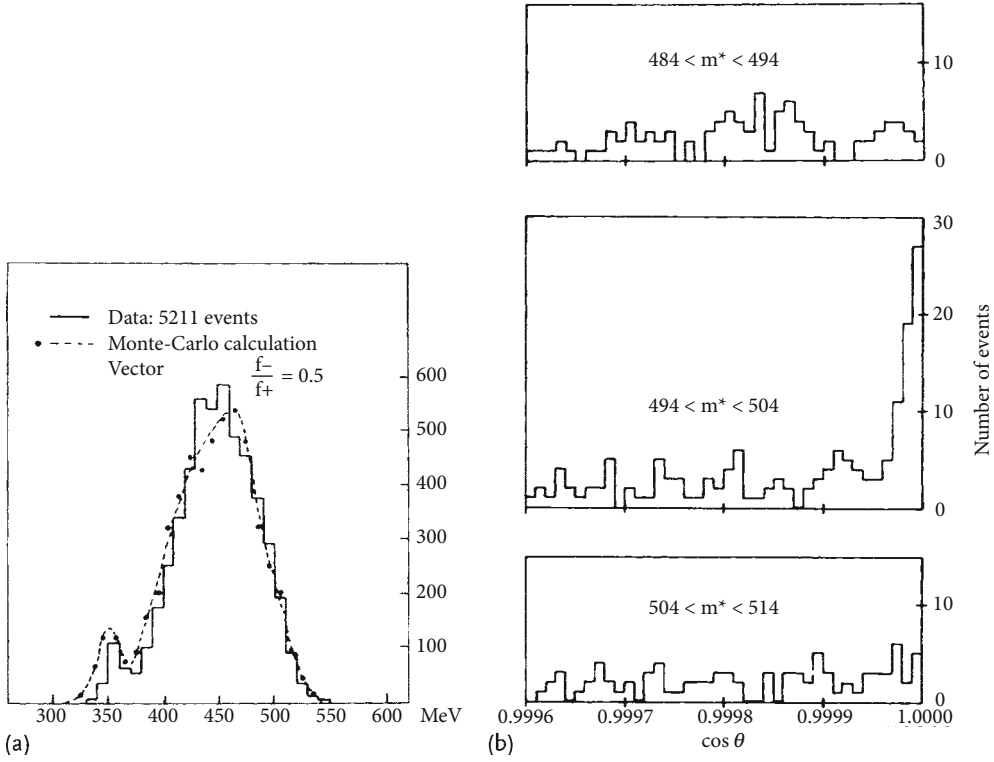


Figure 16.7 (a) Reconstructed mass distribution [99]. (b) Only data in the region $m^* \sim m_K$ show a peak at $\theta = 0$ that is above the background. This is evidence for $K_2 \rightarrow 2\pi$.

16.2

Formalism of CP and CPT Violation

16.2.1

CP, T, CPT Transformation Properties

We have to define CP violation parameters to describe the observed phenomena. As CP, T and CPT are related, we first list transformation properties of CP, T, CPT operations on particle states (see Appendix F). Denoting phase factors $\eta_i = e^{i\xi_i}$, $\bar{\eta}_i = e^{i\bar{\xi}_i}$

$$\text{CP}|q, \mathbf{p}, s_a \rangle_{\text{out}}^{\text{in}} = \eta_{\text{CP}} |\bar{q}, -\mathbf{p}, s_a \rangle_{\text{out}}^{\text{in}} \quad (16.46a)$$

$$\text{T}|q, \mathbf{p}, s_a \rangle_{\text{out}}^{\text{in}} = \eta_{\text{T}} |q, -\mathbf{p}, -s_a \rangle_{\text{in}}^{\text{out}} \quad (16.46b)$$

$$\text{CPT}|q, \mathbf{p}, s_a \rangle_{\text{out}}^{\text{in}} = \eta_{\text{CPT}} |\bar{q}, \mathbf{p}, -s_a \rangle_{\text{in}}^{\text{out}} \quad (16.46c)$$

$$\eta_{\text{CPT}} = \eta_{\text{CP}} \eta_{\text{T}} \quad (16.46d)$$

where $q(\bar{q})$ denotes a particle (antiparticle), $\mathbf{p}$ its momentum and s_α its spin component. Note T and CPT are antiunitary operators and complex conjugation of c-numbers is to be included in their operation (see Sect. 9.2.2). Defining

$$O_i|\bar{q}\rangle = \bar{\eta}_i|O_i\bar{q}\rangle \quad O_i = \text{CP, T, CPT}$$

and using $O_i^2 = 1$, we have

$$\bar{\eta}_i = \eta_i^* \quad (16.47)$$

By requiring CPT = TCP, we have

$$\begin{aligned} \text{CPT}|q, \mathbf{p}, s_\alpha, \text{in}\rangle &= \text{CP}\eta_T|q, -\mathbf{p}, -s_\alpha, \text{out}\rangle = \eta_T\eta_{\text{CP}}|\bar{q}, \mathbf{p}, -s_\alpha, \text{out}\rangle \\ \text{TCP}|q, \mathbf{p}, s_\alpha, \text{in}\rangle &= \text{T}\eta_{\text{CP}}|\bar{q}, -\mathbf{p}, s_\alpha, \text{in}\rangle = \eta_{\text{CP}}^*\eta_T^*|\bar{q}, \mathbf{p}, -s_\alpha, \text{out}\rangle \end{aligned} \quad (16.48)$$

where we have taken the complex conjugate of c-numbers. This means

$$\begin{aligned} \bar{\eta}_{\text{CPT}} &= \eta_{\text{CPT}} = \eta_{\text{CPT}}^* = \eta_{\text{CP}}\eta_T = 1 \\ \bar{\eta}_{\text{CP}} &= \eta_{\text{CP}}^* \\ \bar{\eta}_T &= \eta_T^* = \eta_{\text{CPT}}\eta_{\text{CP}} = \eta_T\eta_{\text{CP}}^2, \quad \eta_T = \eta_{\text{CPT}}\eta_{\text{CP}}^* = \eta_T^*\eta_{\text{CP}}^{*2} \end{aligned} \quad (16.49)$$

Mass Matrix We use $|A\rangle$ to denote a one-particle state and $|B\rangle$ for continuum states to which A can decay weakly. As there is no distinction between “in” and “out” for a one-particle state, $|A : \text{in}\rangle = |A : \text{out}\rangle = |A\rangle$. For spinless kaons in the rest frame, indices other than particle species are irrelevant. Applying T invariance to the amplitude,

$$\begin{aligned} \langle i|H_W|j\rangle &= \langle i|T^{-1}TH_WT^{-1}T|j\rangle = (\eta_i^*\eta_j\langle i|H_W|j\rangle)^* \\ &= \eta_i\eta_j^*\langle j|H_W|i\rangle \\ \therefore \langle n|H_W|K^0\rangle &= \eta_T^*\eta_n\langle K^0|H_W|n\rangle \\ \langle n|H_W|\bar{K}^0\rangle &= \bar{\eta}_T^*\eta_n\langle \bar{K}^0|H_W|n\rangle \end{aligned} \quad (16.50)$$

Applying Eq. (16.50) to the decay matrix in Eq. (16.11),

$$\begin{aligned} \Gamma_{21} &= 2\pi \sum_n \delta(E_0 - E_n) \langle 2|T^{-1}TH_WT^{-1}T|n\rangle \langle n|T^{-1}TH_WT^{-1}T|1\rangle \\ &= 2\pi \sum_{Tn} \delta(E_0 - E_n) (\bar{\eta}_T^*\eta_n)^* \langle Tn|H_W|T2\rangle (\eta_T^*\eta_n) \langle T1|H_W|Tn\rangle \\ &= (\bar{\eta}_T\eta_T^*)2\pi \sum_n \delta(E_0 - E_n) \langle 1|H_W|n\rangle \langle n|H_W|2\rangle = \eta_{\text{CP}}^2\Gamma_{12} \end{aligned} \quad (16.51)$$

where $|Tn\rangle$ denotes T-transformed state defined by Eq. (16.46). As T changes the particle's momentum as well as spin orientation, $|Tn\rangle \neq |n\rangle$, but summation on all the continuum states makes the difference disappear. Similar arguments apply to the mass part in Eq. (16.11), and

$$M_{21} = \eta_{\text{CP}}^2 M_{12}, \quad \Gamma_{21} = \eta_{\text{CP}}^2 \Gamma_{12} \quad (16.52a)$$

This means

$$A_{21} = \eta_{\text{CP}}^2 A_{12}, \quad \text{or} \quad |A_{12}| = |A_{21}| \quad \text{T invariance} \quad (16.52b)$$

It also means

$$\frac{M_{12}^*}{M_{12}} = \frac{\Gamma_{12}^*}{\Gamma_{12}}, \quad \text{or} \quad \text{Im}(M_{12}^* \Gamma_{12}) = 0 \quad (16.52c)$$

No constraints are obtained for A_{11} and A_{22} .

Requiring CPT invariance gives

$$A_{11} = A_{22} \quad \text{CPT invariance} \quad (16.53)$$

No constraints are made for A_{12} or A_{21} . CP invariance gives

$$A_{11} = A_{22} \quad \text{and} \quad |A_{12}| = |A_{21}| \quad \text{CP invariance} \quad (16.54)$$

Symmetry Properties of Decay Amplitudes Denoting the CPT operation as Θ , the application of the requirement of its invariance to a weak decay amplitude for a process $i \rightarrow f$ can be written as

$$\begin{aligned} \langle f : \text{out} | H_W | i \rangle &= \langle f : \text{out} | \Theta^{-1} \Theta H_W \Theta^{-1} \Theta | i \rangle = \langle \Theta f : \text{in} | H_W | \Theta i \rangle^* \\ &= \sum_{\text{out}} (\langle \Theta f : \text{in} | \Theta f : \text{out} \rangle \langle \Theta f : \text{out} | H_W | \Theta i \rangle)^* \\ &= \langle \Theta f : \text{out} | H_W | \Theta i \rangle^* e^{2i\delta_f} \end{aligned} \quad (16.55)$$

where we have used [see Eq. (6.12)]

$$\langle f : \text{out} | f : \text{in} \rangle = \langle f : \text{out} | S | f : \text{out} \rangle = S_{ff} = e^{2i\delta_f} \quad (16.56)$$

δ_f is the scattering phase shift of the f state as well as $\bar{f}$ because the strong interaction is invariant under C transformation. $|\Theta f\rangle$ is a CPT-transformed state as defined in Eq. (16.46). Then Eq. (16.55) means that a decay amplitude and its CPT conjugate are related by

$$\begin{aligned} \langle f : \text{out} | H_W | i \rangle &= A e^{i\delta_f} \\ \langle \Theta f : \text{out} | H_W | \Theta i \rangle &= A^* e^{i\delta_f} \end{aligned} \quad (16.57)$$

where A is a complex number. Similarly, T and CP invariance can be expressed as

$$\begin{aligned} \langle f : \text{out} | H_W | i \rangle &= \langle f : \text{out} | T^{-1} T H_W T^{-1} T | i \rangle = \langle T f : \text{in} | H_W | T i \rangle^* \\ &= \langle T f : \text{out} | H_W | T i \rangle^* e^{2i\delta_f} \end{aligned} \quad (16.58)$$

$$\langle f : \text{out} | H_W | i \rangle = \langle \text{CP } f : \text{out} | H_W | \text{CP } i \rangle \quad (16.59)$$

Neutral K Meson Decays As the K meson is a scalar,

$$|\Theta K^0\rangle = |\text{CP} K^0\rangle = |\bar{K}^0\rangle, \quad |\text{T} K^0\rangle = |K^0\rangle \quad (16.60)$$

Since a momentum-inverted final state can be reproduced by rotation from the original state and there is no specific orientation in the total angular momentum zero state, we have

$$|\Theta f : \text{out}\rangle = |\text{CP} f : \text{out}\rangle = |\bar{f} : \text{out}\rangle, \quad |\text{T} f : \text{out}\rangle = |f : \text{out}\rangle \quad (16.61)$$

Separating the strong phase and defining weak amplitudes A and $\bar{A}$ by

$$\langle f : \text{out} | H_W | K^0 \rangle \equiv A_f e^{i\delta_f}, \quad \langle \bar{f} : \text{out} | H_W | \bar{K}^0 \rangle \equiv \bar{A}_{\bar{f}} e^{i\delta_{\bar{f}}} \quad (16.62)$$

Equations (16.57), (16.58) and (16.59) relate one to the other:

$$\text{CPT} \quad A(\bar{i} \rightarrow \bar{f}) \equiv \bar{A}_{\bar{f}} = A^*(i \rightarrow f) \quad (16.63a)$$

$$\text{T} \quad \begin{cases} A(i \rightarrow f) \equiv A_f = A^*(i \rightarrow f) \\ A(\bar{i} \rightarrow \bar{f}) = A^*(\bar{i} \rightarrow \bar{f}) \end{cases} \quad (16.63b)$$

$$\text{CP} \quad A(\bar{i} \rightarrow \bar{f}) = A(i \rightarrow f) \quad (16.63c)$$

16.2.2

Definition of CP Parameters

Now we are ready to apply the formalism we have developed so far to the neutral K mesons.

Mixing Parameters Since we learned that observed neutral kaons are not eigenstates of CP, we can write

$$\begin{aligned} |K_S\rangle &= \frac{1}{\sqrt{1+|\epsilon_S|^2}} (|K_1\rangle + \epsilon_S |K_2\rangle) \\ &= \frac{1}{\sqrt{2(1+|\epsilon_S|^2)}} \left\{ (1+\epsilon_S) |K^0\rangle + (1-\epsilon_S) |\bar{K}^0\rangle \right\} \\ |K_L\rangle &= \frac{1}{\sqrt{1+|\epsilon_L|^2}} (|K_2\rangle + \epsilon_L |K_1\rangle) \\ &= \frac{1}{\sqrt{2(1+|\epsilon_L|^2)}} \left\{ (1+\epsilon_L) |K^0\rangle - (1-\epsilon_L) |\bar{K}^0\rangle \right\} \end{aligned} \quad (16.64a)$$

$\epsilon_{S,L} \neq 0$ means CP violation. They need not be equal. We decompose them as

$$\epsilon_S = \epsilon + \delta, \quad \epsilon_L = \epsilon - \delta \quad (16.64b)$$

We shall show soon that CPT invariance means $\delta = 0$ and T invariance means $\epsilon = 0$. For the sake of later discussions, we introduce an alternative form for ex-

pressing K_S and K_L :

$$|K_S\rangle = p\sqrt{1-z}|K^0\rangle + q\sqrt{1+z}|\bar{K}^0\rangle \quad (16.65a)$$

$$|K_L\rangle = p\sqrt{1+z}|K^0\rangle - q\sqrt{1-z}|\bar{K}^0\rangle \quad (16.65b)$$

$$|p|^2 + |q|^2 = 1 \quad (16.65c)$$

Then it is easy to show that the mass matrix is expressed as

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} = \begin{bmatrix} M + \frac{z}{2}\Delta_\lambda & -\frac{p}{2q}\sqrt{1-z^2}\Delta_\lambda \\ -\frac{q}{2p}\sqrt{1-z^2}\Delta_\lambda & M - \frac{z}{2}\Delta_\lambda \end{bmatrix} \quad (16.66)$$

$$M = \frac{1}{2}(\lambda_L + \lambda_S), \quad \Delta_\lambda = \lambda_L - \lambda_S$$

Problem 16.3

Derive Eq. (16.66).

CPT and T invariance means $\mathcal{A}_{11} = \mathcal{A}_{22}$ and $|\mathcal{A}_{21}| = |\mathcal{A}_{12}|$, which, in turn, means $z = 0$ (CPT), and $|q/p| = 1$ (T), respectively. In terms of the mass matrix elements, δ, ϵ can also be defined as

$$\delta \equiv \frac{\mathcal{A}_{22} - \mathcal{A}_{11}}{2\Delta_\lambda}, \quad \epsilon \equiv \frac{\mathcal{A}_{21} - \mathcal{A}_{12}}{2\Delta_\lambda} \quad (16.67)$$

The two definitions Eqs. (16.67) and (16.64) are equivalent in first order in ϵ or δ . Since the second-order term never appears in this book, either definition is valid. The former is convenient for theoretical and the latter for experimental evaluations. To first order, z, p, q are related to δ, ϵ by

$$z = -2\delta, \quad (16.68a)$$

$$\frac{p}{q} = \frac{1+\epsilon}{1-\epsilon} \quad \text{or} \quad \epsilon = \frac{p-q}{p+q} \quad (16.68b)$$

For theoretical discussions, ϵ is expressed as

$$\begin{aligned} \epsilon &= -\frac{\text{Im}(M_{12}) - i\text{Im}(\Gamma_{12})/2}{\Delta\Gamma/2 - i\Delta m} \\ &= \frac{\sin\phi_{\text{SW}}}{\Delta m} e^{i\phi_{\text{SW}}} \left[-\text{Im}(M_{12}) + i\frac{\text{Im}(\Gamma_{12})}{2} \right] \end{aligned} \quad (16.69a)$$

$$\phi_{\text{SW}} \equiv \tan^{-1} \frac{2\Delta m}{\Delta\Gamma} = 43.49 \pm 0.08^\circ \quad (16.69b)$$

where ϕ_{sw} is referred as the superweak phase for the reason discussed around Eq. (16.184).

To summarize, CP is violated if ϵ_S or ϵ_L is nonzero. CPT invariance means $\delta = 0$ which, in turn, means $\epsilon_S = \epsilon_L = \epsilon$. In this case, CP violation is equivalent to T violation and $\text{Re}(\epsilon) \neq 0$. Note ϵ itself is a phase dependent variable as we will discuss in more details in Sect. 16.3.2.

Problem 16.4

Prove that the solution of Eq. (16.66), where δ, ϵ are defined by Eq. (16.67), is given by Eq. (16.64) in first order in δ and ϵ .

Bell–Steinberger Relation Taking the inner product of $|K_S\rangle$ and $|K_L\rangle$, we see

$$\langle K_S | K_L \rangle = 2[\text{Re}(\epsilon) - i\text{Im}(\delta)] \quad (16.70)$$

As this equation shows, the observed states K_S and K_L are not orthogonal and mix together. The amount of overlap is a measure of CP and CPT violation. Its value can be estimated approximately as follows. Using the relation

$$\mathcal{A}|K_S\rangle = \lambda_S|K_S\rangle, \quad \mathcal{A}|K_L\rangle = \lambda_L|K_L\rangle \quad (16.71)$$

and multiplying by $\langle K_L|$ and $\langle K_S|$ from the left, respectively, we obtain

$$\langle K_L | \mathcal{A} | K_S \rangle = \lambda_S \langle K_L | K_S \rangle \quad (16.72a)$$

$$\langle K_S | \mathcal{A} | K_L \rangle = \lambda_L \langle K_S | K_L \rangle \quad (16.72b)$$

Subtracting the complex conjugate of Eq. (16.72a) from Eq. (16.72b), we have

$$(\lambda_L - \lambda_S^*) \langle K_S | K_L \rangle = \langle K_S | \mathcal{A} - \mathcal{A}^\dagger | K_L \rangle = -i \langle K_S | \Gamma | K_L \rangle \quad (16.73)$$

Insertion of $\langle K_S | K_L \rangle = 2[\text{Re}(\epsilon) - i\text{Im}(\delta)]$ and Eq. (16.11c) gives

$$2(\bar{\Gamma} + i\Delta m)[\text{Re}(\epsilon) - i\text{Im}(\delta)] = \sum_f A_L(f) A_S^*(f) \rho(f) \quad (16.74)$$

$$A_L(f) = \langle f | H_W | K_L \rangle, \quad A_S^*(f) = \langle K_S | H_W | f \rangle$$

where $\rho(f)$ denotes the phase space volume of the state f . The expression is phase convention independent. The rhs is on shell and the terms are observables, hence can be determined by experiments. Equation (16.74) is called the Bell–Steinberger relation or unitarity condition, as it is derived from the unitarity relation Eq. (16.7). Denoting the partial decay width as $\Gamma(K_{S,L} \rightarrow f) = |A_{S,L}(f)|^2 \rho(f)$ and applying

the Schwartz inequality to Eq. (16.74), we have

$$\begin{aligned}
 2|\overline{F} + i\Delta m||\text{Re}(\epsilon) - i\text{Im}(\delta)| &\leq \sum_f \sqrt{F_L(f)} \sqrt{F_S(f)} \\
 &\leq \sqrt{\left(\sum_f F_S(f)\right) \left(\sum_f F_L(f)\right)} \\
 &= \sqrt{F_S F_L} \\
 \text{or } |\text{Re}(\epsilon) - i\text{Im}(\delta)| &\leq \frac{1}{2} \sqrt{\frac{F_S F_L}{\overline{F}^2 + \Delta m^2}} \\
 &\simeq \sqrt{\frac{F_L}{2F_S}} \simeq 0.03
 \end{aligned} \tag{16.75a}$$

This is the interesting result that a bound can be placed on a CP- and CPT-violating quantity from knowledge of CP-conserving quantities.

Time Development of K^0 and $\overline{K}^0$ Time development of a wave function is described by $|\psi(t)\rangle = e^{-i\mathcal{H}t}|\psi(0)\rangle$. To express it in a more tangible form, we write the inverse of Eq. (16.65)

$$\begin{aligned}
 |K^0\rangle &= \frac{1}{2p} \left(\sqrt{1-z}|K_S\rangle + \sqrt{1+z}|K_L\rangle \right) \\
 |\overline{K}^0\rangle &= \frac{1}{2q} \left(\sqrt{1+z}|K_S\rangle - \sqrt{1-z}|K_L\rangle \right)
 \end{aligned} \tag{16.76}$$

Then the time development of pure K^0 at $t = 0$ is given by

$$\begin{aligned}
 |K^0(t)\rangle &= e^{-i\mathcal{H}t}|K^0\rangle = \frac{1}{2p} \left[\sqrt{1-z}|K_S\rangle e^{-i\lambda_S t} + \sqrt{1+z}|K_L\rangle e^{-i\lambda_L t} \right] \\
 &= [g_+(t) + z g_-(t)]|K^0\rangle - \sqrt{1-z^2} \frac{q}{p} g_-(t) |\overline{K}^0\rangle
 \end{aligned} \tag{16.77a}$$

where

$$g_{\pm}(t) = \frac{1}{2} (e^{-i\lambda_L t} \pm e^{-i\lambda_S t})$$

Similarly

$$|\overline{K}^0(t)\rangle = [g_+(t) - z g_-(t)]|\overline{K}^0\rangle - \sqrt{1-z^2} \frac{p}{q} g_-(t) |K^0\rangle \tag{16.77b}$$

Combining the two equations, we have

$$e^{-i\mathcal{H}t} = \begin{bmatrix} g_+(t) + z g_-(t) & -\frac{p}{q} \sqrt{1-z^2} g_-(t) \\ -\frac{q}{p} \sqrt{1-z^2} g_-(t) & g_+(t) - z g_-(t) \end{bmatrix} \tag{16.78}$$

This is an exact formula, but in most cases approximations to first order in δ, ϵ are more convenient. We list several equivalent formula for later reference. Substituting $z = -2\delta, p/q \simeq (1 + \epsilon)/(1 - \epsilon)$, we have

$$|K^0(t)\rangle = |K^0\rangle\langle K^0|e^{-i\mathcal{A}t}|K^0\rangle + |\bar{K}^0\rangle\langle\bar{K}^0|e^{-i\mathcal{A}t}|K^0\rangle \quad (16.79a)$$

$$= |K^0\rangle[g_+(t) - 2\delta g_-(t)] + |\bar{K}^0\rangle[2(\mathcal{A}_{21}/\Delta_\lambda)g_-(t)] \quad (16.79b)$$

$$= |K^0\rangle[g_+(t) - 2\delta g_-(t)] - |\bar{K}^0\rangle[(1 - 2\epsilon)g_-(t)] + O(\epsilon^2, \delta^2) \quad (16.79c)$$

$$= \frac{1}{\sqrt{2}}[(1 - \epsilon + \delta)e^{-i\lambda_S t}|K_S\rangle + (1 - \epsilon - \delta)e^{-i\lambda_L t}|K_L\rangle] + O(\epsilon^2, \delta^2) \quad (16.79d)$$

$$|\bar{K}^0(t)\rangle = |\bar{K}^0\rangle\langle\bar{K}^0|e^{-i\mathcal{A}t}|\bar{K}^0\rangle + |K^0\rangle\langle K^0|e^{-i\mathcal{A}t}|\bar{K}^0\rangle \quad (16.80a)$$

$$= |\bar{K}^0\rangle[g_+(t) + 2\delta g_-(t)] + |K^0\rangle[2(\mathcal{A}_{12}/\Delta_\lambda)g_-(t)] \quad (16.80b)$$

$$= |\bar{K}^0\rangle[g_+(t) + 2\delta g_-(t)] - |K^0\rangle[(1 + 2\epsilon)g_-(t)] + O(\epsilon^2, \delta^2) \quad (16.80c)$$

$$= \frac{1}{\sqrt{2}}[(1 + \epsilon - \delta)e^{-i\lambda_S t}|K_S\rangle - (1 + \epsilon + \delta)e^{-i\lambda_L t}|K_L\rangle] + O(\epsilon^2, \delta^2) \quad (16.80d)$$

Formulae with CPT Invariance Assumption As $\delta \simeq 0$ experimentally, and since many expressions can be simplified, we assume CPT invariance from now on until we test it later in Sect. 16.4. Setting $z = 0$, we have from Eqs. (16.14) and (16.15)

$$\begin{aligned} \mathcal{A}_{11} = \mathcal{A}_{22} = M, \quad \lambda_{S,L} = \mathcal{A}_{11} \pm \sqrt{\mathcal{A}_{12}\mathcal{A}_{21}} = \mathcal{A}_{11} \pm \frac{q}{p}\mathcal{A}_{12} \\ = \mathcal{A}_{11} \pm \frac{p}{q}\mathcal{A}_{21} \end{aligned} \quad (16.81a)$$

$$\mathcal{A}_{12} = -\frac{\Delta_\lambda}{2}\frac{p}{q}, \quad \mathcal{A}_{21} = -\frac{\Delta_\lambda}{2}\frac{q}{p}, \quad \mathcal{A}_{21} + \mathcal{A}_{12} = -\Delta_\lambda + O(\epsilon^2) \quad (16.81b)$$

$$\frac{q}{p} = \sqrt{\frac{\mathcal{A}_{21}}{\mathcal{A}_{12}}} = \sqrt{\frac{M_{12}^* - i\Gamma_{12}^*/2}{M_{12} - i\Gamma_{12}/2}} \quad (16.81c)$$

$$\Delta_\lambda = -2\sqrt{\mathcal{A}_{12}\mathcal{A}_{21}} = -(\mathcal{A}_{21} + \mathcal{A}_{12}) + O(\epsilon^2) \simeq -2\left(\text{Re}M_{12} + \frac{i}{2}\text{Re}\Gamma_{12}\right) \quad (16.81d)$$

The above equations are obtained simply by setting $z = 0$ in Eq. (16.66). Note, by rephasing [Eq. (16.17)], $\mathcal{A}_{12} \rightarrow e^{2i\xi}\mathcal{A}_{12}$, $\mathcal{A}_{21} \rightarrow e^{-2i\xi}\mathcal{A}_{21}$, but also by Eq. (16.81c),

$q/p \rightarrow (q/p)e^{-2i\xi}$, therefore $(q/p)\mathcal{A}_{12}$ and $(p/q)\mathcal{A}_{21}$ are phase-independent variables. They should be, as Δ_λ is an observable. Then our choice of phase Eq. (16.64a), or equivalently Eq. (16.68b), constrains the phase of $\mathcal{A}_{12}$ and $\mathcal{A}_{21}$, too. Actually, by virtue of Eq. (16.81b), $\mathcal{A}_{12}, \mathcal{A}_{21} = \Delta_\lambda/2 + O(\epsilon)$. Therefore both M_{12} and Γ_{12} are very close to real numbers and their phase is small [$\sim O(\epsilon)$],

$$\therefore \Delta m = m_L - m_S = -2\text{Re}\left(\sqrt{\mathcal{A}_{12}\mathcal{A}_{21}}\right) \simeq -2\text{Re}M_{12} \quad (16.82a)$$

$$\Delta\Gamma = \Gamma_S - \Gamma_L = 4\text{Im}\left(\sqrt{\mathcal{A}_{12}\mathcal{A}_{21}}\right) \simeq 2\text{Re}\Gamma_{12} \quad (16.82b)$$

$$\Delta m \Delta\Gamma = -4\text{Re}(M_{12}\Gamma_{12}^*) \quad (16.82c)$$

$$\text{Im}M_{12} \ll \text{Re}M_{12}, \quad \text{Im}\Gamma_{12} \ll \text{Re}\Gamma_{12} \quad (16.82d)$$

Here $\simeq$ means a small deviation due to CP violation. If CP invariance is assumed, the relations are exact.

Problem 16.5

Derive (16.82).

16.3

CP Violation Parameters

16.3.1

Observed Parameters

2 π Decay Parameters η_{+-} , η_{00} The 2 π CP-violating parameters are defined by

$$\eta_{+-} \equiv \frac{\langle \pi^+ \pi^- | H_W | K_L \rangle}{\langle \pi^+ \pi^- | H_W | K_S \rangle} = |\eta_{+-}| e^{i\phi_{+-}} \quad (16.83a)$$

$$\eta_{00} \equiv \frac{\langle \pi^0 \pi^0 | H_W | K_L \rangle}{\langle \pi^0 \pi^0 | H_W | K_S \rangle} = |\eta_{00}| e^{i\phi_{00}} \quad (16.83b)$$

$|\eta_{+-}|$ and $|\eta_{00}|$ can be determined from measurement of the 2 π decay ratio. The phase angle ϕ_{+-} , ϕ_{00} as well as the sign of Δm can be determined from the interference effect between $K_S \rightarrow 2\pi$ and $K_L \rightarrow 2\pi$. Suppose there is a beam

- 6) Note our definition of Δm and $\Delta\Gamma$ is to make them both positive. This is possible for the neutral kaons because they were measured. However, some authors prefer to use $\Delta m = m_L - m_S$, $\Delta\Gamma = \Gamma_L - \Gamma_S$ or vice versa ($S \leftrightarrow L$) for a general treatment including neutral D and B mesons. With our choice of phase convention and

in the absence of CP violation, $\mathcal{A}_{12} = M_{12} - i\Gamma_{12}/2 = -\Delta_\lambda/2 = -\Delta m - i\Delta\Gamma/2$. The physical quantity does not depend on the phase choice, but some intermediate variables may change, so one must be careful about the sign convention when reading the literature.

expressed by

$$|\psi(t)\rangle = |K_L(t)\rangle + \rho |K_S(t)\rangle = |K_L(0)\rangle e^{-i\lambda_L t} + \rho |K_S(0)\rangle e^{-i\lambda_S t} \quad (16.84)$$

The 2π decay amplitudes are

$$\langle 2\pi | \psi(t) \rangle = [\eta e^{-i\lambda_L t} + \rho e^{-i\lambda_S t}] \langle 2\pi | K_S \rangle \quad (16.85)$$

As the number of 2π decays is proportional to $|\langle 2\pi | \psi(t) \rangle|^2$

$$\begin{aligned} N_{2\pi} &= N_K \left| \rho e^{(i\Delta m - \Gamma_S/2)t} + \eta e^{-(\Gamma_L/2)t} \right|^2 \\ &= N_K \left\{ |\eta|^2 e^{-\Gamma_L t} + |\rho|^2 e^{-\Gamma_S t} + 2e^{-\bar{\Gamma}t} |\eta\rho| \cos(\Delta m t + \phi_\rho - \phi_\eta) \right\} \\ \phi_\rho &= \arg(\rho), \quad \phi_\eta = \arg(\eta_{+-} \text{ or } \eta_{00}), \quad \bar{\Gamma} = (\Gamma_S + \Gamma_L)/2 \end{aligned} \quad (16.86)$$

One can prepare ρ in two ways.

(1) In the first method, a pure $K^0(\bar{K}^0)$ beam is produced at $t = 0$, in which case $\rho = \pm 1$, $\phi_\rho = 0$. If $N_{K^0} = N$, $N_{\bar{K}^0} = \bar{N}$ at $t = 0$, one substitutes $|\rho|^2 \rightarrow 1$ in the second term and $|\rho| \rightarrow (N - \bar{N})/(N + \bar{N})$ in the third term of Eq. (16.86). In order to maximize the interference effect, one needs

$$|\rho| \exp(-\bar{\Gamma}t) \sim |\eta| \sim 2 \times 10^{-3} \quad (16.87)$$

This means that the detector is to be placed in the neighborhood of $l = \beta\gamma t \simeq 12l_S$ where $l_S = \beta\gamma/\Gamma_S$ is the mean decay length. From the time dependence of the interference effect, one can determine both Δm and ϕ . Figure 16.8 shows the 2π

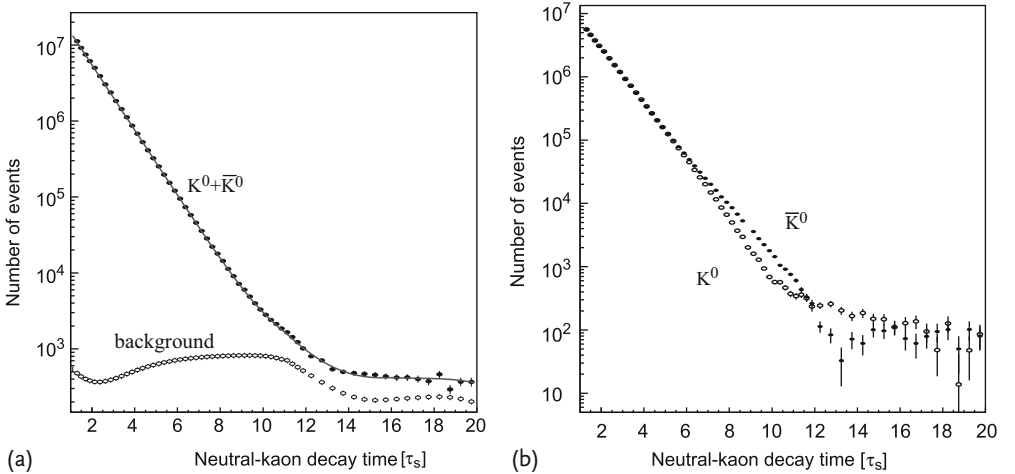


Figure 16.8 Decay curves of pure K^0 and $\bar{K}^0$ beams at $t = 0$. (a) $K^0 + \bar{K}^0$ curve shows no interference. (b) K^0 and $\bar{K}^0$ separately [107].

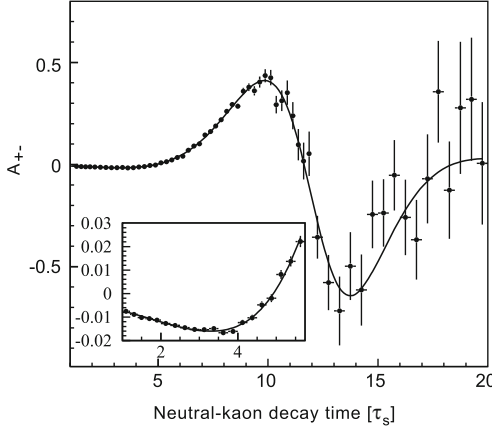


Figure 16.9 Asymmetry curve of neutral $K \rightarrow \pi^+\pi^-$ between initially pure K^0 and $\bar{K}^0$ beams [107].

decay curves. When decay products from both pure K^0 and $\bar{K}^0$ beams at $t = 0$ are added together, there is no interference as is shown in Fig. 16.8a, but separately they show the interference effect. Figure 16.9 gives an asymmetry curve defined by

$$\begin{aligned}
 A_{+-} &= \frac{N_{2\pi}(\rho = -1) - N_{2\pi}(\rho = 1)}{N_{2\pi}(\rho = -1) + N_{2\pi}(\rho = 1)} \\
 &= -\frac{2|\eta_j|e^{(\Gamma_s - \Gamma_l)t/2}\cos(\Delta m t - \phi_j)}{1 + |\eta_j|^2 e^{(\Gamma_s - \Gamma_l)t}}
 \end{aligned} \tag{16.88}$$

The world average values of the decay parameters are given by [311]⁷⁾

$$\begin{aligned}
 |\eta_{+-}| &= 2.233 \pm 0.012 \times 10^{-3} \\
 |\eta_{00}| &= 2.222 \pm 0.012 \times 10^{-3} \\
 \phi_{+-} &= 43.51 \pm 0.05^\circ \\
 \phi_{00} &= 43.52 \pm 0.05^\circ
 \end{aligned} \tag{16.89}$$

Historically, however, both parameters were determined by the second method, using regeneration.

(2) The second method uses a regenerator to produce a K_S beam. Given the regenerator thickness L , we have $\rho = \rho(L)$ given by (16.31). The measured phase is $\Delta m t$ and $\phi_\rho - \phi_\eta$. Since both t and L (which resides in ϕ_ρ) can be varied, ϕ_η as well as the sign of Δm can be determined. If one observes the semileptonic decay mode, one can determine ϕ_ρ because $\phi_\eta = 0$. Usually, the regenerator method is accompanied by systematic errors associated with ϕ_ρ , and it is not quite an ideal method. However, if one measures both $\pi^+\pi^-$ and $\pi^0\pi^0$ modes and takes the

7) These values were determined using all the methods discussed in the following. The values determined from $k \rightarrow 2\pi$ alone are less accurate.

difference, the systematic error associated with ρ almost cancels and a very precise value of $\phi_{00} - \phi_{+-}$ can be obtained. The most recent experimental data give [244]

$$\Delta\phi \equiv \phi_{00} - \phi_{+-} = 0.39 \pm 0.45^\circ \quad (16.90)$$

which is consistent with $\phi_{00} = \phi_{+-}$. A global analysis by the Particle Data Group gives a much better value [311]:

$$\Delta\phi = 0.006 \pm 0.014 \quad (16.91)$$

$\Delta\phi$ gives a measure of CPT violation [Eq. (16.175)], which will be discussed later.

Asymmetry A_L Since the three-body leptonic decay modes are not CP eigenstates, the CP violation effect appears as the difference between the decay mode $e^+ \nu_e \pi^-$ and its CP conjugate $e^- \bar{\nu}_e \pi^+$, that is, as the asymmetry A_L of the kaon with longer life:

$$A_L \equiv \frac{\Gamma(K_L \rightarrow e^+ \nu_e \pi^-) - \Gamma(K_L \rightarrow e^- \bar{\nu}_e \pi^+)}{\Gamma(K_L \rightarrow e^+ \nu_e \pi^-) + \Gamma(K_L \rightarrow e^- \bar{\nu}_e \pi^+)} \quad (16.92)$$

The asymmetry also appears in the muonic decay mode. K_S can also decay to them in principle but, the 2π decay rate being much larger, the branching ratio is small:

$$\frac{\Gamma(K_S \rightarrow e \nu \pi)}{\Gamma_S} \sim \frac{\Gamma(K_L \rightarrow e \nu \pi)}{\Gamma_S} = \text{BR}(K_L \rightarrow e \nu \pi) \frac{\Gamma_L}{\Gamma_S} \sim 10^{-3} \quad (16.93)$$

The CP violation is expected to be even smaller by another factor of 10^{-3} and is more difficult to measure.

The asymmetry can be determined as follows. The transition amplitudes for Kl_3 are defined by

$$A_+ = \langle e^+ \nu \pi^- | T | K^0 \rangle, \quad \bar{A}_- = \langle e^- \nu \pi^+ | T | \bar{K}^0 \rangle \quad : \Delta S = \Delta Q \quad (16.94a)$$

$$A_- = \langle e^- \nu \pi^+ | T | K^0 \rangle, \quad \bar{A}_+ = \langle e^+ \nu \pi^- | T | \bar{K}^0 \rangle \quad : \Delta S = -\Delta Q \quad (16.94b)$$

We assume here that there are no $\Delta S = -\Delta Q$ processes and set $A_- = \bar{A}_+ = 0$. CPT invariance constrains $\bar{A}_- = (A_+)^*$ [Eq. (16.63a)]. The asymmetry of K_S , K_L decays is calculated to give

$$\begin{aligned} A_{S,L} &\equiv \frac{\Gamma(K_{S,L} \rightarrow e^+ \nu_e \pi^-) - \Gamma(K_{S,L} \rightarrow e^- \bar{\nu}_e \pi^+)}{\Gamma(K_{S,L} \rightarrow e^+ \nu_e \pi^-) + \Gamma(K_{S,L} \rightarrow e^- \bar{\nu}_e \pi^+)} \\ &= \frac{|(1 + \epsilon_{S,L})A_-|^2 - |(1 - \epsilon_{S,L})\bar{A}_+|^2}{|(1 + \epsilon_{S,L})A_-|^2 + |(1 - \epsilon_{S,L})\bar{A}_+|^2} \\ &= 2\text{Re}(\epsilon_{S,L}) = 2\text{Re}(\epsilon) \pm 2\text{Re}(\delta) \end{aligned} \quad (16.95)$$

where Eq. (16.64a) was used to derive the second equality. We retain the nomenclature $\epsilon_{S,L} = \epsilon \pm \delta$ to remind us that A_S and A_L are different, although $\delta = 0$ with CPT invariance. It is a constant independent of time if a pure $K_{S,L}$ beam is used. Under normal circumstances, the neutral kaon is produced as a mixture of K^0 and $\bar{K}^0$. Since observing the charge of a decay electron is equivalent to selecting the flavor with the $\Delta S = \Delta Q$ rule, the asymmetry shows the same time dependence as that of the strangeness oscillation. But for long time $t \gg 10/\Gamma_S$ after production, the identity of the parent does not matter and the residual asymmetry agrees with A_L given by Eq. (16.95). Figure 16.10a,b shows asymmetries of e and μ as a function of time. One can recognize interference and residual asymmetry. The asymmetry is given by

$$\begin{aligned} A_L(e) &= 3.34 \pm 0.07 \times 10^{-3} \\ A_L(\mu) &= 3.04 \pm 0.25 \times 10^{-3} \\ A_L(e + \mu) &= 3.32 \pm 0.06 \times 10^{-3} \end{aligned} \quad (16.96)$$

Note the oscillation is not a CP violation effect. It is induced by the mixing and small mass difference of the kaons.

The decay asymmetry A_S is important for deriving information on δ , but it is difficult to determine as it has to be measured immediately after production of K^0 or $\bar{K}^0$. Recently it was measured for the first time by the KLOE group [235], though not so accurately that it can be used for other analysis:

$$A_S(e) = 1.5 \pm 9.6_{\text{stat}} \pm 2.9_{\text{syst}} \times 10^{-3} \quad (16.97)$$

Problem 16.6

Show that by observing the flavor-tagged asymmetry into the same-sign semileptonic mode one can extract ϵ_S , despite the long time passage, to extract the asymmetry value:

$$\begin{aligned} A_{l\pm} &\equiv \frac{\Gamma[\bar{K}^0(t=0) \rightarrow l^\pm \nu_l \pi^\mp] - \Gamma[K^0(t=0) \rightarrow l^\pm \nu_l \pi^\mp]}{\Gamma[\bar{K}^0(t=0) \rightarrow l^\pm \nu_l \pi^\mp] + \Gamma[K^0(t=0) \rightarrow l^\pm \nu_l \pi^\mp]} \\ &\xrightarrow{t \gg \tau_S} 2\text{Re}(\epsilon_S) \end{aligned} \quad (16.98)$$

This is another method of measuring $\epsilon_S = \epsilon + \delta$.

16.3.2 ϵ and ϵ'

The two most important CP-violating parameters that play a central role in the theoretical interpretation of the neutral kaon decays are ϵ and ϵ' . ϵ is a CP-violating

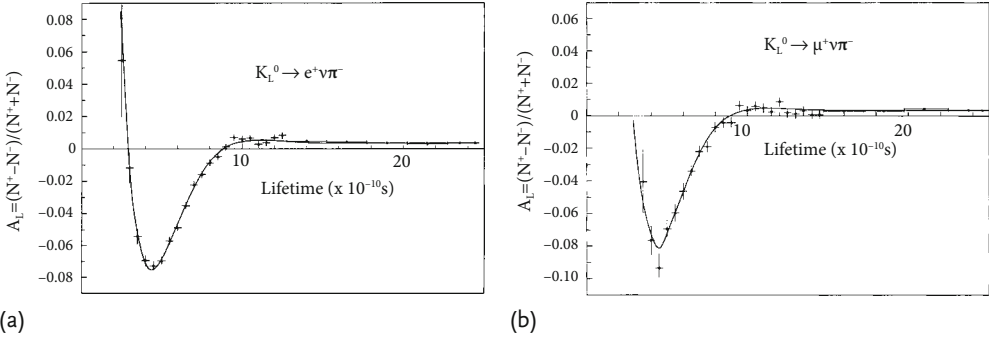


Figure 16.10 Asymmetry A_L [(a) $K_L^0 \rightarrow e^+ \nu \pi^-$, (b) $K_L^0 \rightarrow \mu^+ \nu \pi^-$] oscillates as a function of time. The residual asymmetry is the CP violation effect. The oscillation is not due to CP violation but due to mixing and mass difference [171].

mixing parameter while ϵ' , defined in Eq. (16.103c), expresses a CP violation effect in the decay. ϵ is often referred to as the indirect and ϵ' as the direct CP violation effect.⁸⁾

Phase Convention

We investigate the interplay of CP violation parameters ϵ , ϵ' , η_{+-} , and η_{00} and clarify their roles. Since the neutral kaons decay predominantly to 2π channels, we first analyze the 2π decay mode. The 2π amplitude can be decomposed to isospin states⁹⁾

$$A(K^0 \rightarrow \pi^+ \pi^-) = \frac{1}{\sqrt{3}} (\sqrt{2}A_0 + A_2) \quad (16.100a)$$

$$A(K^0 \rightarrow \pi^0 \pi^0) = \frac{1}{\sqrt{3}} (-A_0 + \sqrt{2}A_2) \quad (16.100b)$$

$$A(K^+ \rightarrow \pi^+ \pi^0) = \sqrt{\frac{3}{2}} A_2 \quad (16.100c)$$

The $I = 1$ state is antisymmetric and is not allowed by Bose statistics. The decay amplitudes, taking into account the final-state interactions, can be expressed as [see

- 8) There is yet another type, referred to as the “mixed interference” CP violation parameter, defined by

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f} \quad (16.99)$$

which appears in the interference between a decay without mixing $K^0 \rightarrow f$ and a decay with mixing $K^0 \rightarrow \bar{K}^0 \rightarrow f$. Such an effect occurs only in decays to final states that are

common to K^0 and $\bar{K}^0$. The CP-violating effect is defined by $\text{Im}(\lambda_f) \neq 0$. In the kaon decays, it is of little use but it plays an important role in neutral B-meson decays.

- 9) We deal with the symmetrized state $\pi^+ \pi^- = (\pi^+ \pi^- + \pi^- \pi^+)/\sqrt{2}$. Normally, $I = 0$ and $I = 2$ components in the neutral two-pion states are expressed as $|I = I_3 = 0\rangle = (\pi^+ \pi^- - \pi^0 \pi^0 + \pi^- \pi^+)/\sqrt{3}$.

Eq. (16.62)]

$$T_I = \langle 2\pi, I | H_W | K^0 \rangle = A_I e^{i\delta_I} \quad (16.101a)$$

$$\bar{T}_I = \langle 2\pi, I | H_W | \bar{K}^0 \rangle = \bar{A}_I e^{i\delta_I} \quad (16.101b)$$

where δ_I are scattering phase shifts of isospin I . CPT invariance means $\bar{A}_I = A_I^*$ and CP violation means $A_I \neq A_I^*$. Next we define

$$\begin{aligned} \alpha &= \frac{\langle I=0 | H_W | K_2 \rangle}{\langle I=0 | H_W | K_1 \rangle} = \frac{\langle 0 | H_W | K^0 \rangle - \langle 0 | H_W | \bar{K}^0 \rangle}{\langle 0 | H_W | K^0 \rangle + \langle 0 | H_W | \bar{K}^0 \rangle} \\ &= \frac{A_0 - \bar{A}_0}{A_0 + \bar{A}_0} = i \frac{\text{Im} A_0}{\text{Re} A_0} \end{aligned} \quad (16.102)$$

Here we have used $\langle 0 |$, $\langle 2 |$ to denote 2π , $I = 0, 2$ eigenstates. We also define CP-violating decay parameters for $I = 0, 2$:

$$\begin{aligned} \epsilon_0 &\equiv \frac{\langle I=0 | H_W | K_L \rangle}{\langle I=0 | H_W | K_S \rangle} = \frac{\langle 0 | H_W | K_2 \rangle + \epsilon \langle 0 | H_W | K_1 \rangle}{\langle 0 | H_W | K_1 \rangle + \epsilon \langle 0 | H_W | K_2 \rangle} \\ &\simeq \epsilon + \alpha = \epsilon + i \frac{\text{Im} A_0}{\text{Re} A_0} \end{aligned} \quad (16.103a)$$

$$\begin{aligned} \epsilon_2 &\equiv \frac{1}{\sqrt{2}} \frac{\langle I=2 | H_W | K_L \rangle}{\langle I=0 | H_W | K_S \rangle} = \frac{1}{\sqrt{2}} \frac{\langle 2 | H_W | K_2 \rangle + \epsilon \langle 2 | H_W | K_1 \rangle}{\langle 0 | H_W | K_1 \rangle + \epsilon \langle 0 | H_W | K_2 \rangle} \\ &\equiv \epsilon' + \epsilon_0 \omega \end{aligned} \quad (16.103b)$$

$$\epsilon' \simeq \frac{1}{\sqrt{2}} \frac{\langle 2 | H_W | K_2 \rangle}{\langle 0 | H_W | K_1 \rangle} - \alpha \omega = i \omega \left[\frac{\text{Im} A_2}{\text{Re} A_2} - \frac{\text{Im} A_0}{\text{Re} A_0} \right] \quad (16.103c)$$

$$\omega \equiv \frac{1}{\sqrt{2}} \frac{\langle 2 | H_W | K_S \rangle}{\langle 0 | H_W | K_S \rangle} \simeq \frac{1}{\sqrt{2}} \frac{\text{Re} A_2}{\text{Re} A_0} e^{i(\delta_2 - \delta_0)} \quad (16.103d)$$

where δ_0, δ_2 are 2π scattering phase shift in $I = 0, 2$ states. Note that ϵ_0 , ϵ_2 and ω (hence ϵ' also) are observables and are independent of the phase convention.

Proof:

$$\begin{aligned} \epsilon_f &= \frac{\langle f | H_W | K_L \rangle}{\langle f | H_W | K_S \rangle} = \frac{\langle f | H_W | \{p | K^0 \rangle - q | \bar{K}^0 \rangle \}}{\langle f | H_W | \{p | K^0 \rangle + q | \bar{K}^0 \rangle \}} = \frac{1 - \lambda_f}{1 + \lambda_f} \\ \lambda_f &= \left(\frac{q}{p} \right) \frac{\langle f | H_W | \bar{K}^0 \rangle}{\langle f | H_W | K^0 \rangle} = \left(\frac{q}{p} \right) \frac{\bar{A}_f}{A_f} \end{aligned} \quad (16.104)$$

By rephasing $|K^0\rangle \rightarrow e^{-i\xi} |K^0\rangle$, $|\bar{K}^0\rangle \rightarrow e^{i\xi} |\bar{K}^0\rangle$, we have $\bar{A}_f/A_f \rightarrow e^{2i\xi} \bar{A}_f/A_f$ and $q/p \rightarrow e^{-2i\xi} (q/p)$ [see discussions following Eqs. (16.81)], therefore λ_f is phase independent. $\square$

The relative amplitude ω violates the $\Delta I = 1/2$ rule. Its absolute value and phase can be obtained from the experimental data. If ρ indicates the phase space

$$|\omega|^2 \simeq \frac{\Gamma(K^+ \rightarrow \pi^+ \pi^0)}{\Gamma(K_S \rightarrow 2\pi)} \simeq \frac{2}{3} \text{BR}(K^+ \rightarrow \pi^+ \pi^0) \frac{\tau_{K_S}}{\tau_{K^+}} = (3.19 \times 10^{-2})^2 \quad (16.105a)$$

$$\text{Re}(\omega) \simeq \frac{1}{12} \left[\frac{\rho_{00}}{\rho_{+-}} \frac{\Gamma(K_S \rightarrow \pi^+ \pi^-)}{\Gamma(K_S \rightarrow \pi^0 \pi^0)} - 2 \right] \quad (16.105b)$$

From the above relations we obtain $\arg(\omega)_{\text{observed}} \simeq \pm 55.6^\circ$, which is in reasonable agreement with that given by Eq. (16.103d), namely [120, 157, 344]

$$\delta_2 - \delta_0 = -47.6 \pm 1.5^\circ \quad (16.106)$$

Since the phase of $|K^0\rangle, |\bar{K}^0\rangle$ can be chosen arbitrarily [Eq. (16.17)], we are free to change ϵ . This can be shown as follows. By rephasing

$$\begin{aligned} \epsilon &= \frac{\mathcal{A}_{21} - \mathcal{A}_{12}}{2\Delta_\lambda} \rightarrow \frac{\mathcal{A}_{21}e^{-2i\xi} - \mathcal{A}_{12}e^{2i\xi}}{2\Delta_\lambda} \\ &\simeq \frac{\mathcal{A}_{21}(1 - 2i\xi) - \mathcal{A}_{12}(1 + 2i\xi)}{2\Delta_\lambda} \simeq \epsilon + i\xi \end{aligned} \quad (16.107)$$

where we have used the fact that CP violation is a small effect [i.e. $O(\epsilon\Delta_\lambda)$] and only needs a small phaseshift to redefine ϵ . Using this freedom, we can change ϵ to a more convenient form. Choosing $\xi = \text{Im}A_0/\text{Re}A_0$, we can make ϵ equal to ϵ_0 . Namely, the CP-violating decay parameter and the mixing parameter are the same, a convention we adopt from now on.¹⁰⁾ In this case one has to bear in mind that $\bar{K}^0$ is defined by

$$\text{CP}|K^0\rangle = e^{2i\xi}|\bar{K}^0\rangle \quad (16.108)$$

Note that in many discussions outside the realm of neutral kaon decays, including those about the Kobayashi–Maskawa matrix, $\xi = 0$ is assumed. So one must be careful about the phase convention. Since we chose to set $\epsilon = \epsilon_0$, which is a phase-convention-independent parameter, we shall use $\tilde{\epsilon}$ when we need the phase-dependent mixing parameter. Note, many authors use ϵ to denote both the phase-independent $I = 0$ decay amplitude and the phase-dependent mixing parameter interchangeably.

Then ϵ defined in Eq. (16.69) is modified to

$$\epsilon = \frac{\sin\phi_{\text{SW}}e^{i\phi_{\text{SW}}}}{\Delta m} \left[-\text{Im}(M_{12}) + i\frac{\text{Im}(\Gamma_{12})}{2} \right] + i\frac{\text{Im}A_0}{\text{Re}A_0} \quad (16.109)$$

10) Historically, Wu–Yang’s phase choice [393] to eliminate $i\text{Im}A_0/\text{Re}A_0$ was adopted, and for this phase choice $\epsilon = \epsilon_0$. We use the same ϵ but $\text{Im}A_0/\text{Re}A_0$ is retained explicitly.

Using

$$i(\sin \phi_{\text{SW}} e^{i\phi_{\text{SW}}})^{-1} = 1 + i \cos \phi_{\text{SW}} / \sin \phi_{\text{SW}} = 1 + i \Delta \Gamma / 2 \Delta m \quad (16.110)$$

ϵ can be rewritten as

$$\begin{aligned} \epsilon &= \sin \phi_{\text{SW}} e^{i\phi_{\text{SW}}} \left[-\frac{\text{Im}(M_{12})}{\Delta m} - \frac{\text{Im}(\Gamma_{12})}{\Delta \Gamma} \right] + i \left(\frac{\text{Im}(\Gamma_{12})}{\Delta \Gamma} + \frac{\text{Im}A_0}{\text{Re}A_0} \right) \\ &\equiv \epsilon_T + i \delta \varphi \end{aligned} \quad (16.111a)$$

$$\delta \varphi \equiv \frac{\text{Im}(\Gamma_{12})}{\Delta \Gamma} + \frac{\text{Im}A_0}{\text{Re}A_0} \simeq \frac{1}{2} [\arg(\Gamma_{12}) - \arg(A_0^* \bar{A}_0)] \quad (16.111b)$$

We shall show that $\delta \varphi$ is a very small number.

Experimental data show

$$\begin{aligned} \text{BR}(K_S \rightarrow 2\pi + (\gamma)) &= 99.89\%, \\ \text{BR}(K_S \rightarrow 3\pi) &= 3.5 \times 10^{-7}, \\ \text{BR}(K_S \rightarrow e^\pm \nu_e \pi^\mp) &= 7.04 \times 10^{-4}, \\ \text{BR}(K_S \rightarrow \mu^\pm \nu_\mu \pi^\mp) &= 4.69 \times 10^{-4} \\ \text{All other decay modes} &< 10^{-5} \end{aligned} \quad (16.112)$$

One notices that 2π channel including photon emissions is overwhelmingly the dominant decay mode. Since $\Gamma_S \gg \Gamma_L$, $\Gamma_S \simeq \Gamma_{2\pi}$, it is a good approximation to use only the 2π intermediate state in the evaluation of Γ_{12} . Note that the leptonic mode ($K \rightarrow l \nu_l \pi$) does not contribute to Γ_{12} if the $\Delta S = \Delta Q$ rule is exact. Setting the phase space volume as ρ and using Eqs. (16.11) and (16.101), we have

$$\begin{aligned} \text{Im}(\Gamma_{12}) &\simeq \left[\text{Im}(\langle K^0 | H_W | I = 0 \rangle \langle I = 0 | H_W | \bar{K}^0 \rangle) + (I = 2 \text{ term}) \right] \rho \\ &= \text{Im}(A_0^* \bar{A}_0 + A_2^* \bar{A}_2) \rho \stackrel{\text{CPT}}{=} \text{Im}[(A_0^*)^2 + (A_2^*)^2] \rho \\ &= -2(\text{Re}A_0 \text{Im}A_0 + \text{Re}A_2 \text{Im}A_2) \rho \end{aligned} \quad (16.113)$$

$$\begin{aligned} \Gamma_S &\simeq \left[|\langle I = 0 | H_W | K^0 + \bar{K}^0 \rangle / \sqrt{2}|^2 + (I = 2 \text{ term}) \right] \rho \\ &\simeq 2[(\text{Re}A_0)^2 + (\text{Re}A_2)^2] \rho \end{aligned} \quad (16.114)$$

$$\begin{aligned} \therefore \frac{\text{Im}(\Gamma_{12})}{\Delta \Gamma} &\simeq \frac{\text{Im}(\Gamma_{12})}{\Gamma_S} \simeq -\frac{\text{Im}A_0}{\text{Re}A_0} + 2\omega^2 \left[\frac{\text{Im}A_0}{\text{Re}A_0} - \frac{\text{Im}A_2}{\text{Re}A_2} \right] \\ &\stackrel{(16.103c)}{=} -\frac{\text{Im}A_0}{\text{Re}A_0} \pm 2|\omega \epsilon'| \end{aligned} \quad (16.115)$$

Using experimental values [$\omega \sim 0.03$, $|\epsilon'|/\epsilon| \simeq 10^{-3}$; see Eq. (16.183)], we see that the second term is at most $\sim 10^{-5}$ of the first term and can be neglected. This proves our promised statement.

We see that ϵ_T is completely aligned along the superweak phase. Using $\Delta m \simeq -2\text{Re}(M_{12})$, $\Delta\Gamma \simeq 2\text{Re}(\Gamma_{12})$ [see Eq. (16.82)], ϵ_T can further be rewritten as

$$\epsilon_T = \sin \phi_{\text{SW}} e^{i\phi_{\text{SW}}} \left[\frac{-\arg(M_{12}^* \Gamma_{12})}{2} \right] \simeq \sin \phi_{\text{SW}} e^{i\phi_{\text{SW}}} \left[\frac{2\text{Im}(M_{12}^* \Gamma_{12})}{\Delta m \Delta\Gamma} \right] \quad (16.116)$$

where $\Delta m \Delta\Gamma = -4\text{Re}(M_{12}^* \Gamma_{12})$ has been used in obtaining the last equality. Note, T invariance means $\text{Im}(M_{12}^* \Gamma_{12}) = 0$ [see Eq. (16.52c)]. Of course it should be as T and CP are equivalent under the assumption of CPT invariance. But it is good to remember that ϵ_T , if defined exactly by the second equality of Eq. (16.116), is intrinsically a T violation parameter and will reappear later in Eq. (16.148) in the discussion of T invariance.

Conventionally, throughout most literatures, ϵ defined by Eq. (16.111) is used. However, as $\delta\varphi$ is small, the following discussions are valid irrespective of whether ϵ or ϵ_T are used, except in trying to determine the value of $\delta\varphi$ itself. Note also, that ϵ itself is a phase dependent variable, but its real part $\text{Re}(\epsilon)$ is an observable and is phase independent as can be seen in the expression of the decay asymmetry parameter A_L [see Eq. (16.98)].

Evaluation of ϵ

To determine ϵ and ϵ' , we express them in terms of observables. We relate the 2π decay amplitudes η_{+-} , η_{00} to the isospin parameters ϵ and ϵ' .

Using the inverse relation of Eq. (16.100) we have

$$\begin{aligned} \langle I = 0 | H_W | K_L \rangle &= \sqrt{\frac{2}{3}} \langle + - | H_W | K_L \rangle - \sqrt{\frac{1}{3}} \langle 00 | H_W | K_L \rangle \\ &= \sqrt{\frac{2}{3}} \eta_{+-} \left[\sqrt{\frac{2}{3}} \langle I = 0 | H_W | K_S \rangle + \sqrt{\frac{1}{3}} \langle I = 2 | H_W | K_S \rangle \right] \\ &\quad - \sqrt{\frac{1}{3}} \eta_{00} \left[-\sqrt{\frac{1}{3}} \langle I = 0 | H_W | K_S \rangle + \sqrt{\frac{2}{3}} \langle I = 2 | H_W | K_S \rangle \right] \end{aligned} \quad (16.117)$$

and dividing by $\langle I = 0 | H_W | K_S \rangle$ we obtain

$$\epsilon = \left(\frac{2}{3} \eta_{+-} + \frac{1}{3} \eta_{00} \right) + \frac{2}{3} (\eta_{+-} - \eta_{00}) \omega \quad (16.118)$$

Similarly

$$\epsilon_2 = \left(\frac{1}{3} \eta_{+-} - \frac{1}{3} \eta_{00} \right) + \frac{1}{3} (\eta_{+-} + 2\eta_{00}) \omega \quad (16.119)$$

Conversely

$$\eta_{+-} = \frac{\epsilon + \epsilon_2}{1 + \omega} \simeq \epsilon + \epsilon', \quad \eta_{00} = \frac{\epsilon - 2\epsilon_2}{1 - 2\omega} \simeq \epsilon - 2\epsilon' \quad (16.120)$$

As $\epsilon' \ll \epsilon$ experimentally [see Eq. (16.183)],

$$\frac{\epsilon'}{\epsilon} \simeq \frac{1}{3} \left(1 - \frac{\eta_{00}}{\eta_{+-}} \right) \quad (16.121)$$

is a good first approximation. Figure 16.11 shows a sketch of Eqs. (16.120).

Experimentally, using Eq. (16.89),

$$|\eta_{+-}| \simeq |\eta_{00}|, \quad \phi_{+-} \simeq \phi_{00} = 43.51 \pm 0.05^\circ \simeq \phi_{SW} \quad (16.122)$$

which means

$$\epsilon \simeq \eta_{+-} \simeq \eta_{00}, \quad (16.123)$$

At this stage, it is convenient to define components parallel and perpendicular to $e^{i\phi_{SW}}$ in the complex plane. Then

$$\epsilon = \epsilon_{\parallel} e_{\parallel} + \epsilon_{\perp} e_{\perp} \quad (16.124a)$$

$$e_{\parallel} = e^{i\phi_{SW}} = \cos \phi_{SW} + i \sin \phi_{SW} \quad (16.124b)$$

$$e_{\perp} = i e^{i\phi_{SW}} = -\sin \phi_{SW} + i \cos \phi_{SW} \quad (16.124c)$$

We will learn later that the components of ϵ, ϵ' perpendicular to $e_{\parallel}$ are related to CPT violation. The experimental observations Eq. (16.123) mean

$$|\epsilon_{\perp}| \ll |\epsilon_{\parallel}| \quad (16.125a)$$

and that ϵ can be determined very accurately by the relations

$$\begin{aligned} |\epsilon| &= (2|\eta_{+-}| + |\eta_{00}|)/3 \\ \phi_{\epsilon} &= (2\phi_{+-} + \phi_{00})/3 \end{aligned} \quad (16.125b)$$

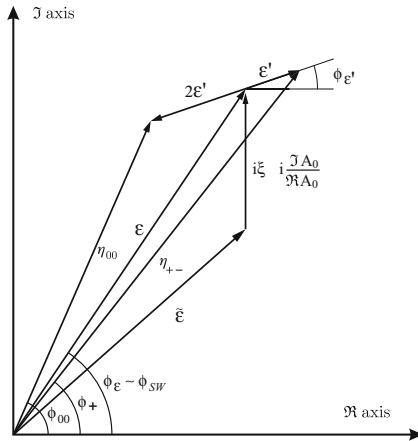


Figure 16.11 Sketch of Eq. (16.120). Not to scale. Since we redefined $\tilde{\epsilon}$ as the mixing parameter, it is phase dependent, but $\epsilon = \epsilon_0$ is not.

which give

$$\begin{aligned} |\epsilon| &= 2.229 \pm 0.012 \times 10^{-3} \\ \phi_\epsilon &= 43.51 \pm 0.05^\circ \end{aligned} \quad (16.126)$$

Direct CP Violation

While ϵ represents an indirect CP violation in mixing, ϵ' is a direct CP violation in decay amplitudes to the $I = 0, 2$ state. To see that ϵ' represents a CP violation in the decay, let us find a necessary condition for the two CP conjugate decay rates to be different. Note, the total decay rate is guaranteed to be the same by CPT invariance, but the partial decay rate to a particular final state f can be different if CP is violated. The partial decay amplitude is expressed as [see Eq. (16.57)]

$$T(\bar{B} \rightarrow \bar{f}) = \sum_j \bar{A}_j e^{i\delta_j} \quad (16.127a)$$

$$T(B \rightarrow f) = \sum_j A_j e^{i\delta_j} \quad (16.127b)$$

where j labels different intermediate states and δ_j is the scattering phase shift. CPT invariance constrains $\bar{A}_j = A_j^*$ and T invariance means $A_j = A_j^*$. Therefore, under the assumption of CPT invariance, CP violation appears as a phase of the amplitude. Let us write

$$\bar{A}_j = A_j^* = |A_j| e^{i\phi_{w_j}} = a_j e^{i\phi_{w_j}}, \quad A_j = a_j e^{-i\phi_{w_j}} \quad (16.128)$$

For the experimental sign of CP violation in the decay

$$\delta\Gamma = \Gamma(\bar{B} \rightarrow \bar{f}) - \Gamma(B \rightarrow f) \neq 0 \quad (16.129)$$

the interference effect is necessary, which requires at least two intermediate states. Writing

$$\begin{aligned} T(\bar{B} \rightarrow \bar{f}) &= a_1 e^{i\delta_1} e^{i\phi_{w_1}} + a_2 e^{i\delta_2} e^{i\phi_{w_2}} \\ T(B \rightarrow f) &= a_1 e^{i\delta_1} e^{-i\phi_{w_1}} + a_2 e^{i\delta_2} e^{-i\phi_{w_2}} \end{aligned} \quad (16.130)$$

then

$$\Gamma(\bar{B} \rightarrow \bar{f}) - \Gamma(B \rightarrow f) = a_1 a_2 \sin(\phi_{w_1} - \phi_{w_2}) \sin(\delta_1 - \delta_2) \quad (16.131)$$

The above equation means that the two states must have both different strong phase and weak phase.

Next we want to express the decay rate difference in terms of the CP violation parameter η_f that we defined previously [Eq. (16.83)].

Problem 16.7

Show the decay asymmetry at time t is given by

$$A_f \equiv \frac{\Gamma(\bar{K}^0 \rightarrow f) - \Gamma(K^0 \rightarrow f)}{\Gamma(\bar{K}^0 \rightarrow f) + \Gamma(K^0 \rightarrow f)} = 2\text{Re}(\epsilon - \eta_f) \quad (16.132)$$

$$\eta_f \equiv \frac{\langle f | H_W | K_L \rangle}{\langle f | H_W | K_S \rangle} \simeq \epsilon + i \frac{\text{Im} A_f}{\text{Re} A_f}$$

which is a restatement that if the final state is reached via a single route, the asymmetry does not appear. Show also

$$A_f(\pi^0 \pi^0) = 2(\epsilon - \eta_{00}) = 4\epsilon', \quad A_f(\pi^+ \pi^-) = 2(\epsilon - \eta_{+-}) = -2\epsilon' \quad (16.133)$$

This explicitly shows that ϵ' represents the difference of two CP conjugate decay rates.

Note, the CP violation parameter ϵ' defined in Eq. (16.103c) can be expressed in the form

$$\begin{aligned} \text{Re}(\epsilon') &= \text{Re} \left(i \omega \left[\frac{\text{Im} A_2}{\text{Re} A_2} - \frac{\text{Im} A_0}{\text{Re} A_0} \right] \right) \\ &\simeq -\frac{1}{\sqrt{2}} \frac{\text{Re} A_2}{\text{Re} A_0} \sin(\delta_2 - \delta_0) \sin(\phi_2 - \phi_0) \end{aligned} \quad (16.134)$$

which is of the form Eq. (16.131) and shows that $I = 0, 2$ states provide the two different routes for decays $K \rightarrow \pi^+ \pi^-$ and $K \rightarrow \pi^0 \pi^0$.

Problem 16.8

Show that ϵ' can also be defined as

$$\text{Re}(\epsilon') = \frac{1}{6} \left(\left| \frac{\bar{A}_{\pi^0 \pi^0}}{A_{\pi^0 \pi^0}} \right| - \left| \frac{\bar{A}_{\pi^+ \pi^-}}{A_{\pi^+ \pi^-}} \right| \right) \quad (16.135)$$

Evaluation of ϵ'

From Eq. (16.120)

$$\epsilon' \simeq \frac{1}{3}(\eta_{+-} - \eta_{00}) \quad (16.136)$$

It is convenient to use the decomposition defined in Eq. (16.124):

$$\frac{\epsilon'_\parallel}{\epsilon} = \text{Re}(\epsilon'/\epsilon), \quad \frac{\epsilon'_\perp}{\epsilon} = \text{Im}(\epsilon'/\epsilon) \quad (16.137)$$

Since

$$\arg(\epsilon') = \arg(\omega) = \frac{\pi}{2} + \delta_2 - \delta_0 = 42.3 \pm 1.5^\circ \simeq \phi_{\text{sw}} \quad (16.138)$$

we have

$$\epsilon'_\perp \ll |\epsilon'_\parallel| \quad (16.139)$$

Using $\eta_{+-} \simeq \eta_{00}$, $\phi_{+-} \simeq \phi_{00} \simeq \phi_{\text{sw}}$

$$\epsilon'_\perp \simeq \frac{|\eta_{+-}|}{3}(\phi_{+-} - \phi_{00}) = -\frac{|\eta_{+-}|}{3}\Delta\phi \quad (16.140)$$

When both η_{+-} and η_{00} are measured it is possible to determine $\epsilon'_\parallel$ in principle, but with poor accuracy. Experimentally $\text{Re}(\epsilon'/\epsilon)$ can be determined directly [51, 244]. Using Eq. (16.120)

$$\text{Re}\left(\frac{\epsilon'}{\epsilon}\right) \simeq \frac{1}{6} \left(\left| \frac{\eta_{+-}}{\eta_{00}} \right|^2 - 1 \right) \simeq \frac{1}{3} \left(\left| \frac{\eta_{+-}}{\eta_{00}} \right| - 1 \right) = 1.65 \pm 0.26 \times 10^{-3} \quad (16.141)$$

where we have used $|\eta_{+-}/\eta_{00}|^2 \simeq 1$ to derive the third equality. $\Delta\phi$ can also be determined accurately using the value of $\text{Re}(\epsilon'/\epsilon)$. Since $\phi_{00}/\phi_{+-} \simeq 1$, Eq. (16.120) means

$$\begin{aligned} \Delta\phi &= \phi_{00} - \phi_{+-} \simeq -3\text{Im}(\epsilon'/\epsilon) \simeq -3\text{Re}\left(\frac{\epsilon'}{\epsilon}\right) \tan(\phi_{\epsilon'} - \phi_{\text{sw}}) \\ &\simeq 0.006 \pm 0.008^\circ \quad [311] \end{aligned} \quad (16.142)$$

16.4

Test of T and CPT Invariance¹¹⁾

We have observed CP violation, but if CPT invariance holds this is equivalent to T violation. The foundation on which CPT invariance is established is very solid, as stated in Chap. 9. It can be proved with very few axioms, such as the principles of quantum mechanics and Lorentz invariance. Namely, to deny CPT invariance is equivalent to denying quantum field theory. One wonders if the observed CP violation is T violation with CPT invariance or genuine CP violation, which holds regardless of CPT invariance.

The neutral kaon system provides test benches for the most rigorous determination of T and CPT symmetry and a variety of other fundamental principles. This is because the mass difference ($\Delta m \simeq 3.5 \times 10^{-6}$ eV) is so small that its effect appears as macroscopic interference phenomena and enables CP and CPT parameters to be determined with an accuracy no other means can provide. In this section, we start with no assumption about CP or CPT invariance.

¹¹⁾ See [48, 68, 121, 269, 299].

16.4.1

Definition of T- and CPT-Violating Amplitudes

The starting point is Eq. (16.55). We first write the decay amplitudes for $K \rightarrow 2\pi$ and $Kl3$ are expressed as sum of two amplitudes:

$$\langle 2\pi, I | H_W | K^0 \rangle = (A_I + B_I) e^{i\delta_I} \quad (16.143a)$$

$$\langle 2\pi, I | H_W | \bar{K}^0 \rangle = (\bar{A}_I + \bar{B}_I) e^{i\delta_I} \quad (16.143b)$$

$$\langle e^+ \nu_e \pi^- | H_W | K^0 \rangle = (a + b) \quad (16.143c)$$

$$\langle e^- \bar{\nu}_e \pi^+ | H_W | \bar{K}^0 \rangle = (\bar{a} + \bar{b}) \quad (16.143d)$$

If CPT invariance holds, $\bar{A} + \bar{B} = (A + B)^*$ [Eq. (16.63a)], so if we rewrite $\bar{A}_I = A_I^*$, $\bar{a} = a^*$, $\bar{B}_I = -B_I^*$, $\bar{b} = -b^*$, A_I, a represent CPT-conserving and B_I, b represent CPT-violating terms. Furthermore, if we write

$$\begin{aligned} A_I &= \text{Re} A_I + i \text{Im} A_I, & B_I &= \text{Re} B_I + i \text{Im} B_I \\ a &= \text{Re} a + i \text{Im} a, & b &= \text{Re} b + i \text{Im} b \end{aligned} \quad (16.144)$$

T invariance means $A_I^* = A_I$, $B_I^* = B_I$ and CP invariance means $\bar{A}_I = A_I$, $\bar{B}_I = B_I$. Consequently, what each term represents is

Re A_I	CPT,	T,	CP
Im A_I	CPT,	$\nexists$,	$\mathcal{CP}$
Re B_I	$\mathcal{CP}$,	T,	$\mathcal{CP}$
Im B_I	$\mathcal{CP}$,	$\nexists$,	CP

If we can demonstrate smallness of both direct and indirect CPT violation parameters, i.e., $|B_I| \ll |A_I|$, $\delta \ll \epsilon$, the assumption of CPT invariance is justified. Similar logic can be applied to a, b .

16.4.2

T Violation

If observables such as transverse polarization ($\sigma \cdot p_1 \times p_2$) in the 3-body decay $K^0 \rightarrow \pi^\pm(p_1)\mu^\mp(p_2)\nu_\mu$ or the electric dipole moment (EDM) are detected, they are a sure sign of T violations [see Sect. 9.2.2] independent of CP or CPT violation. Unfortunately, they have not been observed and there is no other evidence that T invariance is violated. Therefore we investigate whether we can show it in the neutral kaon system. T invariance means [19]

$$T \left[\langle \Psi^{\text{out}}(t_f) | \Phi^{\text{in}}(t_i) \rangle \right] = \langle \Phi^{\text{out}}(t_f) | \Psi^{\text{in}}(t_i) \rangle \quad (16.145)$$

Setting $\Phi = K^0$, $\Psi = \bar{K}^0$, we have

$$|\langle \bar{K}_0^{\text{out}}(t_f) | K_0^{\text{in}}(t_i) \rangle|^2 = |\langle K_0^{\text{out}}(t_f) | \bar{K}_0^{\text{in}}(t_i) \rangle|^2 \quad (16.146)$$

Therefore, time reversal violation can manifest itself in an asymmetry parameter known as the Kabir asymmetry [221]

$$\begin{aligned}
 A_T &= \frac{P[\bar{K}^0(t=0) \rightarrow K^0(t)] - P[K^0(t=0) \rightarrow \bar{K}^0(t)]}{P[\bar{K}^0(t=0) \rightarrow K^0(t)] + P[K^0(t=0) \rightarrow \bar{K}^0(t)]} \\
 &= \frac{|\langle K_0 | e^{-i\mathcal{H}t} | \bar{K}_0 \rangle|^2 - |\langle \bar{K}_0 | e^{-i\mathcal{H}t} | K_0 \rangle|^2}{|\langle K_0 | e^{-i\mathcal{H}t} | \bar{K}_0 \rangle|^2 + |\langle \bar{K}_0 | e^{-i\mathcal{H}t} | K_0 \rangle|^2} \\
 &= \frac{|2(\mathcal{A}_{12}/\Delta_\lambda)g_-(t)A_f|^2 - |2(\mathcal{A}_{21}/\Delta_\lambda)g_-(t)\bar{A}_{\bar{f}}|^2}{|2(\mathcal{A}_{12}/\Delta_\lambda)g_-(t)A_f|^2 + |2(\mathcal{A}_{21}/\Delta_\lambda)g_-(t)\bar{A}_{\bar{f}}|^2} \\
 &\simeq \frac{|\mathcal{A}_{12}|^2\{1 - 2\text{Re}(\gamma)\} - |\mathcal{A}_{21}|^2\{1 + 2\text{Re}(\gamma)\}}{|\mathcal{A}_{12}|^2\{1 - 2\text{Re}(\gamma)\} + |\mathcal{A}_{21}|^2\{1 + 2\text{Re}(\gamma)\}} \\
 &\simeq \frac{|\mathcal{A}_{12}|^2 - |\mathcal{A}_{21}|^2}{|\mathcal{A}_{12}|^2 + |\mathcal{A}_{21}|^2} - 2\text{Re}(\gamma) \tag{16.147a}
 \end{aligned}$$

$$A_f = \langle f | H_W | K^0 \rangle = a + b = a(1 - \gamma)$$

$$\bar{A}_{\bar{f}} = \langle \bar{f} | H_W | \bar{K}^0 \rangle = a^* - b^* = a^*(1 + \gamma^*), \quad \gamma = -b/a \tag{16.147b}$$

where f and $\bar{f}$ are decay final states to tag the flavor. We have used Eqs. (16.79b) and (16.80b) in deriving the third line. No assumption on CPT invariance is made here. Apart from CPT violating parameter $\text{Re}(\gamma)$ which can be tested separately, T violation means $|\mathcal{A}_{12}| \neq |\mathcal{A}_{21}|$, which confirms the statement in Eq. (16.52b).

ϵ_T

A convenient parameter can be defined as

$$\begin{aligned}
 \epsilon_T &\equiv \frac{i}{\Delta\Gamma} \frac{|\mathcal{A}_{12}|^2 - |\mathcal{A}_{21}|^2}{\lambda_L - \lambda_S} = \sin\phi_{\text{sw}} e^{i\phi_{\text{sw}}} \frac{2\text{Im}(M_{12}^* \Gamma_{12})}{\Delta m \Delta\Gamma} \\
 &= \sin\phi_{\text{sw}} e^{i\phi_{\text{sw}}} \frac{1}{2} \sin\{\arg(\Gamma_{12}) - \arg(M_{12})\} \tag{16.148}
 \end{aligned}$$

where ϕ_{sw} , Δm and $\Delta\Gamma$ are defined in Eqs. (16.81). ϵ_T is related to ϵ by $\epsilon = \epsilon_T + i\delta\varphi$ [see Eq. (16.111)¹²⁾]. Using this ϵ_T , the asymmetry A_T is expressed as

$$A_T = 4\text{Re}(\epsilon_T) - 2\text{Re}(\gamma) \tag{16.149}$$

The value of $\text{Re}(\gamma)$ can be evaluated from the asymmetry of semileptonic decays [see Eq. (16.155)] and is known to be consistent with zero. In the analysis here, it was assumed to vanish.

¹²⁾ Note, if ϵ_T is defined by Eq. (16.148), $\delta\varphi$ will be redefined by this relation.

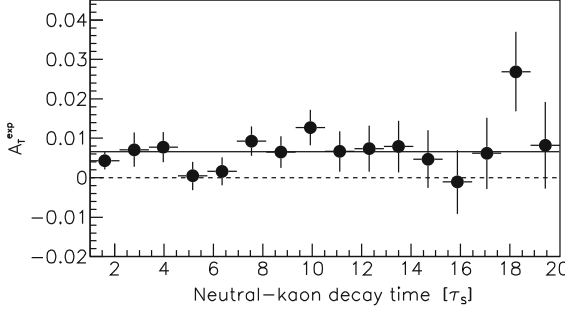


Figure 16.12 Direct T-violation effect observed by CPLEAR [107, 108].

$\epsilon_T \neq 0$ means T violation whether or not CPT invariance is violated. Note A_T is constant and does not depend on time. $K^0 \rightarrow \bar{K}^0$, $\bar{K}^0 \rightarrow K^0$ processes can be identified using lepton charge tagging, i.e. K^0 by $e^+ \nu_e \pi^-$ and $\bar{K}^0$ by $e^- \bar{\nu}_e \pi^+$, assuming the strict $\Delta S = \Delta Q$ rule.¹³⁾ Figure 16.12 shows the CPLEAR measurement [107, 108] of the asymmetry A_T , which is clearly nonzero.¹⁴⁾

This is the first, and so far only, directly observed T-violation effect. The value is determined as

$$4\text{Re}(\epsilon_T) = 6.2 \pm 1.4_{\text{stat}} \pm 1.0_{\text{syst}} \times 10^{-3} \quad (16.150)$$

16.4.3

CPT violation

Re(δ)

CPT invariance requires $\mathcal{A}_{11} = \mathcal{A}_{22}$ [see Eq. (16.15a)], or equivalently $\delta = 0$. Therefore an observable similar to A_T that can test CPT violation is given by

$$\begin{aligned} A_{\text{CPT}} &= \frac{P[\bar{K}^0(t=0) \rightarrow \bar{K}^0(t)] - P[K^0(t=0) \rightarrow K^0(t)]}{P[\bar{K}^0(t=0) \rightarrow \bar{K}^0(t)] + P[K^0(t=0) \rightarrow K^0(t)]} \\ &= \frac{|\langle \bar{K}_0 | e^{-i\mathcal{A}t} | \bar{K}_0 \rangle|^2 - |\langle K_0 | e^{-i\mathcal{A}t} | K_0 \rangle|^2}{|\langle \bar{K}_0 | e^{-i\mathcal{A}t} | \bar{K}_0 \rangle|^2 + |\langle K_0 | e^{-i\mathcal{A}t} | K_0 \rangle|^2} \end{aligned} \quad (16.151)$$

If the final state is tagged by charge detection of the decay lepton, the leptonic decay amplitudes in Eq. (16.94) are accompanied by the direct CPT-violating amplitude $b = -\gamma a$ and are expressed as

$$\langle \bar{K}^0 | e^{-i\mathcal{A}t} | \bar{K}_0 \rangle = a^*(1 + \gamma^*)[g_+(t) + 2\delta g_-(t)] \quad (16.152a)$$

$$\langle K_0 | e^{-i\mathcal{A}t} | K_0 \rangle = a(1 - \gamma)[g_+(t) - 2\delta g_-(t)] \quad (16.152b)$$

13) Experimental analyses are usually carried out without assuming the $\Delta S = \Delta Q$ rule, but so far no effect has been observed.

14) Because of different normalization, what CPLEAR actually measured is $A_T = 4\text{Re}(\epsilon) - 4\text{Re}(\gamma)$.

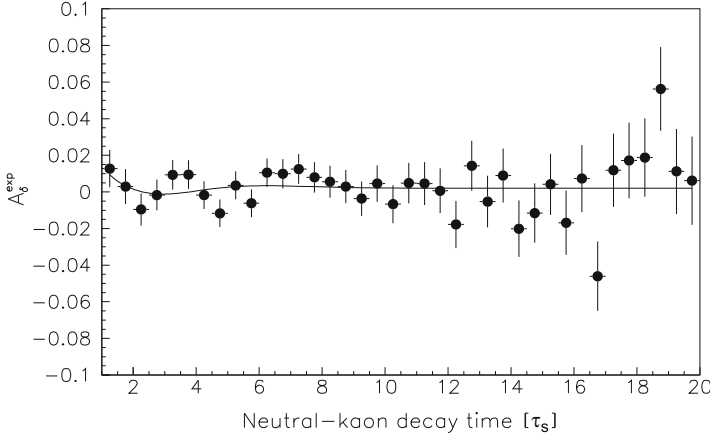


Figure 16.13 Asymmetry A_{δ}^{exp} versus the neutral-kaon decay time in units of τ_S . The solid line represents the result of the fit [107].

$$\begin{aligned}
 \therefore A_{\text{CPT}} &= \frac{|(1+\gamma)^*[g_+(t) + 2\delta g_-(t)]|^2 - |(1-\gamma)[g_+(t) - 2\delta g_-(t)]|^2}{|(1+\gamma)^*[g_+(t) + 2\delta g_-(t)]|^2 + |(1-\gamma)[g_+(t) - 2\delta g_-(t)]|^2} \\
 &= 2\text{Re}(\gamma) + 4 \frac{\text{Re}(\delta) \sinh \Delta \Gamma t/2 + \text{Im}(\delta) \sin \Delta m t}{\cosh \Delta \Gamma t/2 + \cos \Delta m t} \\
 &\xrightarrow{t \gg \tau_S} 4\text{Re}(\delta) + 2\text{Re}(\gamma) \quad (16.152c)
 \end{aligned}$$

$\text{Re}(\delta)$ and $\text{Re}(\gamma)$ can be determined by observing residual asymmetry. Figure 16.13 shows CPLEAR's experimental data on A_{CPT} .¹⁵⁾ The data show no sign of CPT violation.

The obtained values are [107]

$$\text{Re}(\delta) = 2.9 \pm 2.6_{\text{stat}} \pm 0.6_{\text{syst}} \times 10^{-4} \quad (16.153)$$

$\text{Re}(\gamma)$

The value of $\text{Re}(\gamma)$ can be determined using the asymmetry of semileptonic decays. Including $\text{Re}(\gamma)$ in the expression for the asymmetry Eq. (16.95) we obtain

$$A_{S,L} = 2[\text{Re}(\epsilon) \pm \text{Re}(\delta) - \text{Re}(\gamma)] \quad (16.154a)$$

$$\therefore A_S + A_L = 4[\text{Re}(\epsilon) - \text{Re}(\gamma)] \quad (16.154b)$$

A_S was measured in KLOE [235]. However, the most accurate value of $\text{Re}(\gamma)$ comes from the analysis of the Bell–Steinberger relation and is given by

$$\text{Re}(\gamma) = 0.4 \pm 2.5 \times 10^{-3} \quad (16.155)$$

15) Actually what they measured is A_{δ} , which enhances the sensitivity to δ and compensates $\text{Re}(\gamma)$, i.e. $[A_{\delta} \sim 8\text{Re}(\delta)]$. It differs from A_{CPT} by the normalization method. For details see [68, 107].

$\text{Im}(\delta)$ can also be determined in principle from $Kl3$ decay asymmetry by observing its time variation, as is described in Eq. (16.152c), but again its most accurate value can be determined by using the Bell–Steinberger relation.

Analysis Using the Bell–Steinberger Relation*

The Bell–Steinberger relation Eq. (16.74) can be expressed as

$$2 \left(\bar{\Gamma} + i \Delta m \right) [\text{Re}(\epsilon) - i \text{Im}(\delta)] = \sum_f A_L(f) A_S^*(f) \rho(f) \quad (16.156a)$$

$$= \sum_{f=2\pi} \eta_{2\pi} \text{BR}(K_S \rightarrow \pi\pi) \Gamma_S + \sum_{f=3\pi} \eta_{3\pi}^* \text{BR}(K_L \rightarrow 3\pi) \Gamma_L \\ + 2[\text{Re}(\epsilon) - \text{Re}(\gamma) - i \text{Im}(\delta)] \text{BR}(K_L \rightarrow l\nu\pi) \Gamma_L \quad (16.156b)$$

$$\eta_{2\pi} = \frac{\langle 2\pi | H_W | K_L \rangle}{\langle 2\pi | H_W | K_S \rangle}, \quad \eta_{3\pi} = \frac{\langle 3\pi | H_W | K_S \rangle}{\langle 3\pi | H_W | K_L \rangle} \quad (16.156c)$$

In the neutral kaon system only a few modes give significant contributions to the rhs of Eq. (16.156a) [see Eq. (16.112)]. Only the $\pi\pi(\gamma)$, 3π and $\pi l \nu_l$ modes turn out to be relevant up to 10^{-7} level [238]. The second line for the $Kl3$ intermediate states derives from the following calculation:

$$\begin{aligned} & \langle K_S | H_W | e^+ \pi^- \nu \rangle \langle e^+ \pi^- \nu | H_W | K_L \rangle \\ & + \langle K_S | H_W | e^- \pi^+ \nu \rangle \langle e^- \pi^+ \nu | H_W | K_L \rangle \rho \\ & = \frac{1}{2} \left[(1 + \epsilon_S^*) (1 + \epsilon_L) |A(K^0 \rightarrow e^+ \pi^- \nu)|^2 \right. \\ & \quad \left. - (1 - \epsilon_S^*) (1 - \epsilon_L) |A(\bar{K}^0 \rightarrow e^- \pi^+ \nu)|^2 \right] \rho \\ & = \Gamma(K_L \rightarrow l\nu\pi) \frac{1}{2} \left[(1 + \epsilon_S^*) (1 + \epsilon_L) \{1 - 2\text{Re}(\gamma)\} \right. \\ & \quad \left. - (1 - \epsilon_S^*) (1 - \epsilon_L) \{1 + 2\text{Re}(\gamma)\} \right] \\ & = \{2\text{Re}(\epsilon) - 2i\text{Im}(\delta) - 2\text{Re}(\gamma)\} \text{BR}(K_L \rightarrow l\nu\pi) \Gamma_L \end{aligned} \quad (16.157)$$

Dividing both sides of Eqs. (16.156a) and (16.156b) by Γ_S gives

$$2 \left(\frac{\bar{\Gamma}}{\Gamma_S} + i \tan \phi_{\text{SW}} \right) [\text{Re}(\epsilon) - i \text{Im}(\delta)] = \sum_f \eta_{2\pi} \text{BR}(K_S \rightarrow 2\pi) + \frac{\Gamma_L}{\Gamma_S} \\ \times \left[\sum_f \eta_{3\pi}^* \text{BR}(K_L \rightarrow 3\pi) + \text{BR}(K_L \rightarrow l\nu\pi) 2 \{ \text{Re}(\epsilon) - \text{Re}(\gamma) - i \text{Im}(\delta) \} \right] \quad (16.158)$$

As $\text{BR}(K_S \rightarrow 2\pi) \approx 1$, $\Gamma_L/\Gamma_S \simeq 1.75 \times 10^{-3}$, the contribution of the second line is much smaller. Since the above equation is a relation of complex numbers, we can determine two variables, $\text{Im}(\delta)$ and $\text{Re}(\gamma)$ treating $\text{Re}(\epsilon)$ as given. Alternatively, by using Eq. (16.154) we can replace $[\text{Re}(\epsilon) - \text{Re}(\gamma)]$ by an observable quanti-

ty ($A_S + A_L$) on the rhs. Then we can determine $\text{Re}(\epsilon)$ and $\text{Im}(\delta)$. Analyses by CPLEAR/KLOE [107, 238] give

$$\text{Re}(\epsilon) = 159.6 \pm 1.3 \times 10^{-5}, \quad \text{Im}(\delta) = 0.4 \pm 2.1 \times 10^{-5} \quad (16.159)$$

Note on $\Delta S = -\Delta Q$ Components

So far we have completely neglected the $\Delta S = -\Delta Q$ contribution as it does not exist in the quark model. If one is so skeptical as to doubt CPT invariance, it is logical to include the $\Delta S = -\Delta Q$ contribution [denoted as x , $\bar{x}$ or $x_{\pm} = (x \pm \bar{x})/2$], too. If its contribution is included, from a purely experimental point of view we have to add $\Delta S = -\Delta Q$ amplitudes for leptonic decays to those expressed in Eqs. (16.143):

$$\langle e^+ \nu_e \pi^- | H_W | K^0 \rangle = a + b \quad \Delta S = \Delta Q \quad (16.160a)$$

$$\langle e^- \bar{\nu}_e \pi^+ | H_W | \bar{K}^0 \rangle = a^* - b^* \quad \Delta S = \Delta Q \quad (16.160b)$$

$$\langle e^- \bar{\nu}_e \pi^+ | H_W | K^0 \rangle = c + d \quad \Delta S = -\Delta Q \quad (16.160c)$$

$$\langle e^+ \nu_e \pi^- | H_W | \bar{K}^0 \rangle = c^* - d^* \quad \Delta S = -\Delta Q \quad (16.160d)$$

We define $\Delta S = \Delta Q$ rule-breaking parameters by

$$x = \frac{c^* - d^*}{a + b} \simeq \frac{c^* - d^*}{a} (1 + \gamma), \quad \bar{x} = \frac{c^* + d^*}{a - b} \simeq \frac{c^* + d^*}{a} (1 - \gamma) \quad (16.161)$$

We can also separate them into CPT-conserving $x_+ = (x + \bar{x})/2$ and CPT-violating $x_- = (x - \bar{x})/2$ terms. These parameters enter in the asymmetries $A_{\Delta m}$, A_T , A_{CPT} and can be determined if the asymmetries are reanalyzed including these parameters. So far there is no sign of them and their upper-limit values are given by [107, 238]

$$\begin{aligned} \text{Re}(x) &= 1.9 \pm 4.1 \times 10^{-3} \\ \text{Im}(x) &= 1.2 \pm 1.9 \times 10^{-3} \\ \text{Re}(x_-) &= 0.2 \pm 1.3 \times 10^{-2} \\ \text{Im}(x_+) &= 1.2 \pm 2.2 \times 10^{-2} \end{aligned} \quad (16.162)$$

$K^- \bar{K}^0$ Mass Difference and Decay Width Difference

We learned that CPT invariance means

$$\text{CPT invariance: } \rightarrow \mathcal{A}_{11} = \mathcal{A}_{22} \quad (16.163a)$$

$$\text{or } m_{K^0} = m_{\bar{K}^0}, \quad \Gamma_{K^0} = \Gamma_{\bar{K}^0} \quad (16.163b)$$

Let us express the mass and width difference in terms of the CPT-violating parameter δ . The CPT-violating mixing parameter is given by

$$\delta = \frac{\mathcal{A}_{22} - \mathcal{A}_{11}}{2\Delta_\lambda} = \frac{\sin \phi_{\text{SW}} e^{i\phi_{\text{SW}}}}{2\Delta m} \left[i(M_{11} - M_{22}) + \frac{\Gamma_{11} - \Gamma_{22}}{2} \right] \quad (16.164)$$

$$M_{11} - M_{22} = m_{K^0} - m_{\bar{K}^0}, \quad \Gamma_{11} - \Gamma_{22} = \Gamma_{K^0} - \Gamma_{\bar{K}^0}$$

It is convenient to decompose δ into $\delta_{\parallel}$ and $\delta_{\perp}$ as in Eq. (16.124):

$$\text{Re}(\delta) = \delta_{\parallel} \cos \phi_{\text{SW}} - \delta_{\perp} \sin \phi_{\text{SW}} \quad (16.165a)$$

$$\text{Im}(\delta) = \delta_{\parallel} \sin \phi_{\text{SW}} + \delta_{\perp} \cos \phi_{\text{SW}} \quad (16.165b)$$

where $\delta_{\parallel}$, $\delta_{\perp}$ are signed real numbers. Then

$$\begin{aligned} m_{K^0} - m_{\bar{K}^0} &= M_{11} - M_{22} = \frac{2\Delta m}{\sin \phi_{\text{SW}}} \delta_{\perp} = \Delta \Gamma \{-\text{Re}(\delta) \tan \phi_{\text{SW}} + \text{Im}(\delta)\} \\ &\simeq 2\Delta m \{-\text{Re}(\delta) + \text{Im}(\delta)\} \end{aligned} \quad (16.166a)$$

$$\begin{aligned} \Gamma_{K^0} - \Gamma_{\bar{K}^0} &= \Gamma_{11} - \Gamma_{22} = \frac{4\Delta m}{\sin \phi_{\text{SW}}} \delta_{\parallel} = 2\Delta \Gamma \{\text{Re}(\delta) + \text{Im}(\delta) \tan \phi_{\text{SW}}\} \\ &\simeq 2\Delta \Gamma \{\text{Re}(\delta) + \text{Im}(\delta)\} \end{aligned} \quad (16.166b)$$

As we have already determined $\text{Re}(\delta)$ and $\text{Im}(\delta)$, the mass and decay width difference can also be evaluated. Inserting values of $\text{Re}(\delta)$ and $\text{Im}(\delta)$ in Eqs. (16.153) and (16.159), we obtain [107]

$$\begin{aligned} \delta_{\parallel} &= 1.9 \pm 2.0 \times 10^{-4} \\ \delta_{\perp} &= -1.5 \pm 2.0 \times 10^{-4} \end{aligned} \quad (16.167)$$

and subsequently from Eqs. (16.166) we have [107, 236]

$$m_{K^0} - m_{\bar{K}^0} = 0.5 \pm 5.8 \times 10^{-19} \text{ GeV} \quad (16.168a)$$

$$\Gamma_{K^0} - \Gamma_{\bar{K}^0} = -1.5 \pm 2.0 \times 10^{-18} \text{ GeV} \quad (16.168b)$$

Dividing Eq. (16.168) by the mass and width of the kaon, one finds the CPT violation can be at most [311]

$$\frac{m_{K^0} - m_{\bar{K}^0}}{m_{K^0}} \lesssim 0.8 \times 10^{-18} \quad (16.169a)$$

$$\frac{\Gamma_{K^0} - \Gamma_{\bar{K}^0}}{\Gamma_S} = 5.2 \pm 5.8 \times 10^{-4} \quad (16.169b)$$

Equation (16.169a) is considered the most stringent proof of CPT invariance at this time. Equation (16.169b) is a measure of direct CPT violation [see also Eq. (16.175a)].

Direct CPT Violation in 2π Decays*

The observed value $\arg(\omega) = -55.6^\circ$ is in reasonable agreement with $\delta_2 - \delta_0 = -47.6 \pm 1.5^\circ$. Therefore the expression for ω without CPT invariance assumption

$$\omega = \frac{1}{\sqrt{2}} \frac{\text{Re} A_2 + i \text{Im} B_2}{\text{Re} A_0 + i \text{Im} B_0} \quad (16.170)$$

indicates $\text{Im} B_I \ll \text{Re} A_I$. With this assumption, the expressions for A_L , α , ϵ , ϵ' in Eqs. (16.95) and (16.102)–(16.103) are modified as follows:

$$\frac{1}{2} A_L \rightarrow \text{Re}(\epsilon) - \left\{ \text{Re}(\delta) - \frac{\text{Re} b}{\text{Re} a} \right\} = \text{Re}(\epsilon) - \left\{ \text{Re}(\delta) + \text{Re}(\gamma) \right\} \quad (16.171a)$$

$$i \frac{\text{Im} A_0}{\text{Re} A_0} \rightarrow \frac{i \text{Im} A_0 + \text{Re} B_0}{\text{Re} A_0 + i \text{Im} B_0} \simeq i \frac{\text{Im} A_0}{\text{Re} A_0} + \left\{ \frac{\text{Re} B_0}{\text{Re} A_0} \right\} \quad (16.171b)$$

$$\epsilon_0 = \tilde{\epsilon} - \delta + i \frac{\text{Im} A_0}{\text{Re} A_0} \rightarrow \epsilon_T + i \delta \varphi - \left\{ \delta - \frac{\text{Re} B_0}{\text{Re} A_0} \right\} \quad (16.171c)$$

$$\begin{aligned} \epsilon' &\rightarrow i \omega \left[\frac{\text{Im} A_2 - i \text{Re} B_2}{\text{Re} A_2 + i \text{Im} B_2} - \frac{\text{Im} A_0 - i \text{Re} B_0}{\text{Re} A_0 + i \text{Im} B_0} \right] \\ &\simeq i \omega \left(\frac{\text{Im} A_2}{\text{Re} A_2} - \frac{\text{Im} A_0}{\text{Re} A_0} \right) + \omega \left\{ \frac{\text{Re} B_2}{\text{Re} A_2} - \frac{\text{Re} B_0}{\text{Re} A_0} \right\} \end{aligned} \quad (16.171d)$$

Those in large braces $\{\dots\}$ are newly added CPT-violating terms. Relations between redefined ϵ , ϵ' and η_{00} , η_{+-} from Eq. (16.120) do not change.

If we approximate $\Gamma_{K^0} - \Gamma_{\bar{K}^0}$ by its 2π intermediate states, a calculation similar to what we carried out to derive $\Gamma_S \simeq 2(\text{Re} A_0)^2 \rho$ (see Eq. (16.113)) leads to

$$\begin{aligned} \Gamma_{K^0} - \Gamma_{\bar{K}^0} &\simeq \rho \left[\left\{ |A(K^0 \rightarrow \pi^+ \pi^-)|^2 + |A(K^0 \rightarrow \pi^0 \pi^0)|^2 \right\} \right. \\ &\quad \left. - \left\{ |A(\bar{K}^0 \rightarrow \pi^+ \pi^-)|^2 + |A(\bar{K}^0 \rightarrow \pi^0 \pi^0)|^2 \right\} \right] \\ &= \rho \left[\left(|A_0 + B_0|^2 + |A_2 + B_2|^2 \right) \right. \\ &\quad \left. - \left(|A_0^* - B_0^*|^2 + |A_2^* - B_2^*|^2 \right) \right] \\ &= 4\rho [\text{Re} A_0 \text{Re} B_0 + \text{Re} A_2 \text{Re} B_2] \\ &\simeq 2\Gamma_S \left[\frac{\text{Re} B_0}{\text{Re} A_0} + 2|\omega|^2 \frac{\text{Re} B_2}{\text{Re} A_2} \right] \\ &\simeq 2\Gamma_S \frac{\text{Re} B_0}{\text{Re} A_0} \end{aligned} \quad (16.172)$$

Substituting Eq. (16.172) into Eq. (16.164), we obtain

$$\tilde{\delta} \equiv \delta - \frac{\text{Re} B_0}{\text{Re} A_0} \simeq i \sin \phi_{\text{sw}} e^{i\phi_{\text{sw}}} \left[\frac{m_{K^0} - m_{\bar{K}^0}}{2\Delta m} + \frac{\text{Re} B_0}{\text{Re} A_0} \right] \quad (16.173)$$

Relations between η , ϵ , ϵ' , δ and $\text{Re} B_0/\text{Re} A_0$ are depicted in Fig. 16.14.

We see that $\tilde{\delta} \parallel \delta_\perp$, namely $\tilde{\delta}$ is almost perpendicular to the superweak direction in the 2π dominance limit ($\delta\varphi = 0$). In this approximation,

$$|\tilde{\delta}| \simeq |\epsilon| \tan(\phi_\epsilon - \phi_{\text{sw}}) \quad (16.174)$$

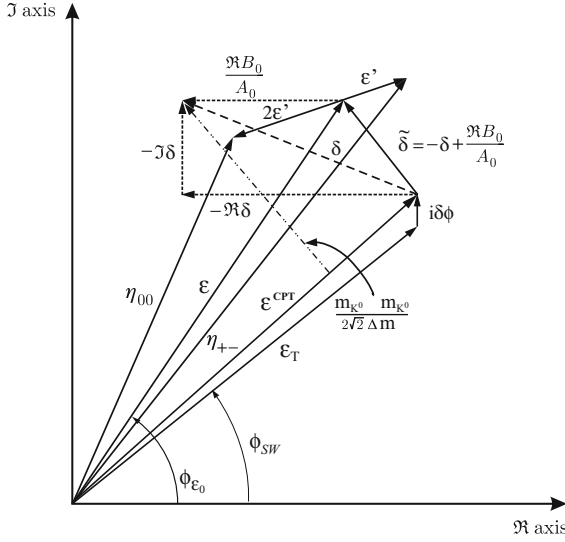


Figure 16.14 Diagram showing ϵ when CPT invariance is not assumed (not to scale). CPT-invariant ϵ is denoted as ϵ^{CPT} . In the 2π dominance approximation, $\delta\varphi = 0$, the CPT-violating term $\tilde{\delta}$ is almost orthogonal to ϵ .

Experimentally, $\phi_\epsilon - \phi_{\text{SW}} \simeq \phi_{+-} - \phi_{\text{SW}} = 0.61 \pm 0.62 \pm 1.01^\circ$.

An estimate of the CPT-violating 2π decay amplitude can be obtained as follows. From Eq. (16.172)

$$\frac{\text{Re } B_0}{\text{Re } A_0} \simeq \frac{\Gamma_{K^0} - \Gamma_{\bar{K}^0}}{2\Gamma_S} = \frac{\delta_{\parallel}}{\cos \phi_{\text{SW}}} \quad (16.175a)$$

As was discussed in Eq. (16.138), $\arg(\epsilon') \simeq \phi_{\text{SW}}$ and Eq. (16.171d) tells us that

$$\omega \left(\frac{\text{Re } B_2}{\text{Re } A_2} - \frac{\text{Re } B_0}{\text{Re } A_0} \right) \simeq -\epsilon'_{\perp} \stackrel{(16.140)}{\simeq} \frac{|\eta_{+-}|}{3} \Delta\phi \quad (16.175b)$$

Substituting experimental values of $\Delta\phi$ and $\delta_{\parallel}$ allows $\text{Re } B_0/\text{Re } A_0$ and $\text{Re } B_2/\text{Re } A_2$ to be estimated from Eqs. (16.175) [107]

$$\frac{\text{Re } B_0}{\text{Re } A_0} = 2.6 \pm 2.9 \times 10^{-4}, \quad \frac{\text{Re } B_2}{\text{Re } A_2} = 1.3 \pm 4.5 \times 10^{-4} \quad (16.176)$$

We conclude that under the present circumstances, there is no sign of CPT violation.

16.4.4

Possible Violation of Quantum Mechanics

Theoretically, CPT invariance is deeply rooted in quantum field theory. Its violation means that one of Lorentz invariance, locality, unitarity or the principle of quantum

mechanics is not correct. Unitarity is related to the conservation of probability and Lorentz invariance (or its extension, Poincaré invariance) to conservation of energy or angular momentum, therefore the most suspicious was violation of locality in the strong interaction. However, the success of the Standard Model makes the possibility very small. With the advent of quantum gravity as exemplified by black hole radiation, it was pointed out that the existence of a quantum gravitational field might possibly affect the propagation of quantum mechanical waves [201]. In other words, when we discuss CPT violation, we have to dare to intrude into the sacred area of quantum mechanics. Ellis and others in [130, 210] proposed experiments to test the breakdown of quantum coherence due to quantum gravity. In their treatment, CPT violation is incorporated in the breakdown of quantum mechanics, where a pure state develops into a mixed state. They expressed the effect in terms of three parameters α, β, γ . These parameters have dimensions of mass and are expected to have values $\sim m_K^2/m_{\text{Planck}} \sim 2 \times 10^{-20}$ GeV. Phenomenologically, their effect most conspicuously appears as the modification (destruction) of macroscopic interference effects, for example in neutron propagation and $K \rightarrow 2\pi, 3\pi, \pi l \nu$. In the neutral kaon decays, the effect could appear as the loss of coherence of the neutral kaon wave function. Figure 16.15 shows how this quantum decoherence effect would appear in the asymmetry of $K \rightarrow 2\pi$ and $K \rightarrow l\pi\nu$ [107], which we have already shown in Fig. 16.9 and Fig. 16.1. So far no breakdown of quantum mechanics has been observed. The experimental data can set upper limits to these parameters, which are approaching the estimated values of $\sim 10^{-20}$ GeV.

Another test bench for the possible breakdown of quantum mechanics is provided by DAPHNE and the B-factory, where correlated pairs of $K^0-\bar{K}^0$ (or K_S-K_L) and $B^0-\bar{B}^0$ are copiously produced [56, 236]. Here, one can also test quantum entanglement or the Einstein–Podolsky–Rosen (EPR) paradox [128] in connection with quantum decoherence. We refer to [60].

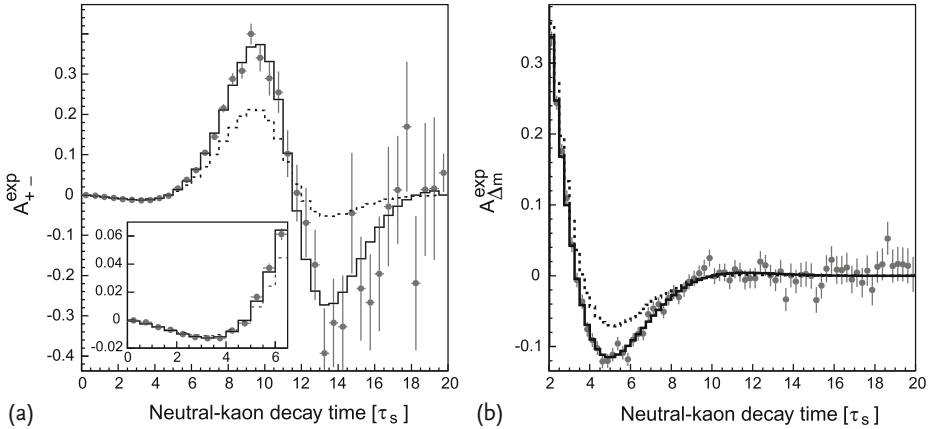


Figure 16.15 The measured decay-rate asymmetries (a) $A_{2\pi}$ and (b) $A_{\Delta m}$ analyzed for a possible loss of coherence. The dashed lines represent the expected asymmetries with positive values of α, β and γ , which are 10 times larger than the obtained upper limits [107].

16.5

Experiments on CP Parameters

Since all CP- and/or CPT-violation tests depend so much on precision experiments, we need to study how they are carried out and understand their limitations. We should also be alert in devising improved or new techniques or proposing new methods. Most of the CP-violation measurements made before 2000 were carried out by fixed-target experiments, in which a high-energy proton beam bombarded a target to produce neutral kaons (charged particles were removed by sweeping magnets). Because of the collimated kaon beam with high momentum, most experiments had to optimize their apparatus for specific purposes. With the advent of collider accelerators, a comprehensive study of CP phenomena became possible. For some specific purposes, however, where high flux of the kaon beam is required, the fixed-target experiments still have an advantage, although by no means an easy one. It took two large groups (NA31 $\rightarrow$ NA48 and E731 $\rightarrow$ KTeV) 20 years to determine just one number, the direct CP violation parameter ϵ' . We will describe how the experiments were done.

Two collider detector experiments on kaons are worth mentioning: CPLEAR at CERN and KLOE at Frascati in Italy. Since they deal with low-energy kaons, experiments with a large solid angle (almost 4π) and covering 0–20 K_S decay regions were possible.

The former used a low-energy antiproton beam to tag neutral kaons $K^0 (\bar{K}^0)$ by $\bar{p}p \rightarrow K^\pm \pi^\mp$, and the latter produced correlated kaon pairs by $e^- + e^+$ at $\sqrt{s} = m_\phi$ in a reaction

$$e^- + e^+ \rightarrow \phi \rightarrow K^0 \bar{K}^0 \text{ or } K^+ K^- \quad \text{with } J^{PC} = 1^{--} \quad (16.177)$$

The detector concept is very similar in both experiments. Here we describe the CPLEAR detector.

16.5.1

CPLEAR

CPLEAR was an experiment to study CP violation in kaon decays. Tagged K^0 and $\bar{K}^0$ were produced by stopping low-energy $\bar{p}$ in a pressurized (at 16 bar) hydrogen gas target. The beam had a flux of $10^6 \bar{p} \text{ s}^{-1}$ and its momentum was 200 MeV/c with a spread of 5×10^{-4} . The accumulated total number of $\bar{p}$ was 1.1×10^{13} . The use of liquid hydrogen was avoided to minimize the amount of matter in the decay volume. The incoming $\bar{p}$ beam was delivered by the Low Energy Antiproton Ring (LEAR) facility at CERN and the size of the beam was about 3 mm FWHM. K^0 and $\bar{K}^0$ were identified by the reactions

$$\begin{aligned} \bar{p} \text{ at rest} + p &\rightarrow (K^- \pi^+) + K^0 \\ &\rightarrow (K^+ \pi^-) + \bar{K}^0 \end{aligned} \quad (16.178)$$

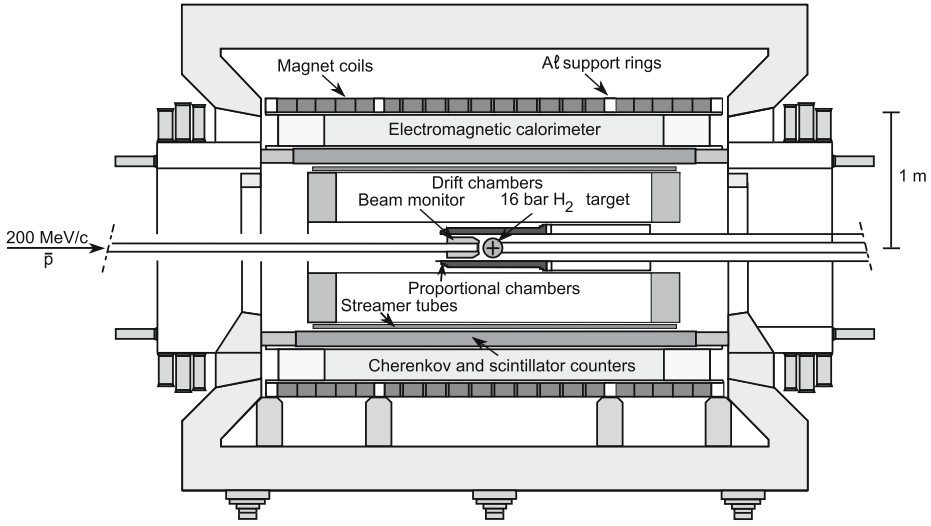


Figure 16.16 CPLEAR detector. Longitudinal view [107, 109].

each having a branching ratio of $\approx 2 \times 10^{-3}$. The conservation of strangeness in the strong interaction dictates that a K^0 is accompanied by a K^- and a $\bar{K}^0$ by a K^+ . The highest momentum of the kaon was $\sim 750 \text{ MeV}/c$, the mean decay length of K_S was 4 cm and the cylinder volume with radius of $\approx 60 \text{ cm}$ corresponded to roughly 20 K_S mean flight paths, covering the interesting interference region. The whole detector was embedded in a warm i.e. normal conducting solenoidal magnet, which provided a 0.44 T uniform field [109].

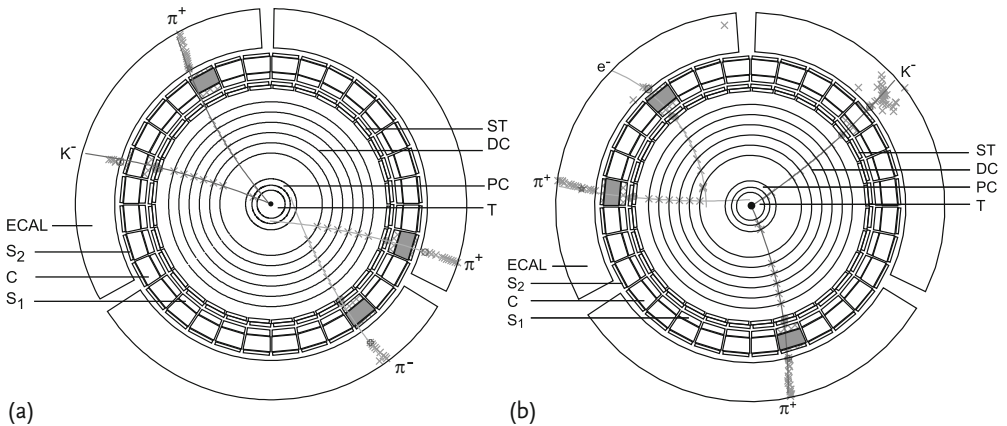


Figure 16.17 Display of CPLEAR events. $\bar{p}p \rightarrow K^- \pi^+ K^0$ with the neutral kaon decaying to (a) $\pi^+ \pi^-$ and (b) $e^- \pi^+ \bar{\nu}$. Parts of detectors: T; gas target, S₂, C, S₁; Cherenkov

and scintillator counters that make a particle identifier. Others can be identified by comparing with Fig. 16.16 [107, 109].

The general layout is shown in Figure 16.16. A series of cylindrical tracking detectors (drift chambers, streamer tubes, etc.) provided information about the trajectories of charged particles to determine their charge signs, momenta and positions. The reconstructed track resolution was $\lesssim 350 \mu\text{m}$ in r and $r\phi$ and 2 mm in z ; $0.05\% \lesssim \Delta p/p \lesssim 0.1\%$. Particle identification (PID) was carried out by a combination of a threshold-type Cherenkov detectors, to identify charged kaons, and scintillation counters, which measured the time of flight and dE/dx . The outermost detector was a lead/gas-tube sampling calorimeter (ECAL) to detect photons produced in π^0 decays.

Two transverse views of the detector together with typical event displays are shown in Fig. 16.17. Both of them show $K^-\pi^+$ tracks originating in the hydrogen target. In Fig. 16.17a, K^0 has decayed to two charged pions while in (b) it has decayed according to $K^0 \rightarrow e^-\pi^+\nu$. The total number of events measured was 6.4×10^5 .

16.5.2

NA48/KTeV

It may be surprising to learn that at the turn of the century, 30 years after their discovery, all the known CP-violation phenomena were still explained by a single number ϵ and were consistent with the superweak model, which denied the existence of direct CP violation in the decay or in any place other than in neutral kaon decays. Although the Standard Model predicted its existence and demonstrated its validity in the B-meson factory experiments later, proof of direct CP violation, also known as ϵ' , was one of the biggest issues at the turn of the century. The issue was settled by two definitive experiments, NA48 [282] at CERN and KTeV [244] at Fermilab. They are good examples of how modern state-of-the-art experiments using fixed targets are carried out and deserve some detailed explanation.

Assuming CPT invariance, the effect can be tested experimentally by measuring the double ratio of K_L and K_S decaying to both $\pi^+\pi^-$ and $\pi^0\pi^0$:

$$R = \frac{\Gamma(K_L \rightarrow \pi^+\pi^-)/\Gamma(K_S \rightarrow \pi^+\pi^-)}{\Gamma(K_L \rightarrow \pi^0\pi^0)/\Gamma(K_S \rightarrow \pi^0\pi^0)} = \left| \frac{\eta_{+-}}{\eta_{00}} \right|^2 \simeq 1 + 6\text{Re}\left(\frac{\epsilon'}{\epsilon}\right)$$

$$\text{Re}(\epsilon'/\epsilon) \simeq 1.6 \times 10^{-3} \quad (16.179)$$

The above number (10^{-3}) may not seem impressive, but in terms of decay asymmetry, which is represented by the value of $\text{Re}(\epsilon')$ itself (see Problem 16.7) rather than its ratio to ϵ , one realizes the difficulty of the experiment:

$$\delta_f = \frac{\Gamma(\bar{K}^0 \rightarrow f) - \Gamma(K^0 \rightarrow f)}{\Gamma(\bar{K}^0 \rightarrow f) + \Gamma(K^0 \rightarrow f)}$$

$$= \begin{cases} -2\text{Re}(\epsilon') = -5.2 \times 10^{-6} & f = \pi^+\pi^- \\ +4\text{Re}(\epsilon') = 10.4 \times 10^{-6} & f = \pi^0\pi^0 \end{cases} \quad (16.180)$$

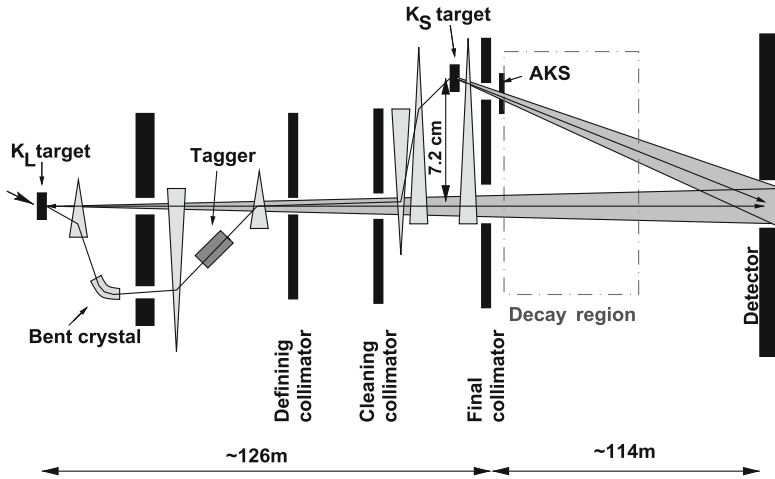


Figure 16.18 Scheme of the beam arrangement for the NA48 experiments. The K_L beam comes straight from the K_L target. The attenuated primary beam is bent and hits the K_S target 120 m downstream of the K_L target.

AKS is the K_S anti-counter located at the exit of the K_S collimator. The two beams are aimed at the center of the detector with crossing angle 0.6 mrad. Note the difference between the vertical and horizontal scales [282].

In such experiments, one needs not only to accumulate high statistics using precision detectors, which can only be constructed using modern technology, but also to minimize the systematic error, which constrains the ultimate precision. Measurement of the double ratio has the inherent advantage of canceling many systematics. By measuring the four decay modes simultaneously, any time variation or background correction common to all four modes cancels out.

Both experiments are similar in detector concept, differing mainly in the preparation of the beam. NA48 makes separate K_L and K_S lines that converge at the detector. KTeV makes two parallel K_L beam lines and uses one to regenerate a K_S beam. Figure 16.18 shows the layout of the NA48 beams. A proton beam ($\sim 1.5 \times 10^{12}$ protons per pulse, 450 GeV/c), delivered by the SPS at CERN, impinges on a target that is located 126 m upstream of the decay volume. The charged components are removed by sweeping magnets. The neutral beam consists almost entirely of photons, neutrons, and K_L^0 mesons, with some few residual K_S^0 , Λ and Ξ particles at the highest energies. The remaining primary beam goes through a bent crystal and about 10^{-5} of it, satisfying the channeling condition,¹⁶⁾ goes through the tagger and hits the K_S target just 6 m upstream of the decay volume. The tagger is used to separate K_S events from those of K_L origin in the time domain. The energy of the neutral kaons is 70–170 GeV with average 110 GeV. The whole target and the collimator system are aligned along an axis pointing to the

16) If a particle goes through a single crystal at an angle close to a major crystal direction, coherent scattering on the lattice nuclei forces the particle to follow the lattice direction. The role of the bent crystal is similar to that of a bending magnet, with reduced intensity and excellent selection of well-defined emittance. Its effect was used to separate desired particles from the halo of the beam [122].

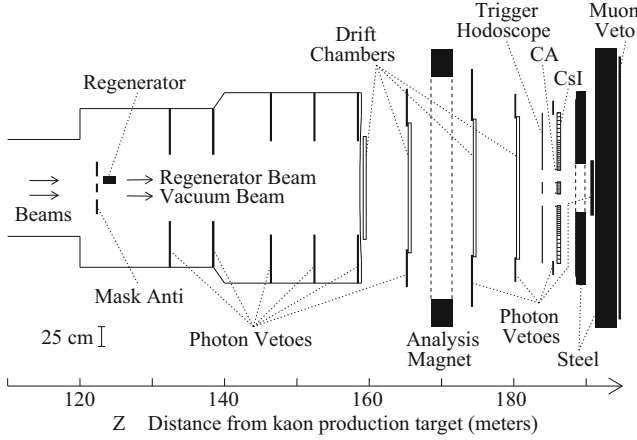


Figure 16.19 Top view of the KTeV (E832) detector. The evacuated decay volume extends to $z = 159$ m [244].

center of the detector 120 m away, such that the two beams intersect at this point with an angle of 0.6 mrad.

In KTeV, a primary proton beam from the Fermilab Tevatron [3×10^{12} protons per pulse, 800 GeV/c in a 20 s extraction cycle (“spill”)], on the other hand, makes two identical K_L beams, one of which hits the regenerator target 120 m downstream of the primary target. The average kaon momentum is 70 GeV/c, somewhat lower than that of NA48 because of the different primary energy and beam distance from production to decay point. The regenerator consists of 84 scintillator modules of $10 \times 10 \times 2$ cm³, each viewed by a photomultiplier. The last module defines the beginning of the decay volume. The total length is 170 cm, corresponding to about two hadronic interaction lengths, and the magnitude of the regeneration amplitude is $|\rho| \sim 0.03$.

For the detector we mainly explain the KTeV apparatus, with comparisons where appropriate, because the concepts of KTeV and NA48 are very similar. Figure 16.19 shows the top view of the KTeV(E832) detector. The evacuated decay volume extends to $z = 159$ m.

Two typical event displays, one showing $K \rightarrow \pi^+ \pi^-$ and the other $\rightarrow 2\pi^0$, are shown in Fig. 16.20.

Reconstruction of $K \rightarrow \pi^+ \pi^-$ A magnetic spectrometer composed of a magnet, which gives a kick of 412 MeV/c, drift chambers and triggering hodoscopes reconstructs trajectories of charged particles with the momentum resolution $\sigma_p/p = [1.7 \oplus (p/14)] \times 10^{-3}$, where p is the momentum in GeV/c. Kaon identification was performed by reconstructing the invariant mass (16.44) of the two charged particles (see Fig. 16.20, left), using the same method as the first CP-violation experiment by Fitch and Cronin, but with far better accuracy. The position resolution is ~ 100 μ m and the mass resolution is $\sim \delta m_K \approx 1.6$ MeV/c². Within the ac-

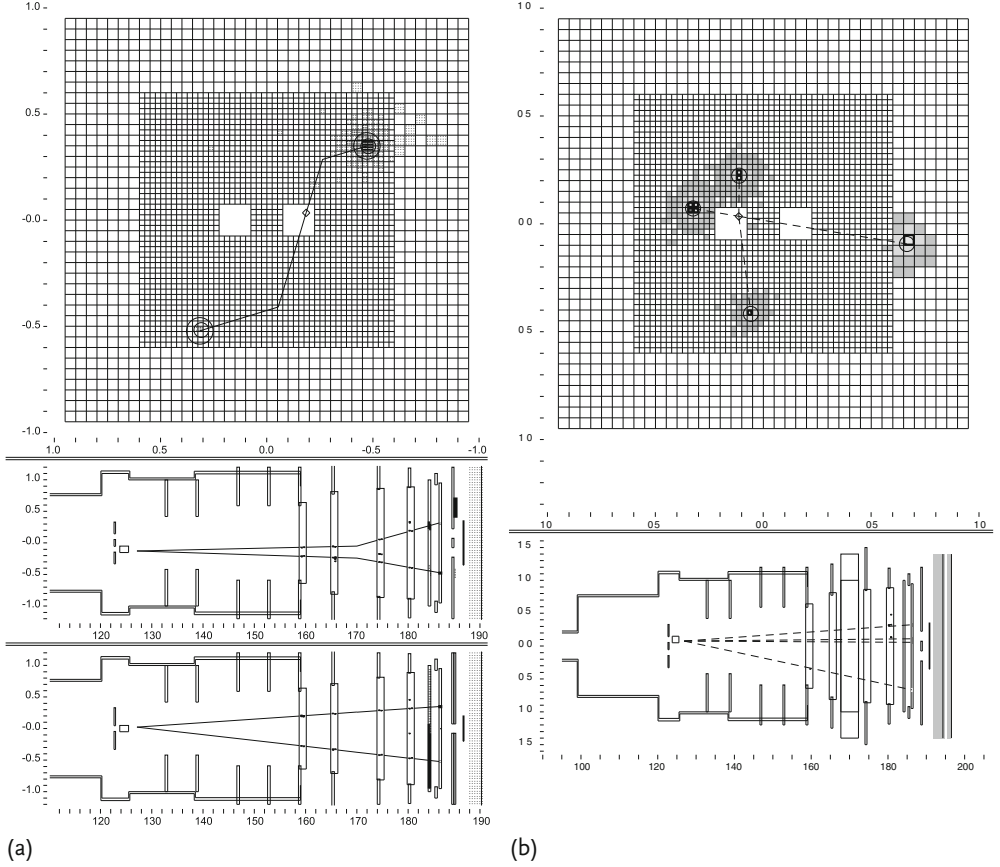


Figure 16.20 Event display: Upper figures show CsI configuration and deposited energy clusters. Lines are trajectories (shown in lower figures) of reconstructed pions projected on the CsI plane with vertex at crossing. (a) $K \rightarrow \pi^+\pi^-$; (b) $K \rightarrow \pi^0\pi^0 \rightarrow 4\gamma$ [391].

ceptance window of $488 < m_K < 508 \text{ MeV}/c^2$, backgrounds for $K_L \rightarrow \pi^+\pi^-$ (0.1%) come mainly from $K \rightarrow \pi e(\mu)\nu$, where $e(\mu)$ were misidentified as pions. The p_T^2 distribution can separate $K_S \rightarrow 2\pi$ from the coherent regenerated beam (cut on $p_T^2 < 250 \text{ MeV}/c^2$) from that of the background (diffraction scattering), as was shown in Fig. 16.5. It also tells us that major background for $K_S \rightarrow \pi^+\pi^-$ (0.087%) comes from regenerator scattering.

Reconstruction of $K \rightarrow \pi^0\pi^0$ A CsI electromagnetic calorimeter is composed of 2232 inner cells ($2.5 \times 2.5 \text{ cm}^2$) and 868 larger outer cells ($5 \times 5 \text{ cm}^2$), covering a total area of $1.9 \times 1.9 \text{ m}^2$ with two holes in the center through which the beams pass. The configuration is shown in Fig. 16.20. The depth of each crystal is 50 cm ($27 X_0$). The energy resolution is $\sigma_E/E = 2\%/\sqrt{E} \oplus 0.4\%$, where E is in GeV. $K \rightarrow \pi^0\pi^0 \rightarrow 4\gamma$ deposits energy at four places (see Fig. 16.20, right). The energy E_1, E_2 of the two

photons satisfies the relation

$$\begin{aligned} m_{\pi^0}^2 &= (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 = 2E_1 E_2 (1 - \cos \theta) \approx 4E_1 E_2 \theta^2 \\ &\approx 4E_1 E_2 r_{12}^2 / z^2 \end{aligned} \quad (16.181)$$

where z is the distance to the decay position, r_{12} is the distance between the two energy clusters in the detector and θ is their opening angle from the decay point. Assuming any 2γ 's come from π^0 decay, the vertex position is reconstructed. Out of three ($=4C_2/2$) pair combinations made of 4γ 's, two with the minimum separation are chosen and m_K is reconstructed (Fig. 16.21b). NA48 uses a liquid krypton ionization chamber and its energy resolution is $\sigma_E/E = 3.2 \pm 0.2\%/\sqrt{E} \oplus 9 \pm 1\%/E \oplus 0.42 \pm 0.05\%$. The vertex position was determined with a method similar to KTeV using the formula $d = \sum_{i>j} \sqrt{E_i E_j} r_{ij} / m_K$ with assumed m_K . The resolution of the longitudinal vertex position was on the order of 20–30 cm for KTeV and ~ 50 cm for NA48. Backgrounds for $K_L \rightarrow 2\pi^0$ (0.484%) come mainly from $3\pi^0$, where 2γ 's are missing, and for $K_S \rightarrow 2\pi^0$ (1.235%) mainly from the regenerator and collimator scattering. The backgrounds for neutral decay modes are 10 times larger than those for charged decay modes and are one of the major sources of errors. This is due to the energy resolution of the electromagnetic calorimeter and inapplicability of the p_T cut to eliminate the regenerator scatterings.

Other Detectors At the back end are located steel absorbers and muon identifiers. The decay volume and some parts of the detector are equipped with an extensive array of photon veto counters to reduce trigger rates and backgrounds and to define sharp apertures and edges that limit the detector acceptance.

Overall Correction Separation of regenerated $K_S \rightarrow 2\pi$ events from those of CP-violating $K_L \rightarrow 2\pi$ is physically clear as the two beams are ~ 20 cm apart. The solid angle difference due to separation and the fact that they are not aligned along the center of the detector could be a potential source of systematic errors. The regenerator also produces some background. KTeV observes four decay modes simultaneously, by frequently switching the regenerator position alternately between the two beam positions and also the polarity of the magnet. Any time fluctuation cancels when the ratio is taken and the geometrical difference averaged out. Note, however, some backgrounds contribute differently to the four modes, hence do not cancel and have to be corrected carefully for each mode. Use of high resolution detectors is considered essential for removing such background.

The total numbers of events accumulated are

$$\begin{array}{rcl} N(K_L \rightarrow \pi^0 \pi^0) & \approx & 3.4 \times 10^6 \\ N(K_S \rightarrow \pi^0 \pi^0) & \approx & 5.6 \times 10^6 \\ N(K_L \rightarrow \pi^+ \pi^-) & \approx & 11.2 \times 10^6 \\ N(K_S \rightarrow \pi^+ \pi^-) & \approx & 19.4 \times 10^6 \\ \hline \text{Total} & \approx & 40 \times 10^6 \end{array} \quad (16.182)$$

Before $\text{Re}(\epsilon'/\epsilon)$ can be extracted from these data, backgrounds must be identified and subtracted, and acceptance corrections must be determined. Both these tasks

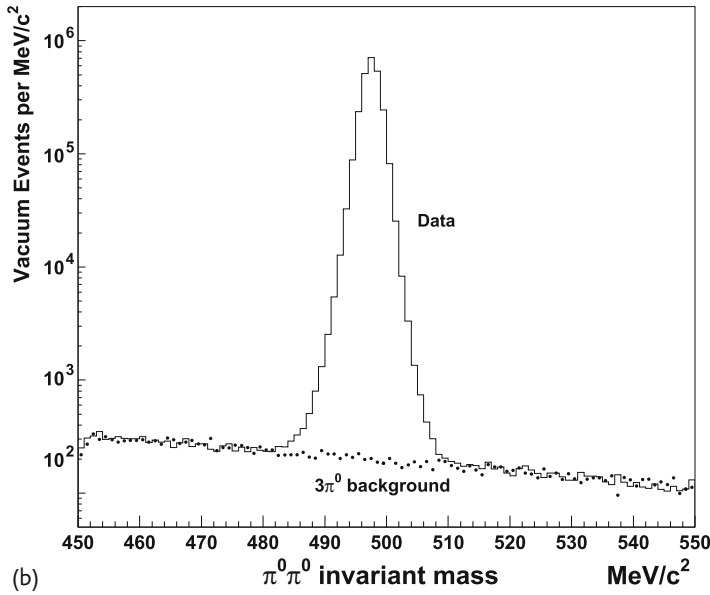
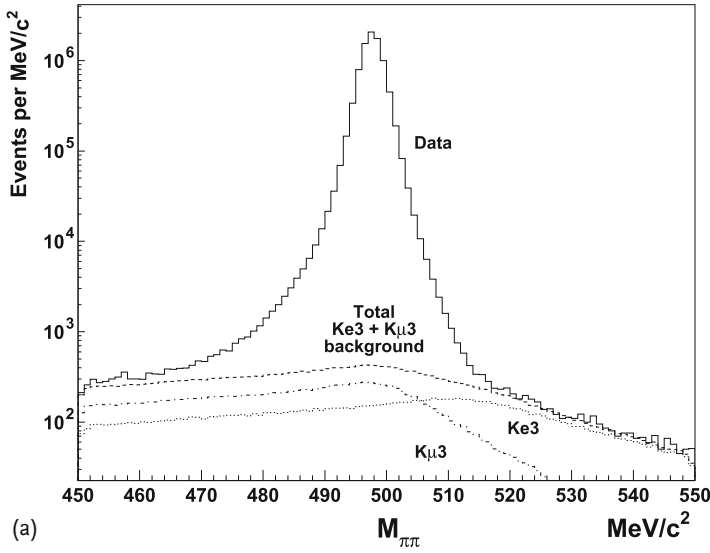


Figure 16.21 Reconstructed mass distribution of 2π from vacuum (K_L) beam. (a) $\pi^+\pi^-$ with estimated backgrounds by Monte Carlo. (b) $\pi^0\pi^0$. Backgrounds are mainly from $3\pi^0$ with two photons missing [244, 347].

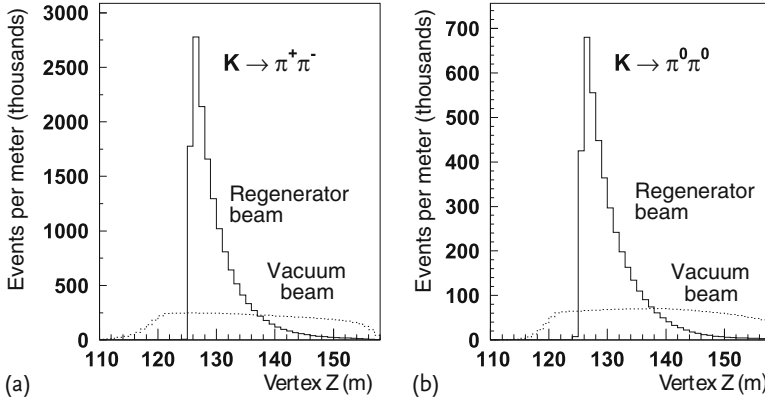


Figure 16.22 z-Vertex distribution of reconstructed (a) $K \rightarrow \pi^+\pi^-$ and (b) $K \rightarrow \pi^0\pi^0$ for the vacuum beam (dashed) and regenerator beam (line histogram) [244].

require Monte Carlo simulations of the experiment. Especially important is the decay vertex distribution, which differs widely between K_L and K_S , as shown in Fig. 16.22. While the K_L decay distribution is almost flat in the decay volume, that of K_S halves in about 2 m. This could be a potential source of serious systematic errors. NA48 uses a weighted z-distribution for K_L vertices [weight factor $\sim e^{-\Delta\Gamma_S t} = e^{(z/\beta\gamma)(1/\tau_S - 1/\tau_L)}$] to make them similar, which contributes to cancellation of many systematics but at the expense of statistics. KTeV, on the other hand, relies on elaborate Monte Carlo simulations to assure the correctness of reconstruction. Figure 16.23 shows examples of decay distributions for $K \rightarrow \pi^+\pi^-$ and $K_L \rightarrow \pi e(\mu)\nu$. The agreement is excellent, but the degree of disagreement indicates the limits of systematic errors, too. In the end the largest single error comes from the overall reproducibility of the events, exemplified by the K_L decay distribution in Fig. 16.23.

Both NA31/NA48 and E731/KTeV spent nearly twenty years in deriving their final results. During this time, they both learned from experience and from each other. The single biggest issue common to both experiments was to determine geometrical efficiency and, in the end, their differences were minor compared to their similarity.

Extraction of $\text{Re}(\epsilon'/\epsilon)$ K_S components include regenerated $K_S \propto \rho$ and CP-violating $K_S \propto \eta$, which are not distinguishable. The number of events is proportional to the square of the sum $\eta + \rho$ [see Eq. (16.86)] and is a function of ρ/η . The fitting must be carried out including not only $\text{Re}(\epsilon'/\epsilon)$ but also Δm , τ_S , the phases ϕ_{00} , ϕ_{+-} and ρ . The accuracy can be cross-checked by floating well-known parameters such as τ_S and Δm to see if they agree with the known values. The end

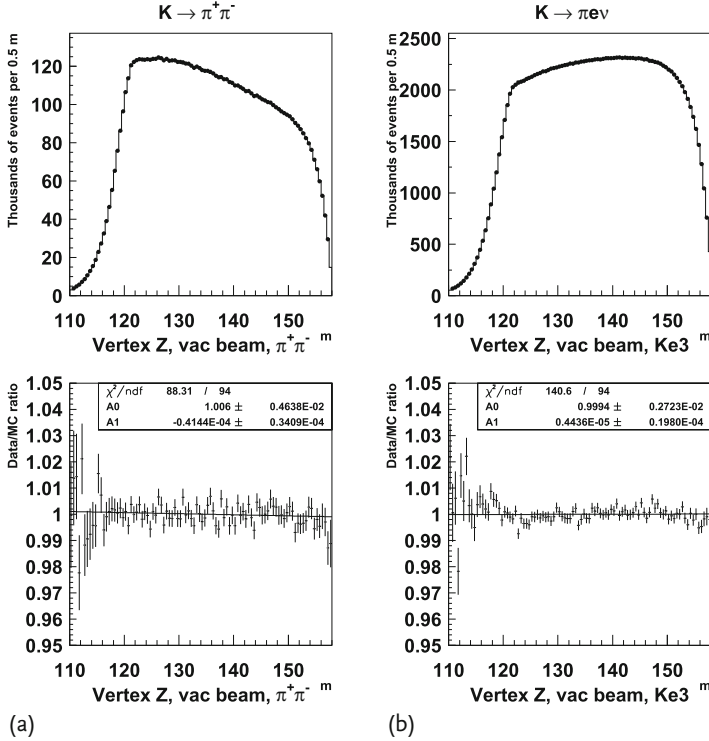


Figure 16.23 Vertex distributions of K_L from $K \rightarrow \pi^+ \pi^-$ (a) and $K \rightarrow \pi l \nu$ (b) for both reconstructed data and Monte Carlo (MC) simulations. Lower figures show the ratio data/MC, indicating good agreement [244].

results of KTeV and NA48 are [51, 244, 245, 311]

$$\text{Re}(\epsilon'/\epsilon) = 2.07 \pm 0.26 \times 10^{-3} \quad \text{KTeV} \quad (16.183a)$$

$$\text{Re}(\epsilon'/\epsilon) = 1.47 \pm 0.22 \times 10^{-3} \quad \text{NA48} \quad (16.183b)$$

$$\text{weighted average} = 1.65 \pm 0.23 \times 10^{-3} \quad \text{PDG08} \quad (16.183c)$$

which clearly established the nonzero value of ϵ' .

16.6

Models of CP Violation

In discussing models of CP violation, a primary assumption is CPT invariance. Experimentally, the relative strength of the CPT-violating strong interaction is estimated from Eq. (16.169a) to be $\lesssim 10^{-18}$. Dividing this number by α , the upper limit of CPT violation in the electromagnetic interaction is found to be $< 10^{-16}$,

and by dividing by $G_F m_K^2$, that of the weak interaction $< 10^{-14}$. These values give the most stringent test of CPT violation. For the electromagnetic interaction, there exists a test by the magnetic moment of the electron

$$\begin{aligned} [g(e^+) - g(e^-)]/g_{\text{average}} &= -0.5 \pm 2.1 \times 10^{-12} \\ [g(\mu^+) - g(\mu^-)]/g_{\text{average}} &= -0.11 \pm 0.12 \times 10^{-8} \end{aligned} \quad (16.184)$$

but all other tests give at most $\sim 10^{-4}$. CPT symmetry is known to hold with excellent accuracy.

Theoretically, CPT invariance has rock-solid foundations, as described in Sect. 16.4.4. It is very hard to construct a self-consistent theory without CPT invariance. Therefore, we assume its invariance in the discussions below. In the following paragraphs we list models that accommodate CP violation (denoted as $\mathcal{CP}$). The models are classified according to where one seeks the origin of CP violation. The reason for listing models other than the Standard Model is because the Kobayashi–Maskawa model is considered OK within the Standard Model, but not sufficient to explain the matter–antimatter asymmetry of the universe, and theories beyond the Standard Model, including GUT (grand unified theory) and superstring models, are hot issues in the world of high-energy physics.

CP Violation in the Strong Interaction The model assumes that $\mathcal{CP}$ occurs in two steps. First, K_L decay to X in a CP-conserving interaction, followed by $X \rightarrow 2\pi$ in the strong interaction. If we let f be the strength of the transition amplitude $X \rightarrow 2\pi$, f has to be $\sim 10^{-3}$ to explain $|\eta| \sim 10^{-3}$ and is referred to as the millistrong force. Since CPT invariance is assumed, the interaction is one of four possible processes:

$$(1) \ P, \mathcal{C}, \mathcal{T} \quad (2) \ \bar{P}, \mathcal{C}, T \quad (3) \ \bar{P}, C, \mathcal{T} \quad (4) \ \bar{P}, \mathcal{C}, \mathcal{T}$$

In the electromagnetic transitions of atomic levels, a small mixture of different parity components induces interference with the main component of the amplitude and rotates the direction of polarization. This tiny polarization shift is experimentally observed, but can be explained quite well by the neutral current effect of the Standard Model, hence models (2),(3) and (4), which include P violation, are excluded [50, 70]. The possibility of millistrong interaction remains in principle, but no one has conceived of a natural way to embed CP violation in QCD and the possibility does not seem to be taken seriously.

CP Violation in the Electromagnetic Interaction If the electromagnetic interaction is the cause of CP violation, $\alpha \sim |\eta|$ means CP has to be violated maximally. However, experimentally observed asymmetry values in the decay of η mesons $\eta \rightarrow \pi^+\pi^-\pi^0$, $\pi^+\pi^-\gamma$ give $0.9 \pm 1.7 \times 10^{-3}$, $9 \pm 4 \times 10^{-3}$ [311]. Therefore C violation in the electromagnetic interaction is small if any. The P-violating component does not exist beyond that allowed by the Standard Model, as was explained in the previous section. Therefore, the possibility of CP violation originating in the electromagnetic interaction is excluded.

Superweak Model The superweak model assumes an interaction H_{SW} with $\Delta S = 2$ is the cause of CP violation. The CP-violating term $\langle K^0 | H_{\text{SW}} | \bar{K}^0 \rangle$ appears in the mass matrix A_{12} . Hermiticity of H_{SW} dictates the CP-violating component must satisfy $A_{21} = A_{12}^*$. Then from Eq. (16.67)

$$\epsilon = \frac{A_{21} - A_{12}}{2\Delta_\lambda} = -\frac{\text{Im}(\langle K^0 | H_{\text{SW}} | \bar{K}^0 \rangle)}{(\Gamma S - \Gamma_L)/2 - i\Delta m} \quad (16.185)$$

Therefore, the relation $\phi_\epsilon = \phi_{\text{SW}} = \tan^{-1}(2\Delta m/\Delta \Gamma)$ holds exactly (strictly speaking, $+\pi$ is possible theoretically but excluded experimentally). Calling the strength of the superweak model $\sim f G_F$, we have $f \sim \epsilon \Delta m / G_F m_K^3 \sim 10^{-11}$. The effect appears only in the mass matrix and does not appear in the $\Delta S = 1$ decay amplitude, hence $\epsilon' = 0$. Therefore, the following equalities are exact in the superweak model:

$$|\eta_{+-}| = |\eta_{00}|, \quad \phi_{+-} = \phi_{00} = \phi_\epsilon = \phi_{\text{SW}} \quad (16.186)$$

In this model, all the CP-violating effects can be described by a single parameter ϵ and do not appear in other mesons, including B mesons. Although the superweak model was finally excluded by NA48/KTeV and by B-factory experiments, historically it played a prominent role for nearly 40 years.

Milliweak Interaction The milliweak model assumes the CP-violating effect occurs in the $\Delta S = 1$ interaction, whose strength is $\sim 10^{-3}$. There are varieties of models. One model increases the number of Higgs particles to introduce complex numbers in the Lagrangian. Another is a left-right symmetric model, in which parity violation is assumed spontaneous. Since the Standard Model has been established, these have become exotic models that go beyond the Standard Model but are discussed in GUT in the context of the matter-dominant universe.

Origin of CP Violation in the Standard Model The CP violation effect in the Standard Model belongs to the class of milliweak interactions. CP violation can be introduced into the Lagrangian by making the coupling constant complex. In the Standard Model, the weak interaction is mediated by $W^\pm$ and Z^0 . They couple to weak isospin doublets and the charged current interaction Lagrangian has the form of Eq. (15.200), namely

$$\begin{aligned} \mathcal{L}_{W1} &= -\frac{g_w}{\sqrt{2}} \bar{\Psi}_L \gamma^\mu (\tau_+ W_\mu^+ + \tau_- W_\mu^-) \Psi_L \\ \Psi_L &= \begin{bmatrix} u \\ d' \end{bmatrix}, \quad \begin{bmatrix} c \\ s' \end{bmatrix}, \quad \begin{bmatrix} t \\ b' \end{bmatrix} \end{aligned} \quad (16.187)$$

If we denote u_j ($j = 1, \dots, 3$) = (u, c, t) and d_k ($k = 1, \dots, 3$) = (d, s, b) , the interaction Lagrangian of the charged W boson can be written as

$$\mathcal{L}_W = -\frac{g_w}{\sqrt{2}} \sum \left[\bar{u}_{jL} \gamma^\mu V_{jk} d_{kL} W_\mu^+ + \bar{d}_{kL} \gamma^\mu V_{jk}^* u_{jL} W_\mu^- \right] \quad (16.188)$$

where V_{jk} is the KM (Kobayashi–Maskawa) matrix element introduced in Eq. (15.160b). If we apply CP transformation [Appendix F, Eq. (F.4)] to the above Lagrangian, it is transformed to

$$\mathcal{L}_W = -\frac{g_W}{\sqrt{2}} \sum \left[\bar{d}_{kL} \gamma^\mu V_{jk} u_{jL} W_\mu^- + \bar{u}_{jL} \gamma^\mu V_{jk}^* d_{kL} W_\mu^+ \right] \quad (16.189)$$

We see that CP invariance requires the KM matrix V_{jk} to be real. As the field u_j, d_k has the freedom to change its phase, many KM elements V_{jk} can be made real using this freedom. Let us see how many complex numbers the KM matrix has. In general, a $N \times N$ complex matrix has $2N^2$ independent real variables. Unitarity requirement reduces the number to N^2 . In this N^2 -dimensional space of real numbers, $N C_2 = N(N-1)$ space rotations can be defined in terms of rotation angle θ_j . The rest of the variables become phase angles ϕ_k . On the other hand, the $2N$ quark fields have the freedom to change their phase as

$$U \equiv \begin{bmatrix} u \\ c \\ t \end{bmatrix} \rightarrow F_U U = \begin{bmatrix} e^{i\phi_u} & & \\ & e^{i\phi_c} & \\ & & e^{i\phi_t} \end{bmatrix} U \quad (16.190a)$$

$$D \equiv \begin{bmatrix} d \\ s \\ b \end{bmatrix} \rightarrow F_D D = \begin{bmatrix} e^{i\phi_d} & & \\ & e^{i\phi_s} & \\ & & e^{i\phi_b} \end{bmatrix} D \quad (16.190b)$$

This transforms

$$V \rightarrow F_U^* V F_D, \quad \text{namely} \quad V_{jk} \rightarrow e^{-i(\phi_{u_j} - \phi_{d_k})} V_{jk} \quad (16.191)$$

$2N - 1$ phases, excluding the common overall phase, can be removed by redefinition of the quark fields. In the end,

$$N^2 - \frac{N(N-1)}{2} - (2N-1) = \frac{(N-1)(N-2)}{2} \quad (16.192)$$

phases remain. Therefore for $N \leq 2$, all the matrix elements can be made real and no CP violation effect appears unless put in explicitly by hand. As CP is violated, N must be ≥ 3 was the original motivation of introducing the KM matrix. As the measurement of Z^0 decay width shows that the number of neutrinos with $m_\nu < m_Z/2$ is limited to three (see Sect. 15.7.4), it is natural to think $N = 3$, as is required by the Standard Model. In this case the KM matrix is expressed by three rotation angles and one phase angle, as in Eq. (15.162). In this convention, the dominant contribution of the phase δ is contained in V_{ub} and V_{td} .

Therefore, in the Standard Model, Feynman diagrams that contain V_{ub} or V_{td} are CP-violating amplitudes. Referring to the Standard Model description of $\bar{K}^0 \leftrightarrow K^0$ transition amplitude, Fig. 16.3, when the intermediate states j or $\bar{j}$ are top quarks, the transition $t \leftrightarrow d$ attaches V_{td} to the amplitude and it becomes a CP-violating amplitude. But the decay amplitudes are on-mass shell; no top quarks appear in the decay diagrams in the lowest order approximation, referred to as the tree diagram,

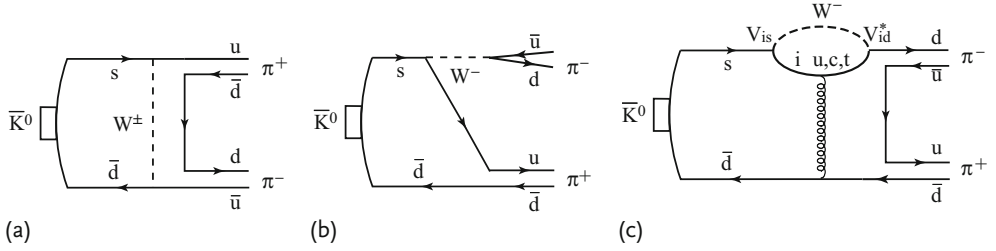


Figure 16.24 $K \rightarrow \pi^+ \pi^-$ decay amplitudes. Tree diagrams (a,b) do not contain KM matrix element V_{td} . The penguin diagram (c) contains the top quarks in the intermediate states. The coupling V_{td} is complex, hence diagram (c) violates CP.

shown in Fig. 16.24a,b. In order to insert the top quark in the intermediate states, the higher order diagram or the so-called penguin diagram has to be introduced (Fig. 16.24c).

The Standard Model clarifies the dynamical structure of CP violation and its calculation gives the right order of magnitude for the experimental values, but owing to QCD complexity accurate estimates are hard to obtain.

17

Hadron Structure

17.1

Historical Overview

We have succeeded in constructing the quark model from hadron spectroscopy. What it can tell us is about the static nature of the hadron, for example its mass, electric charge, spin, isospin and strangeness. They are all characteristics that can be observed from outside. The quark model cannot tell us about the inner structure of the hadron. It cannot tell us whether an elementary particle with fractional charge really exists inside the hadron. In order to investigate the inner structure of the hadron, we need a probe that can go inside the hadron and show us its structure. Qualifications for the probe include (1) its own properties are well known, (2) it has small size and (3) it can penetrate deep inside the hadron. From this point of view, leptons are useful. One may compare the situation with the situation at the beginning of the 20th century, when there were several models of atomic structure: the α particle was used as a probe, and the investigation of scattering patterns from a golden target led to the Rutherford model being established.

In the 1960s, the world of particle physics was in confusion. In the field of strong interactions, confidence in field theory was weak and the idea of particle democracy based on the analyticity of the scattering amplitude was at its peak. Above all the bootstrap theory told us that all particles are equal, no elementary particles exist and that they themselves are building blocks of other particles, namely they are their own fundamental constituents. Part of the reasoning came from an experimental observation of the nucleon form factor. A series of experiments by Hofstadter et al. [81, 206] showed that, unlike the atom, the nucleon does not have a core and the whole structure is soft like a cloud or jelly. There was no reason to believe in the existence of a substructure inside the nucleon. Unlike 50 years previously, this time Thomson's viewpoint seemed more pertinent than Rutherford's. In addition to denying subatomic structure, particle physicists at that time found the unusual properties of the quark, such as fractional charge, incompatible with their common sense. Even Gell-Mann, the quark proposer himself, said for a while that it should be considered as a mathematical tool rather than a real model.

The situation changed drastically when deep inelastic scattering data were presented in 1969. This was the Rutherford experiment brought up to date in the sense

that it clarified the parton structure of the nucleon, namely, that the nucleon is composed of point-like particles. Series of deep inelastic scattering data, beginning with those from the MIT-SLAC group, clarified that the nucleon is a quark composite and that the quark really has fractional charge, which paved the way for quantum chromodynamics (QCD) to appear as the theory of the strong interaction.

This chapter explains the process of clarifying the inner structure of the nucleon to be a composite of point-like particles generically called partons, and the role of deep inelastic scattering of electrons (and muons) and neutrinos in identifying the partons as quarks and gluons.

17.2

Form Factor

Momentum Transfer as a Measure of the Size We Can Probe The electron is the best known and most easily obtainable probe of all time. In order to use the electron as a probe, we need to know the formula for its reaction with other particles. The basic equation is the Rutherford scattering formula introduced in Sect. 6.3.1:

$$\begin{aligned}
 d\sigma &= \frac{2\pi}{v} |H_{fi}|^2 \delta(E - E_i) \frac{d^3\mathbf{p}}{(2\pi)^3} = \frac{1}{(2\pi)^2} \frac{m p}{v} |V_{fi}|^2 d\Omega \\
 \frac{d\sigma}{d\Omega} &= \left| \frac{m}{2\pi} \int d\mathbf{r} V(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} \right|^2 = \frac{4Z^2\alpha^2 m^2}{Q^4} \\
 m, Z &: \text{mass and electric charge of the target} \\
 V(\mathbf{r}) &= \frac{Ze^2}{4\pi} \frac{1}{r} = \frac{Z\alpha}{r} \\
 Q^2 &= |\mathbf{Q}|^2 = |\mathbf{k}_i - \mathbf{k}_f|^2 = 4k^2 \sin^2 \frac{\theta}{2} \\
 \mathbf{k}, \theta &: \text{momentum and scattering angle of the electron}
 \end{aligned} \tag{17.1}$$

$\mathbf{Q}$ is the momentum that the electron gives to the target, or the strength of the impact, and is called the momentum transfer. Equation (17.1) holds when the target is a point particle. If the target is spread and has a charge distribution $\rho(\mathbf{r})$, the potential is modified to

$$V(\mathbf{r}) \rightarrow V'(\mathbf{r}) = \int d\mathbf{r}' V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') \tag{17.2}$$

and the matrix element to

$$\begin{aligned}
 V_{fi} &\rightarrow V'_{fi} = \int d\mathbf{r} V'(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} = \left[\int d(\mathbf{r} - \mathbf{r}') V(\mathbf{r} - \mathbf{r}') e^{i\mathbf{Q}\cdot(\mathbf{r} - \mathbf{r}')} \right] \\
 &\quad \times \int d\mathbf{r}' \rho(\mathbf{r}') e^{i\mathbf{Q}\cdot\mathbf{r}'} \\
 &= V_{fi} F(Q^2)
 \end{aligned} \tag{17.3a}$$

$$F(Q^2) = \int d\mathbf{r} \rho(\mathbf{r}) e^{i\mathbf{Q} \cdot \mathbf{r}} \quad (17.3b)$$

where, for simplicity, the distribution was assumed to be spherically symmetric. $F(Q^2)$ is called the form factor and is a measure of the target spread. Normalizing the charge distribution by

$$\int d\mathbf{r} \rho(\mathbf{r}) = 1 \quad (17.4)$$

we have

$$F(0) = 1 \quad (17.5)$$

From the above discussion, we know that the cross-section formula for an extended target is obtained from the point target by multiplying by the form factor (squared)

$$d\sigma(\theta) = d\sigma_{pt} |F(Q^2)|^2 \quad (17.6)$$

The above equation says that if we know the cross section formula of a point target, we can obtain information about the target spread by measuring the scattering pattern. As is clear from Eq. (17.3b), the part of $\rho(\mathbf{r})$ where $\mathbf{Q} \cdot \mathbf{r} \gg 1$ does not contribute to the integral because of violent oscillation due to $\exp(i\mathbf{Q} \cdot \mathbf{r})$. This means that the form factor is a measure of total charge integrated over the region $r < 1/Q$. In other words, the electron has penetrated to the depth given by $r \simeq 1/Q$. The maximum value of the momentum transfer Q gives the limit to which we can probe the inner structure of the target.

Target Size When $\mathbf{Q} \cdot \mathbf{r} \ll 1$, the form factor can be expanded as

$$\begin{aligned} F(Q^2) &= \int d\mathbf{r} \rho(\mathbf{r}) [1 + (i\mathbf{Q} \cdot \mathbf{r}) - (1/2)(i\mathbf{Q} \cdot \mathbf{r})^2; \dots] \\ &= 1 - \frac{Q^2}{6} \langle r^2 \rangle + \dots \end{aligned} \quad (17.7a)$$

$$\langle r^2 \rangle = \int d\mathbf{r} \rho(\mathbf{r}) |\mathbf{r}|^2 = -6 \left. \frac{dF(Q^2)}{dQ^2} \right|_{Q^2=0} \quad (17.7b)$$

This means that if we can measure the cross section for small values of $|Q|$, the information we can obtain is limited to the total charge ($Z\alpha$) and the size of its spread. The larger the value of $|Q|$ we can obtain, i.e., the larger the energy and scattering angle, the deeper we can probe the inner structure.

Examples of Form Factors It helps to understand the physical meaning of the form factor if we know the spectral shape of typical spatial distributions. We list a few representative distributions.

Spherical Distribution $\rho(r) = \rho_0$ for $r \leq R$ and $\rho(r) = 0$ for $r > R$ (Fig. 17.1)

$$F(\xi) = \frac{3(\sin \xi - \xi \cos \xi)}{\xi^3}, \quad \xi = QR \quad (17.8a)$$

This is similar to light going through a pinhole making several bright and dark rings and is called a diffraction pattern. It is a typical wave characteristic that commonly occurs on going around behind a clear-cut boundary.

Nuclear Charge Distribution The nucleus looks like a blurred sphere with radius $R = r_0 A^{1/3}$. Figure 17.2b shows scattering data for an electron incident on calcium. It resembles the diffraction pattern, but somewhat smoothed. A charge distribution that reproduces the observed pattern well is depicted in Fig. 17.2a. R represents the nuclear size, $t = (4 \ln 3)a$ the thickness of the surface or the degree of blurriness. One sees valleys become shallower as the surface is blurred. Fitting parameters to the data, we obtain, roughly, $r_0 \simeq 1.03 \times 10^{-13}$ cm, $t \simeq 2.4 \times 10^{-13}$ cm.

Exponential Distribution When the distribution becomes harder smoothly towards the core, it may be represented by an exponential or a Gaussian:

$$\rho(r) = \rho_0 \exp(-r/R) \quad (17.9a)$$

$$F(Q^2) = \frac{1}{(1 + Q^2/R^2)^2} \quad (17.9b)$$

A good example is the sun, which is a gas sphere with intensity slowly rising towards the core. Note: The reader is reminded that the Yukawa potential $V(r) \sim e^{\mu r}/r$ (which is a singular distribution and hence cannot be normalized) has the Fourier transform $V(Q^2) \sim 1/(1 + Q^2/\mu^2)$, which resembles the exponential form factor but is softer.

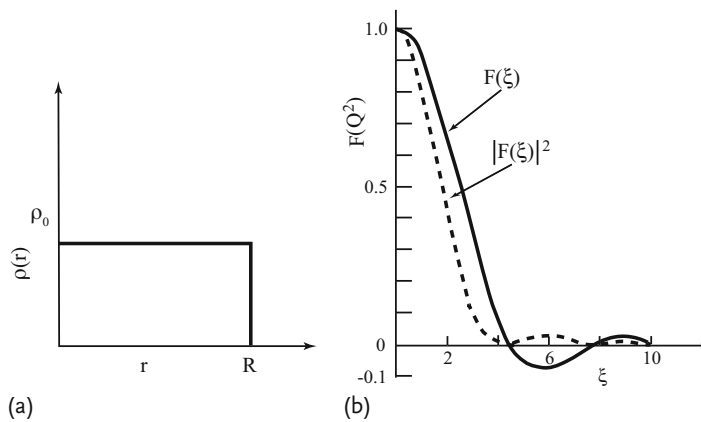


Figure 17.1 (a) $\rho(r) = \rho_0$ for $r \leq R$ and $\rho(r) = 0$ for $r > R$. (b) $F(\xi) = 3(\sin \xi - \xi \cos \xi)/\xi^3$. When there is a clear-cut boundary, the form factor shows a typical diffraction pattern.

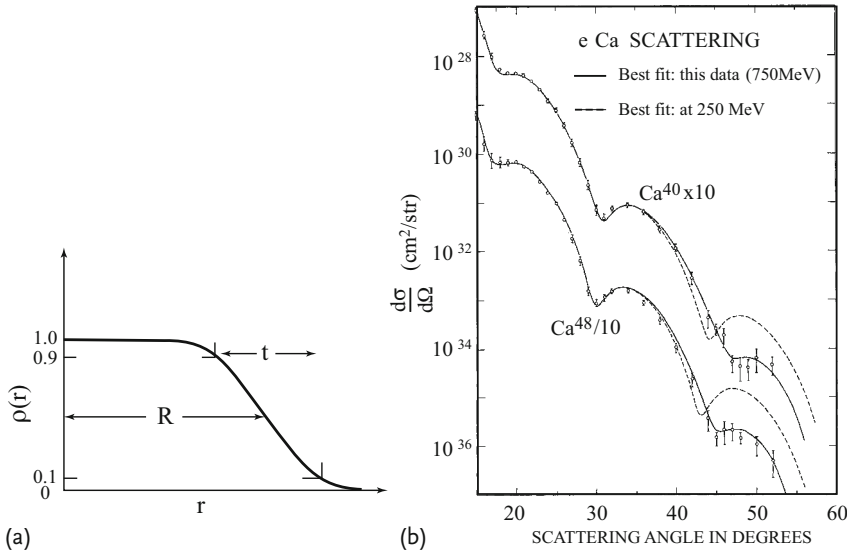


Figure 17.2 (a) $\rho(r) = \rho_0/[1 + \exp(r-R)/a]$, $t = (4 \ln 3)a$. (b) e -Ca scattering data. From fitted curves $R = r_0^{1/3}$, $r_0 = 1.03 \times 10^{-13}$ cm, $t = 2.4 \times 10^{-13}$ cm are obtained [57].

Gaussian Distribution

$$\rho(r) = \rho_0 \exp(-r^2/2\sigma^2) \quad (17.10a)$$

$$F(Q^2) = \exp(-Q^2\sigma^2/2) \quad (17.10b)$$

Problem 17.1

Derive Eqs. (17.9) and Eqs. (17.10).

17.3

e-p Elastic Scattering

Electromagnetic Form Factor of the Nucleon The proton has spin $s = 1/2$ and a spatially spread structure. If it were a point particle, namely a Dirac particle, the elastic scattering cross section with an electron

$$e(\omega, \mathbf{k}) + p(E, \mathbf{p}) \rightarrow e(\omega', \mathbf{k}') + p(E', \mathbf{p}') \quad (17.11)$$

could be obtained from the e - μ scattering cross section in the laboratory (LAB) frame [Eq. (7.20)] by replacing the target mass by the proton mass M :

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{LAB}} = \sigma_M \frac{\omega'}{\omega} \left[1 - \frac{q^2}{2M^2} \tan^2 \frac{\theta}{2} \right] \quad (17.12a)$$

$$\sigma_M = \sigma_{\text{Mott}} = \frac{4\alpha^2 \omega'^2}{Q^4} \cos^2 \frac{\theta}{2} \quad (17.12b)$$

$$q^2 = -Q^2 = (k - k')^2 \simeq -2\omega\omega'(1 - \cos \theta)$$

$$\omega' = \omega/[1 + \omega(1 - \cos \theta)/M] \quad (17.12c)$$

$Q^2 = -q^2$ is a relativistic version of Q^2 . The formula is valid for the electron mass set to $m = 0$. The first term in $[\dots]$ represents scattering by the electric Coulomb field created by the target and the second by the magnetic field. When the target spread is taken into account, the charge distribution and the magnetic moment distribution are generally different and two different form factors need to be introduced. In order to see how to include them, we express the most general proton current in the form

$$\begin{aligned} q J_p^\mu(x) &= e \langle p' | J^\mu(x) | p \rangle \equiv e J_p^\mu(p', p) e^{-iq \cdot x} \\ &= e \bar{u}(p') \left[F_1(q^2) \gamma^\mu + \frac{\chi}{2M} F_2(q^2) i \sigma^{\mu\nu} q_\nu \right] u(p) e^{-iq \cdot x} \end{aligned} \quad (17.13)$$

From Lorentz invariance alone, other terms proportional to $q^\mu = (p' - p)^\mu$, $(p' + p)^\mu$ are allowed. However, using current conservation $\partial_\mu J_p^\mu = 0$ and the Dirac equation that the plane wave functions $u(p)$, $\bar{u}(p')$ satisfy, it can be proved that the above two terms are enough.

Problem 17.2

Show that a term proportional to $q^\mu = (p' - p)^\mu$, if it exists in Eq. (17.13), is forbidden by current conservation. Next, prove the equality which is valid when sandwiched between $\bar{u}(p')$ and $u(p)$:

$$\gamma^\mu = \frac{1}{2M} [(p' + p)^\mu + i \sigma^{\mu\nu} q_\nu] \quad (17.14)$$

which is referred to as the Gordon equality.

The Gordon equality shows that the term proportional to $(p' + p)^\mu$ is already included in Eq. (17.13). To understand the physical meaning of Eq. (17.13), rewrite it as

$$e J_p^\mu(p', p) = e \bar{u}(p') \left[F_1(q^2) \frac{(p' + p)^\mu}{2M} + \frac{F_1(q^2) + \chi F_2(q^2)}{2M} i \sigma^{\mu\nu} q_\nu \right] u(p) \quad (17.15)$$

and take nonrelativistic limit $[p'^\mu, p^\mu \rightarrow (M, 0)]$. Then expressing $A_\mu = (\phi, -\mathbf{A})$

$$(p' + p)^\mu A_\mu \rightarrow 2M\phi, \quad i \sigma^{\mu\nu} q_\nu A_\mu \rightarrow i \sigma^{ij} q_j A_i \sim \boldsymbol{\sigma} \cdot \nabla \times \mathbf{A} = \boldsymbol{\sigma} \cdot \mathbf{B} \quad (17.16a)$$

If the form factors, are normalized as $F_1(0) = 1$, $F_2(0) = 1$, we have

$$e J_p^\mu(p', p) A_\mu \rightarrow e\phi + (1 + \chi) \frac{e}{2M} \boldsymbol{\sigma} \cdot \mathbf{B} \quad (17.16b)$$

From this equation, we see the first term $\sim \gamma^\mu$ contains magnetic moment (Dirac moment) interaction as well as the Coulomb interaction, but the contribution of the anomalous magnetic moment is contained in the second term $\sim i\sigma^{\mu\nu}q_\nu$. As the proton has charge e and magnetic moment $\mu_p = (1 + \chi_p)e/2M$

$$F_{1p}(0) = 1, \quad F_{2p}(0) = 1, \quad \chi = \chi_p \quad (17.17)$$

Similarly, as the neutron has total charge 0, although the inside distribution may not vanish,

$$F_{1n}(0) = 0, \quad F_{2n}(0) = 1, \quad \chi = \chi_n \quad (17.18)$$

The experimental values are given by

$$\chi_p = 1.79, \quad \chi_n = -1.91 \quad (17.19)$$

Rosenbluth Formula Using the proton current Eq. (17.13) and the electron current

$$q j_e^\mu(x) = -e \langle k' | j_e^\mu(x) | k \rangle = -e \bar{u}(k') \gamma^\mu u(k) e^{-iq \cdot x} \quad (17.20)$$

$$q = k - k'$$

The electron-proton elastic scattering amplitude is given in the one-photon-exchange approximation as

$$iT_{fi} = -e^2 \int d^4x j_e^\mu(x) \left(\frac{-ig_{\mu\nu}}{q^2} \right) J_p^\nu(x) \quad (17.21)$$

and the cross section becomes

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{LAB}} = \sigma_M \frac{\omega'}{\omega} \left[\left(F_1^2 - \frac{\chi^2 q^2}{4M^2} F_2^2 \right) - \frac{q^2}{2M^2} (F_1 + \chi F_2)^2 \tan^2 \left(\frac{\theta}{2} \right) \right] \quad (17.22)$$

Defining

$$G_E(q^2) = F_1(q^2) + \frac{\chi^2 q^2}{4M^2} F_2(q^2), \quad G_M(q^2) = F_1(q^2) + \chi F_2(q^2) \quad (17.23)$$

the cross section Eq. (17.22) is rewritten as

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{LAB}} = \sigma_M \frac{\omega'}{\omega} \left[\frac{G_E^2 + \tau G_M^2}{1 + \tau} + 2\tau G_M^2 \tan^2 \left(\frac{\theta}{2} \right) \right] \quad (17.24a)$$

$$\tau = \frac{Q^2}{4M^2} = -\frac{q^2}{4M^2} \quad (17.24b)$$

$$Q^2 = -(k - k')^2 = 4\omega\omega' \sin^2 \frac{\theta}{2}, \quad \omega' = \frac{\omega}{1 + (\omega/M)(1 - \cos \theta)} \quad (17.24c)$$

This is referred to as the Rosenbluth formula. In the limit $q^2 \rightarrow 0$, $G_E \rightarrow 1$, $G_M \rightarrow 1 + \chi$, $G_E(q^2)$, $G_M(q^2)$ are considered as expressing the electric charge and magnetic moment distributions, respectively.

Early data on elastic e - p scattering came from electron beams at Stanford university and SLAC (Stanford Linear Accelerator Center). As the cross section has the form

$$\frac{1}{\sigma_M} \frac{d\sigma}{d\Omega} = A(Q^2) + B(Q^2) \tan^2 \left(\frac{\theta}{2} \right) \quad (17.25)$$

by changing angle θ at fixed Q^2 , A and B and consequently G_E , G_M can be obtained as a function of Q^2 . Figure 17.3 shows data on the proton magnetic moment.

The observed data can be fitted well with experimental formulae in the range $0 \leq Q^2 \leq 30 (\text{GeV}/c)^2$

$$G_E^p = \frac{G_M^p}{\mu_p} = \frac{G_M^n}{\mu_n} = \frac{1}{[1 + (Q^2[(\text{GeV}/c)^2]/0.71)]^2} \quad (17.26)$$

$$G_E^n(Q^2) = 0$$

Comparing the expression with Eq. (17.9), we see the charge distribution inside the neutron is neutral, as it is outside, and all other distributions are very smooth from outside to the core and well represented by an exponential shape. This means the size of the nucleon is about

$$\sqrt{\langle r^2 \rangle} \simeq 0.7 \times 10^{-13} \text{ cm} \simeq 1/(2m_\pi) \quad (17.27)$$

If the nucleon has a core and has charge λe , the form factor should approach λ asymptotically. However, the data do not show such a tendency and can be interpreted as that the charge distribution is very soft right to the core.

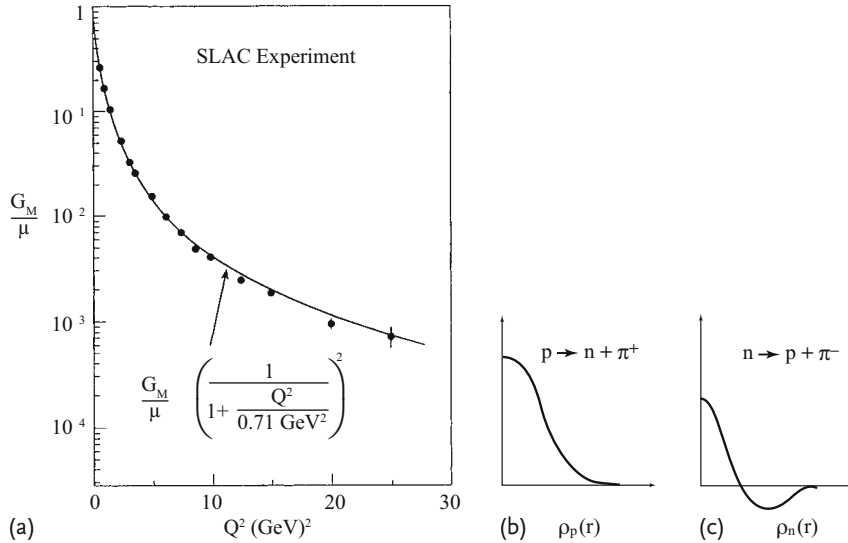


Figure 17.3 (a) Form factor of the proton magnetic moment [364]. (b,c) Expected charge distributions of the proton and the neutron in early models.

Before the appearance of the quark model, the nucleon was considered to emit and reabsorb pions

$$p \rightarrow n + \pi^+, \quad n \rightarrow p + \pi^- \quad (17.28)$$

In this image, the charge distribution is expected to be of the form depicted in Fig. 17.3b,c. The proton charge distribution will be positive at the center and the neutron distribution is expected to be positive in the core and negative in the periphery. The data do not support this concept. In the quark model, $p = uud$, $n = udd$. Since both charge and magnetic moment distribution are considered to represent that of the constituent quarks, the similarity of G_E^p , G_M^p and G_M^n are easily explained. The net charge of the neutron vanishes everywhere if the three udd distributions are the same. As the three quarks are separate and do not make any particular hard core if distributed smoothly, the exponential form and absence of a hard core are also explained.

17.4

Electron Proton Deep Inelastic Scattering

17.4.1

Cross-Section Formula for Inelastic Scattering

In quasielastic scattering, where the final state is composed of a scattered electron and a recoiled particle with mass M_X

$$e(k) + P(p) \rightarrow e(k') + X(P_F) \quad \text{generally } M_X \neq M_p \quad (17.29)$$

there is a one-to-one correspondence between the scattered electron angle θ and the energy loss $\nu \equiv \omega - \omega'$ if the incident energy ω and M_X are fixed. Measurement of one variable determines the other, too (Problem 17.3).

Problem 17.3

Fixing the incident electron energy on a target at rest and setting $W = M_X$, derive the relation

$$\nu \equiv \omega - \omega' = \frac{Q^2}{2M} + \frac{W^2 - M^2}{2M} \quad (17.30)$$

between the momentum transfer Q^2 and scattered energy ω' . Since Q^2 is a function of ω , ω' and $\cos \theta$, this equation means the scattering angle θ is uniquely determined once ω' is known. Specially for elastic scattering ($W = M$), show that

$$\nu = \frac{Q^2}{2M} = \frac{\omega \omega'}{M} (1 - \cos \theta) \quad (17.31)$$

Assume the electron mass $m = 0$.

For inelastic scattering, $W (= M_X > M)$ is arbitrary and takes continuous values within the kinematically allowed region. In this region, the scattering angle θ and the scattered energy ω' are independent. If we observe only scattered electrons, the cross section has the generic form

$$\frac{d^2\sigma}{d\Omega d\omega'} = \sigma_M \left[W_2(Q^2, \nu) + 2 W_1(Q^2, \nu) \tan^2 \left(\frac{\theta}{2} \right) \right] \quad (17.32)$$

Here, W_1 , W_2 are called structure functions, analogous to form factors in elastic scattering. Unlike the elastic case, they are functions of both Q^2 and ν . Equation (17.32) can be derived from Eq. (17.21). The transition amplitude for inelastic scattering can be written as

$$iT_{fi} = i(2\pi)^4 \delta(k + p - k' - P_F) j_{e\mu} \frac{e^2}{q^2} J_p^\mu \quad (17.33a)$$

$$j_{e\mu} = \bar{u}(k') \gamma_\mu u(k), \quad J_p^\mu = \langle X(P_F) | J_{EM}^\mu(0) | p(p) \rangle \quad (17.33b)$$

Following the standard procedure, the cross section can be derived as

$$d\sigma = \frac{(4\pi\alpha)^2}{4(k \cdot p)} L_{\mu\nu} \frac{1}{Q^4} 4\pi M W^{\mu\nu} \frac{d^3k'}{(2\pi)^3 2\omega'} \quad (17.34a)$$

$$L_{\mu\nu} = \frac{1}{2} \sum_{\text{spin}} \bar{u}(k') \gamma_\mu u(k) \bar{u}(k) \gamma_\nu u(k') = 2 \left[k'_\mu k_\nu + k'_\nu k_\mu + (q^2/2) g_{\mu\nu} \right] \quad (17.34b)$$

$$W^{\mu\nu} \equiv \frac{(2\pi)^4}{4\pi M} \sum_F \delta^4(p + q - P_F) \left[\frac{1}{2} \sum_{\text{spin}} J_p^\mu J_p^{\nu *} \right] \quad (17.34c)$$

As J_p^μ is hermitian

$$W^{\mu\nu} = W^{\nu\mu *} \quad (17.35)$$

$W^{\mu\nu}$ is a Lorentz tensor. Moreover, it is a function of q and p only, because the proton spin average has been taken. Therefore it can be expanded by the available Lorentz tensors $q^\mu q^\nu$, $p^\mu p^\nu$, $p^\mu q^\nu$, $p^\nu q^\mu$, $i\epsilon^{\mu\nu\rho\sigma} p_\rho q_\sigma$. The current conservation $\partial_\mu J^\mu \sim q_\mu J^\mu = 0$ constrains

$$q_\mu W^{\mu\nu} = q_\nu W^{\mu\nu} = 0 \quad (17.36)$$

Three Lorentz tensors satisfy this condition:

$$\left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right), \quad \left(p^\mu - \frac{(p \cdot q) q^\mu}{q^2} \right) \left(p^\nu - \frac{(p \cdot q) q^\nu}{q^2} \right), \quad (17.37)$$

$$i\epsilon^{\mu\nu\rho\sigma} p_\rho q_\sigma$$

The third term is an antisymmetric tensor and has opposite parity relative to the first two. Since the electromagnetic interaction conserves parity, this term does not

enter. In fact, by multiplying by $L_{\mu\nu}$, it is easily proved that their product vanishes. Therefore, $W^{\mu\nu}$ can be expressed as

$$W^{\mu\nu} = \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right) W_1(Q^2, \nu) + \left(p^\mu - \frac{(p \cdot q) q^\mu}{q^2} \right) \left(p^\nu - \frac{(p \cdot q) q^\nu}{q^2} \right) \frac{W_2(Q^2, \nu)}{M^2} \quad (17.38)$$

which defines W_1 and W_2 . Using

$$q^\mu L_{\mu\nu} = q^\nu L_{\mu\nu} = 0 \quad (17.39)$$

we can show

$$L_{\mu\nu} W^{\mu\nu} = 4\omega\omega' \left[W_2 \cos^2 \frac{\theta}{2} + 2 W_1 \sin^2 \frac{\theta}{2} \right] \quad (17.40)$$

From this we finally obtain the desired formula

$$\begin{aligned} \frac{d^2\sigma}{d\Omega d\omega'} &= \frac{4\alpha^2}{Q^4} \omega'^2 \left[W_2(Q^2, \nu) \cos^2 \left(\frac{\theta}{2} \right) + 2 W_1(Q^2, \nu) \sin^2 \left(\frac{\theta}{2} \right) \right] \\ &= \sigma_M \left[W_2 + 2 W_1 \tan^2 \left(\frac{\theta}{2} \right) \right] \end{aligned} \quad (17.41)$$

Problem 17.4

Prove Eq. (17.41).

Problem 17.5

Assuming the proton to be a structureless Dirac particle, and setting the final state of the proton as p' (i.e. elastic scattering), calculate the structure function by inserting $J_p^\mu = e\bar{u}(p')\gamma^\mu u(p)$ in Eq. (17.34c), or by comparing Eq. (17.12a) and Eq. (17.41), and show that

$$W_2(Q^2, \nu) = \delta \left(\nu - \frac{Q^2}{2M} \right) \quad (17.42a)$$

$$W_1(Q^2, \nu) = \frac{Q^2}{4M^2} \delta \left(\nu - \frac{Q^2}{2M} \right) \quad (17.42b)$$

Hint: Use the following formula:

$$\int d\omega' \delta \left(\nu - \frac{Q^2}{2M} \right) = \frac{\omega'}{\omega} \quad (17.43)$$

17.4.2

Bjorken Scaling

In 1969, the SLAC-MIT group carried out an experiment in which an electron beam was incident on a proton target and the intensity of the scattered electron was measured. Figure 17.4 [149] shows some of the data: a differential cross section as a function of final hadronic mass W for a fixed incident energy ω of 10 GeV and the scattering angle $\theta = 6^\circ$. Here, one sees a sharp peak (scaled down by $1/8.5$) at the left side corresponding to elastic scattering, and a few other quasielastic peaks, corresponding to resonance production

$$e + p \rightarrow e + N^* \quad M^* > M \quad (17.44)$$

Below the peaks one sees a broad hill-like bump, which constitutes inelastic scattering, where multiple particles are produced with invariant mass W continuously changing in the range allowed kinematically. One also refers to this region as the continuum.

Data similar to in Fig. 17.4 sliced at some narrow range of W band and plotted as a function of Q^2 are shown in Fig. 17.5 [76]. One sees that the cross section of elastic scattering falls off rapidly as a function of Q^2 , which is a manifestation of the target being an extended object, but those of inelastic scattering at large W (loosely dubbed deep inelastic scattering) show a relatively slow decrease.

Actually, such asymptotic behavior was expected by Bjorken on a theoretical basis [62]. A more quantitative description of the asymptotic behavior at large hadron

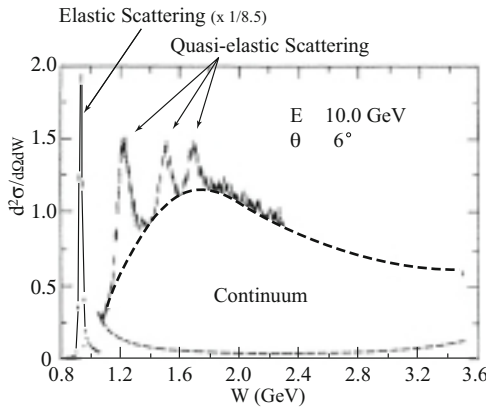


Figure 17.4 $d^2\sigma/d\Omega d\omega'$ ($E_2 = 10$ GeV, $\theta = 6^\circ$), is plotted as a function of W , the invariant mass of the hadronic final states [69, 149, 227]. The elastic peak ($W = M$) is scaled down by $1/8.5$. Its radia-

tive tail is also drawn. The broad bump under the dashed line labeled continuum indicates inelastic scattering. Dashed line is drawn only to guide eyes.

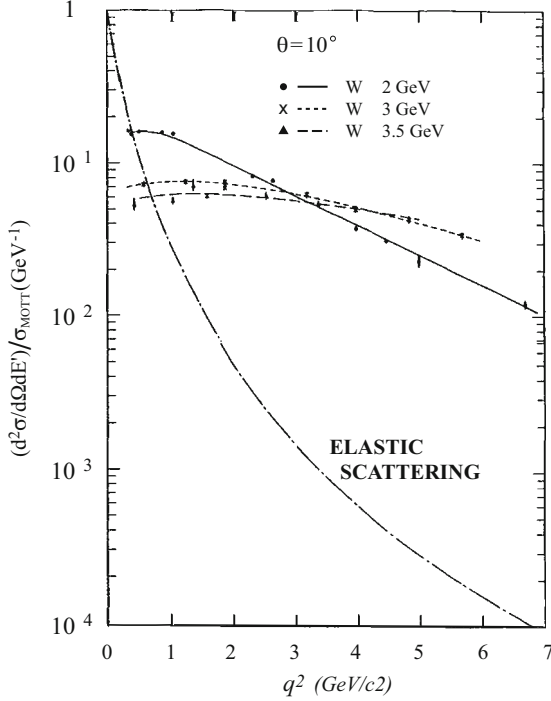


Figure 17.5 $(d^2\sigma/d\Omega d\omega')/\sigma_M$ in GeV^{-1} [76].

mass W is expressed

$$\begin{aligned} &\text{in the limit } \omega, Q^2, \nu \rightarrow \infty \\ &\text{and for } 0 < x \equiv \frac{Q^2}{2M\nu} < 1 \end{aligned} \quad (17.45)$$

as

$$\begin{aligned} \nu W_2(Q^2, \nu) &\rightarrow F_2(x), \\ M W_1(Q^2, \nu) &\rightarrow F_1(x) \end{aligned} \quad (17.46)$$

Equation (17.45) is often referred to as the scaling limit. Equations (17.46) mean that in the limit, the structure functions are not functions of two independent variables but depend only on their ratio, which is referred to as Bjorken scaling. In Figure 17.6, $M W_1$ and νW_2 are plotted as functions of $\delta = (W^2 + Q^2)/Q^2 = 1/x + M^2/Q^2$, which is $\sim 1/x$ in the scaling limit. The variable δ is used to take account of the mass effect in the low-energy region.

They do not depend on Q^2 , as is shown in Fig. 17.7, a clear manifestation of Bjorken scaling. Bjorken's reasoning for the scaling is as follows. Since $M W_1$, νW_2 are dimensionless quantities, physically fixed variables such as the proton mass

1) These F_1, F_2 are not to be confused with the form factors that appeared in Eq. (17.13).

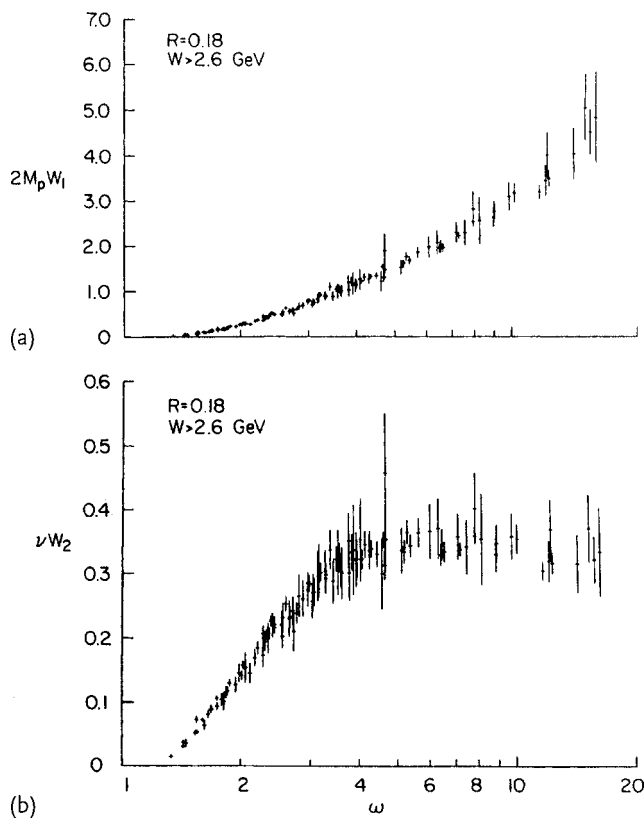


Figure 17.6 Bjorken scaling. (a) Structure functions plotted as a function of δ (denoted as ω) $= (W^2 + Q^2)/Q^2 = 1/x + M^2/Q^2$ [148, 149]. They do not depend on Q^2 , as is shown in Fig. 17.7. The plot consists of data with $W > 1.8 \text{ GeV}$.

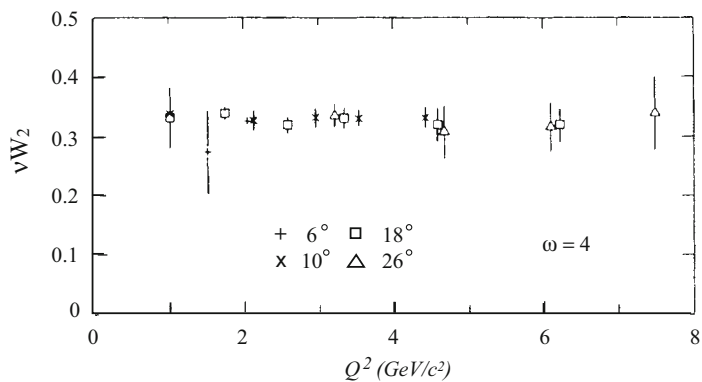


Figure 17.7 Bjorken scaling. νW_2 does not depend on Q^2 . The plot consists of data with $W > 1.8 \text{ GeV}$.

can be neglected in the scaling limit and they should depend only on an available dimensionless variable, which is x .

We will explain the physical meaning of the Bjorken scaling in terms of partons in the next section. In summary, the soft structure inferred from observation of elastic scattering had to be reconsidered in light of the deep inelastic scattering data.

17.5

Parton Model

The parton model was proposed by Feynman to analyze high-energy hadron collisions [142, 143] and applied by Bjorken and Paschos to e - p inelastic scattering [65] (see also [394]).

17.5.1

Impulse Approximation

Let us assume that the nucleon is a collection of point-like particles (partons) and that the scattering cross section can be approximated as the sum of the scattering by partons (impulse approximation). Denoting the partons by q_i , the impulse approximation means

$$\left. \frac{d\sigma}{d\Omega} \right|_{eN} = \sum_i \left. \frac{d\sigma}{d\Omega} \right|_{eq_i} \quad (17.47)$$

Summation of the individual scattering events requires that each scattering is independent of and does not affect the others. For this assumption to be valid, the time τ_{int} necessary to finish the interactions must be sufficiently short compared to the time τ_{free} , which is a measure of each parton flying freely

$$\tau_{\text{int}} \ll \tau_{\text{free}} \quad (17.48)$$

As the parton is inside the nucleon, it is bound to other partons. The motion of a parton induces a drag effect on other partons due to forces working between them. The drag time would be of the order of the inverse of the binding energy

$$\tau_{\text{free}} \sim \frac{1}{\text{binding energy}} \simeq \frac{1}{E_{\text{free}} - E_{\text{tot}}} > \frac{1}{M} \quad (17.49)$$

The magnitude of the binding energy is expected to be at most the nucleon mass (M). It is not small, but fixed, and does not increase with the nucleon energy $E = \sqrt{p^2 + M^2}$. Viewed from the incident electron, the high-energy nucleon looks like a flat pancake (Fig. 17.8a) because of Lorentz contraction. The contraction factor is given by $1/\gamma \simeq M/E$.

The interaction time is considered to be the time that the electron spends in the nucleon or that it takes to pass through the nucleon. As the nucleon has the size

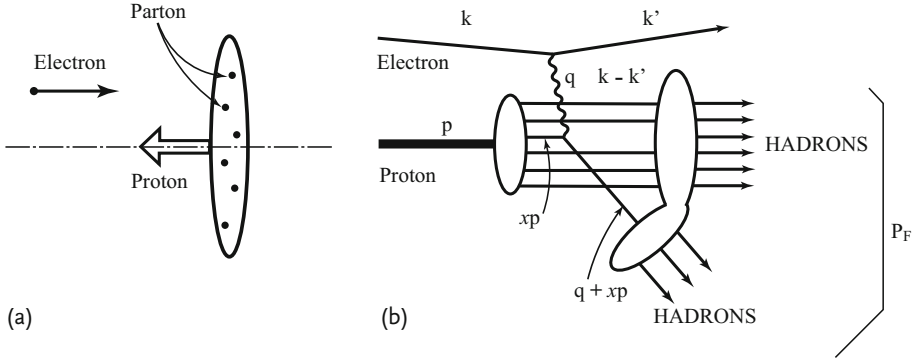


Figure 17.8 (a) Because of Lorentz contraction, the target proton looks like a pancake viewed from the incoming electron. The interaction time, which is the time for the electron to pass through the nucleon, is much shorter than the partons' mutual interaction time (time to move transverse to the electron's direction of motion). Partons are free in this

approximation. (b) Parton kinematics. Only a parton with fractional momentum x of the parent proton interacts with the electron. Other partons are spectators and pass through without interacting. Partons are considered free before and after scattering. When they fly away beyond the hadron size, they transform themselves into many hadrons (hadronization).

$\sim 1/m_\pi$ in its rest frame, the interaction time is given by

$$\tau_{\text{int}} \simeq \frac{1}{m_\pi} \frac{M}{E} \quad (17.50)$$

Consequently, for high energies where

$$E \gg M^2/m_\pi \sim 6.3 \text{ GeV} \rightarrow \tau_{\text{int}} \ll \tau_{\text{free}} \quad (17.51)$$

the condition Eq. (17.48) for the impulse approximation will be satisfied.

Next we rewrite the condition in terms of the parton variables. Because of Lorentz contraction, the longitudinal momentum $p_{\parallel}$ of the nucleon is much larger than the transverse momentum $p_{\perp}$ of the parton inside the nucleon. Therefore, for the i -th parton's longitudinal momentum we have $p_{i\parallel} \gg p_{i\perp} \simeq \langle p_{\perp} \rangle$, where $\langle p_{\perp} \rangle$ is the average transverse momentum of the parton inside the nucleon. Its value ($\sim 330 \text{ MeV}/c$ [335]) can be estimated from transverse momentum distributions of the secondary particles in p - p collisions (see Fig. 13.30). When the electron collides with the proton, the latter is excited to a virtual state with mass W . The inverse of the energy transfer to the nucleon is a measure of the interaction time and is given by

$$\frac{1}{\tau_{\text{int}}} \sim q^0 \sim \sqrt{p^2 + W^2} - \sqrt{p^2 + M^2} \sim \frac{W^2 - M^2}{2p} = \frac{2M\nu - Q^2}{2p} \quad (17.52)$$

where we have used Eq. (17.30) in deriving the last equality. On the other hand, if we set the longitudinal parton momentum as fraction x of the parent nucleon

($p_{i\parallel} = x_i p_{\parallel}$), the time for which the nucleon can be separated to free partons would be given by

$$\frac{1}{\tau_{\text{free}}} \sim \Delta E = \sqrt{p^2 + M^2} - \left\{ \sqrt{(x_1 p_{1\parallel})^2 + p_{1\perp}^2 + \mu_1^2} + \sum_{i \geq 2} \sqrt{(x_i p_{i\parallel})^2 + p_{i\perp}^2 + \mu_i^2} \right\} \quad (17.53)$$

where μ_i is the mass of the i -th parton and we have attached subscript 1 to the parton that interacts with the electron. If the nucleon has infinite energy [or approximately higher energy than that satisfying Eq. (17.51)], we may consider that all the partons move in the same direction as the parent nucleon (i.e. $x_i > 0$). Since the sum of momenta of all the partons has to be equal to that of the nucleon, we use

$$0 < x_i < 1, \quad \sum x_i = 1$$

$$\sqrt{(x_i p_{i\parallel})^2 + p_{i\perp}^2 + \mu_i^2} \sim x_i p_{\parallel} + \frac{\mu_i^2 + p_{i\perp}^2}{2x_i p} \quad (17.54)$$

and calculate ΔE , which gives

$$\Delta E \sim \frac{1}{2p} \left[M^2 - \sum_i \left(\frac{\mu_i^2 + p_{i\perp}^2}{2x_i} \right) \right] \quad (17.55)$$

The above equation means that the impulse approximation is not valid for $x_i \rightarrow 0$, $x_i \rightarrow 1$. A measure for the approximation to be valid is given by

$$1 \gg x_i \gg \mu/p, \quad p_{\perp}/p \quad (17.56)$$

The parton has energy-momentum given by

$$p_i^{\mu} = \left(\sqrt{(x_i p_{\parallel})^2 + p_{i\perp}^2 + \mu_i^2}, p_{i\perp}, x_i p_{\parallel} \right) \quad (17.57)$$

but within the above conditions Eq. (17.57) can be approximated by $p_i^{\mu} \simeq x_i p^{\mu}$. Using the condition that the parton is free before and after the scattering ($p_i^2 = \mu_i^2$, see Fig. 17.8b)

$$\mu_1^2 = (x_1 p)^2 = (q + x_1 p)^2 \quad (17.58a)$$

$$\therefore x \equiv x_i = -\frac{q^2}{2(p \cdot q)} = \frac{Q^2}{2M\nu} \quad (17.58b)$$

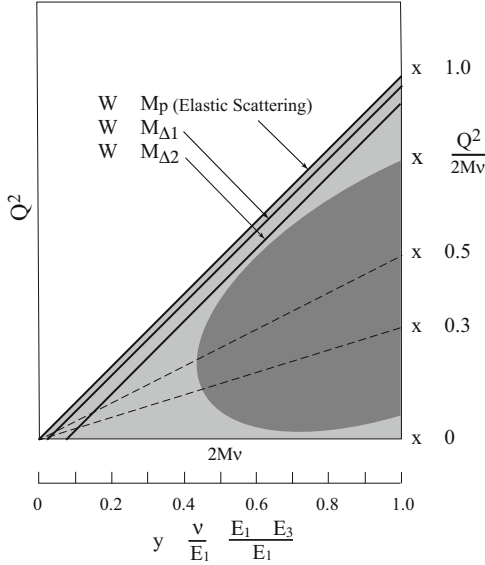


Figure 17.9 Kinematics of deep inelastic scattering. The gray triangular area indicates the allowed region. Elastic scattering is represented by the line with $W = M_p$ ($x = 1$) and quasielastic by $W = M_{\Delta_1}, M_{\Delta_2}$. The darker

gray half-oval corresponds to the deep inelastic scattering (DIS) region. The higher the energy, the closer the DIS region comes to the boundary.

In summary, conditions for the impulse approximation to hold are

$$E \gg M, \quad Q^2 \gg M^2, \quad 2M\nu \gg M^2 \quad (17.59a)$$

$$0 < x = \frac{Q^2}{2M\nu} < 1 \quad (17.59b)$$

These are the same conditions as for Bjorken scaling to hold. The scattering that satisfies the conditions is called deep inelastic scattering. The coordinate frame satisfying the conditions is called the infinite momentum (p_∞) frame. Figure 17.9 illustrates the allowed region for Q^2 , $2M\nu$, and how x , y , W are related to them. The variable y is introduced for later applications. According to the observed data, the scaling seems to hold early for $Q \gtrsim 1 \text{ GeV}/c$ (Fig. 17.6b). We next investigate the physical meaning of the structure functions.

17.5.2

Electron–Parton Scattering

Scaling Variables As will be clarified later, the partons consist of quarks and gluons. The gluons are electrically neutral and do not contribute to electron–proton scattering. Here, we assume the parton is a quark. It is a point-like particle having spin $s = 1/2$, mass μ and electric charge $q_{\text{parton}} = \lambda_q e$. The scattering cross section

in the parton's rest frame is given by making the changes $M \rightarrow \mu$, $\alpha \rightarrow \lambda_q \alpha$ in Eq. (17.12)

$$\left. \frac{d\sigma}{d\Omega} \right|_{eq, \text{LAB}} = \frac{4(\lambda_q \alpha)^2}{Q^4} \frac{\omega'^3}{\omega} \cos^2\left(\frac{\theta}{2}\right) \left[1 + \frac{Q^2}{2\mu^2} \tan^2\left(\frac{\theta}{2}\right) \right] \quad (17.60)$$

In reality, the parton is inside the nucleon with relative momentum given by Eq. (17.59b). Therefore we need to define Lorentz-invariant variables and rewrite Eq. (17.60) in terms of them. Let us first show that the variables so far treated as those in the laboratory frame can, in fact, be considered as Lorentz-invariant variables. Neglecting both proton and electron masses in the limit $E \gg M$, the following equalities hold in the LAB frame:

$$s \equiv (k + p)^2 \simeq 2(k \cdot p) = 2\omega M \quad (17.61a)$$

$$y \equiv \frac{(p \cdot q)}{(p \cdot k)} = \frac{\omega - \omega'}{\omega} = \frac{\nu}{\omega} = \frac{2M\nu}{s} \quad (17.61b)$$

$$x \equiv \frac{Q^2}{2M\nu} = \frac{Q^2}{s\gamma} = \frac{s(1-y)\sin^2(\theta/2)}{M^2\gamma} \quad (17.61c)$$

$$\sin^2\left(\frac{\theta}{2}\right) = \frac{M^2 x \gamma}{s(1-y)}, \quad \cos^2\left(\frac{\theta}{2}\right) = 1 - \frac{M^2 x \gamma}{s(1-y)} \simeq 1 \quad (17.61d)$$

One sees that variables in the laboratory frame [ω , ν , $\sin^2(\theta/2)$, etc.] can be expressed in terms of Lorentz-invariant variables (s , x , y).

Problem 17.6

Confirm

$$0 < \frac{Q^2}{2M\nu} (= x) < 1, \quad 0 < \frac{2M\nu}{s} (= y) < 1 \quad (17.62)$$

Parton Distribution Function For elastic scattering in the laboratory frame (LAB), θ and ω' are related as given in Eq. (17.31). Marking variables in the center of mass frame (CM) with asterisk, we can rewrite y in CM as

$$y = \frac{(p \cdot q)}{(p \cdot k)} = \frac{E^*(\omega^* - \omega'^*) - (-\mathbf{k}^*) \cdot (\mathbf{k}^* - \mathbf{k}'^*)}{E^*\omega^* - (-\mathbf{k}^* \cdot \mathbf{k}^*)} \simeq \frac{1 - \cos \theta^*}{2} \quad (17.63)$$

We obtain

$$\frac{d\sigma}{d\gamma} = \omega \left. \frac{d\sigma}{d\omega'} \right|_{\text{LAB}} = \frac{4\pi M^2}{s(1-y)^2} \left. \frac{d\sigma}{d\Omega} \right|_{\text{LAB}} = 4\pi \left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} \quad (17.64)$$

The above equation holds for two-body scattering with target mass M . When we apply it to the electron-parton two-body system, we make the replacements $M \rightarrow \mu$, $s \rightarrow s_q = (k + p_q)^2$. Noting that for elastic scattering

$$Q^2 = 2\mu\nu = s_q y \quad (17.65)$$

we rewrite Eq. (17.60) to obtain

$$\left. \frac{d\sigma}{dy} \right|_{eq} = \frac{4\pi\lambda_q^2\alpha^2}{Q^4} s_q \left[(1-y) + \frac{y^2}{2} \right] \quad (17.66)$$

As the parton's momentum in the nucleon is given by $x p$, we have

$$s_q = (k + x p)^2 \simeq 2x(k \cdot p) \simeq x s \quad (17.67)$$

Writing the i -th parton's momentum distribution function in the nucleon as $q_i(x)$, the probability that the momentum is in the interval from x to $x + dx$ is given by $q_i(x)dx$. Then the scattering cross section of the electron with the i -th parton in the nucleon is given by

$$\frac{d\sigma}{dy} = \frac{4\pi\alpha^2}{Q^4} s \left[(1-y) + \frac{y^2}{2} \right] \lambda_i^2 x q_i(x) dx \quad (17.68)$$

Consequently, the electron–nucleon scattering cross section is obtained by summing over all the partons and becomes

$$\frac{d^2\sigma}{dx dy} = \frac{4\pi\alpha^2}{Q^4} s \left[(1-y) + \frac{y^2}{2} \right] \sum_i \lambda_i^2 x q_i(x) dx \quad (17.69)$$

Note that the cross section is factorized as a product of two functions each containing x or y only. We have derived the e – p scattering cross section as a sum of e – q scatterings. Note, the partons are elastically scattered by the electron, but the recoil parton is not an observable by itself. The quarks, also known as partons, are confined, and when they come out of the nucleon they are observed as pions and many other hadrons. However, characteristics of e – q elastic scattering can be obtained by observing the behavior of the electron only. Note that Eq. (17.69) is expressed only in terms of electron variables, which can be observed experimentally. Rewriting the generic expression Eq. (17.41) for inelastic scattering in terms of x and y (Problem 17.7), we have

$$\left. \frac{d^2\sigma}{dx dy} \right|_{\text{inel}} = \frac{4\pi\alpha^2}{Q^4} s \left[\nu W_2(1-y) + 2x M W_1 \left(\frac{y^2}{2} \right) \right] \quad (17.70)$$

This equation becomes in the scaling limit

$$\left. \frac{d^2\sigma}{dx dy} \right|_{\text{inel}} = \frac{4\pi\alpha^2}{Q^4} s \left[F_2(x)(1-y) + 2x F_1(x) \left(\frac{y^2}{2} \right) \right] \quad (17.71)$$

Comparing Eq. (17.71) with Eq. (17.69), we see that in the quark–parton model the structure function is given by

$$F_2(x) = x \sum \lambda_i^2 q_i(x) \quad (17.72a)$$

$$F_1(x) = \frac{1}{2} \sum \lambda_i^2 q_i(x) \quad (17.72b)$$

We have proved that by considering e - p inelastic scattering as a sum of e - q elastic scatterings, the structure functions $M W_1(Q^2, \nu)$ and $\nu W_2(Q^2, \nu)$, both of which are functions of two independent variables, become, in the scaling limit, functions of $x = Q^2/2M\nu$ only. Bjorken scaling is evidence that the hadrons have substructure and are made of point-like particles. We have also succeeded in finding a physical interpretation of the structure functions in Bjorken scaling. They represent the momentum distribution of partons in the nucleon.

Problem 17.7

Show that Eq. (17.41) becomes Eq. (17.70) if one rewrites it in terms of x and y .

17.6

Scattering with Equivalent Photons

In this section we try to interpret deep inelastic scattering in terms of equivalent photons, those emitted by the electron. It offers a different perspective and helps to give an intuitive interpretation of the picture. Changing the way of expression does not necessarily increase information, but in some cases it is easier to interpret. Besides, the viewpoint presented in this section can be directly extended to QCD applications.

17.6.1

Transverse and Longitudinal Photons

The formalism we have developed so far is based on the assumption that one-photon exchange gives a good approximation. Then what the nucleon sees is not the electron but a photon. If we take the view that the nucleon interacts with the photon, the electron is just a supplier of photons. The cross section can be considered as a convolution of the photon-nucleon interaction cross section and the intensity of the photon provided by the electron integrated over the photon spectrum. The cross section formula expressed this way is called the Weizsäcker-Williams approximation [78, 96].

The interaction of the photon with a charged particle is given by

$$\mathcal{L}_{\text{INT}} = -Qe \int d^4x A_\mu(x) J^\mu(x) \quad (17.73)$$

the transition amplitude is given by

$$T_{fi} = (2\pi)^4 \delta^4(q + p - P_F) \epsilon_\mu(\lambda) \langle P_F | J^\mu(0) | p \rangle \quad (17.74)$$

and the photon absorption cross section in the LAB frame can be written as

$$\begin{aligned}
 d\sigma_\gamma &= \frac{e^2}{4KM} \sum_F (2\pi)^4 \delta^4(\epsilon + p - P_F) \\
 &\quad \times \frac{1}{2} \sum_{\text{spin}, \lambda} \langle P_F | \epsilon_\mu(\lambda) J^\mu | p \rangle \langle P_F | \epsilon_\nu(\lambda) J^\nu | p \rangle^* \\
 &= \frac{4\pi^2 \alpha}{K} \sum_\lambda \epsilon_\mu(\lambda) \epsilon_\nu(\lambda)^* W^{\mu\nu}
 \end{aligned} \tag{17.75}$$

where $W^{\mu\nu}$ is the same structure function for the e - p inelastic scattering as given by Eq. (17.34c). K is the photon energy in the LAB frame, M the target mass and $\epsilon_\mu(\lambda)$ the polarization vector of helicity λ .

There are two problems associated with the photon being virtual. One is the appearance of an additional ($\lambda = 0$) polarization. The virtual photon is like a massive vector particle but its four-momentum q is space-like ($q^2 < 0$). The polarization vectors have to be orthogonal to the photon's energy-momentum vector q , namely,

$$q \cdot \epsilon(\lambda) = q_\mu \epsilon^\mu(\lambda) = 0 \quad \lambda = \pm, 0 \tag{17.76}$$

If the particle is a massive free vector particle, q is time-like ($q^2 > 0$) and all three polarization vectors satisfying Eq. (17.76) are space-like. But the virtual photon is space-like, and one of the polarization vectors must be time-like. Consequently, we define three polarization vectors as

$$\epsilon^\mu(\pm) = \mp(0, 1, \pm i, 0)/\sqrt{2} \tag{17.77a}$$

$$\epsilon^\mu(0) = \frac{1}{\sqrt{-q^2}} (q_3, 0, 0, q_0) = \frac{1}{\sqrt{-q^2}} \left(\sqrt{\nu^2 - q^2}, 0, 0, \nu \right) \tag{17.77b}$$

It is possible to make $\nu = 0$ by choosing a frame referred to as the Breit frame, which we will discuss soon. In this frame

$$\epsilon^\mu(0) = (1, 0, 0, 0) \tag{17.78}$$

For this reason, the photon with $\lambda = 0$ is sometimes called a scalar photon.²⁾ The polarization vectors satisfy orthogonality and completeness conditions:

$$\epsilon(\lambda) \cdot \epsilon(\lambda') = \epsilon_\mu(\lambda) \epsilon^\mu(\lambda') = (-1)^\lambda \delta_{\lambda\lambda'} \tag{17.79a}$$

$$\sum_\lambda (-1)^{\lambda+1} \epsilon^\mu(\lambda) \epsilon^\nu(\lambda) = - \left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2} \right) \tag{17.79b}$$

The metric is different from the ordinary definition of the completeness condition. The metric used here takes into account that $\epsilon_\mu(0)$ is time-like [$\epsilon(0) \cdot \epsilon(0)^* = +1$]. See also arguments in Appendix C.

2) Confusingly though, it is conventionally referred to as the longitudinal photon.

Problem 17.8

Show that the polarization vector of the scalar photon can be expressed as

$$\epsilon^\mu(0) = a \left(p^\mu - \frac{(p \cdot q)}{q^2} q^\mu \right), \quad a = \left(p^2 - \frac{(p \cdot q)^2}{q^2} \right)^{-1/2} \quad (17.80)$$

where p^μ is the four-momentum of the target proton.

The second problem is the definition of the incident photon energy K . For the virtual photon, any variable that converges to the real photon energy in the limit $Q^2 \rightarrow 0$ can be a candidate. Here we adopt a definition of the photon energy in terms of the flux Eq. (6.84) [168].³⁾

$$4KM|_{\text{LAB}} = 4[(q \cdot p)^2 - q^2 p^2]^{1/2} = 4M(\nu^2 + Q^2)^{1/2} \quad (17.82)$$

$$\therefore K = \sqrt{\nu^2 + Q^2}$$

Using Eqs. (17.75), (17.77) and (17.79), we can express the cross sections for scattering by transverse and longitudinal photons as

$$\sigma_T \equiv \frac{1}{2} \frac{4\pi^2\alpha}{K} \sum_{\lambda=\pm 1} \epsilon_\mu(\lambda) \epsilon_\nu(\lambda)^* W^{\mu\nu} = \frac{4\pi^2\alpha}{K} W_1 \quad (17.83a)$$

$$\sigma_L \equiv \frac{4\pi^2\alpha}{K} \epsilon_\mu(0) \epsilon_\nu(0)^* W^{\mu\nu} = \frac{4\pi^2\alpha}{K} W_L \quad (17.83b)$$

$$W_L = \left[\left(1 + \frac{\nu^2}{Q^2} \right) W_2 - W_1 \right] \quad (17.83c)$$

In the region where the scaling law holds

$$\sigma_T \rightarrow \frac{4\pi^2\alpha}{KM} F_1 \quad (17.84a)$$

$$\sigma_L \rightarrow \frac{4\pi^2\alpha}{KM} F_L \quad (17.84b)$$

$$F_L = \frac{1}{2x} (F_2 - 2xF_1) \quad (17.84c)$$

Problem 17.9

Derive Eq. (17.83).

3) Hand's definition [194], which can be defined by observable variables of the final state only, is used frequently, too:

$$K = (W^2 - M^2)/2M = \nu - Q^2/2M \quad (17.81)$$

17.6.2

Spin of the Target

Callan–Gross Relation As is clear from Eq. (17.72), when the target has spin $s = 1/2$, the equality

$$F_2 = 2x F_1 \quad (17.85)$$

holds, which is referred to as the Callan–Gross relation [85]. In this case, $\sigma_L/\sigma_T = 0$. On the other hand, if the target has spin $s = 0$, there is no magnetic scattering, which means $F_1 = 0$ or $\sigma_T/\sigma_L = 0$. In other words, interpreting the reaction as that by the photon, the target spin can be easily recognized.

Breit Frame An intuitive explanation can be given in the Breit frame (also called the brick wall frame). In this frame (Fig. 17.10), the electron, moving in the x – z plane, changes its momentum in the z -direction only, without changing p_x . In other words, there is no energy transfer and it gives only recoil momentum to the target. On the other hand, the target proton (or parton inside) comes in from the $+z$ direction, receives recoil momentum only and bounces back in the $+z$ direction. In this frame, both particles move as if they had bounced off a brick wall.

In terms of the virtual photon, it comes in from $-z$ direction and is absorbed by the proton. Since a transversely polarized photon gives spin component $s_z = 1$ to the target, a $s = 0$ target cannot absorb the photon and $\sigma_T = 0$. If the target has $s = 1/2$, spin transfer $\Delta s_z = 1$ causes spin flip. If the photon is longitudinally polarized, spin transfer $\Delta s_z = 0$ may also be possible, allowing it to interact with both photons. But notice that the electromagnetic interaction is of vector type, which keeps the particles chirality invariant. This means the particle preserves its helicity (namely its spin has to be flipped) in the mass-zero limit and forces $\sigma_L = 0$. Exper-

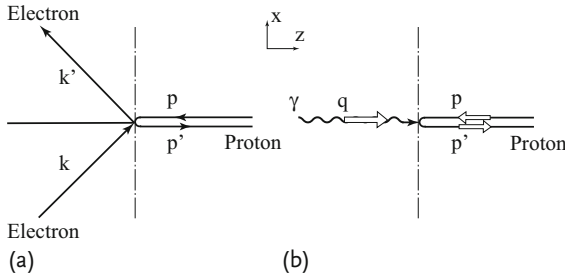


Figure 17.10 In the Breit frame, the electron and the proton recoil from each other without changing energy and p_x but changing p_z , as if they had bounced off a brick wall. The proton enters from the $+z$ direction and recoils back in the $+z$ direction, while the photon ejected by the electron enters with transverse momentum q , which is absorbed by the proton.

As the electron's spin is $1/2$ and the virtual photon's 1 (if transverse wave) or 0 (if longitudinal wave), angular momentum conservation dictates that only the transverse wave can be absorbed by the target parton. Therefore $\sigma_L = 0$. On the other hand, if the target parton has spin 0 , it can absorb the longitudinal wave only.

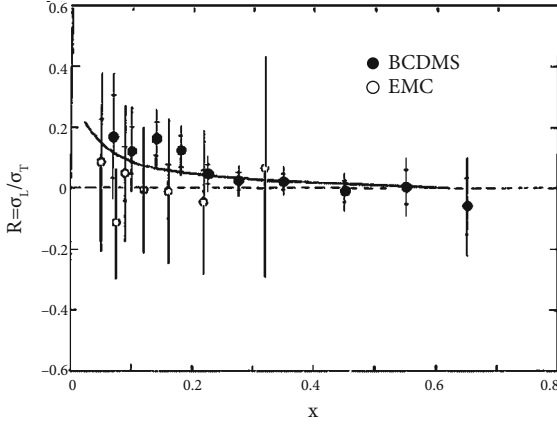


Figure 17.11 The Callan–Gross equality means $R = \sigma_L/\sigma_T = 0$. Observation shows it is generally true except in the region $x \lesssim 0.1$ [16, 59].

imentally, partons in the proton have dominant $s = 1/2$ components (Fig. 17.11) as expected.

Problem 17.10

When the target has mass μ or transverse momentum $p_\perp^2$, show that

$$\frac{\sigma_L}{\sigma_T} = \frac{4(\langle p_\perp^2 \rangle + \mu^2)}{Q^2} \quad (17.86)$$

Hint: Calculate the cross section Eq. (17.83) in the Breit frame. In this frame the virtual photon has momentum only and no energy. The scalar polarization is expressed in a simple form, too:

$$q^\mu = (0, 0, 0, \sqrt{-q^2}), \quad \epsilon^\mu(0) = (1, 0, 0, 0) \quad (17.87)$$

Note, if $\mu^2 + p_\perp^2 \sim O(0.25 \text{ GeV}^2)$, then $\sigma_L/\sigma_T \simeq O(1/Q^2)$. If they have x dependence, Fig. 17.11 can be explained.

17.6.3

Photon Flux

Highlighting the role of the photon in the deep inelastic scattering, we also introduce the concept of the effective photon flux that the electron provides.

We can rewrite the cross section as

$$\frac{d^2 d\sigma}{d\Omega d\omega'}(ep \rightarrow e'X) = \Gamma(\sigma_T + \epsilon\sigma_L) \quad (17.88a)$$

$$\Gamma = \frac{\alpha K}{2\pi^2 Q^2} \frac{\omega'}{\omega} \frac{1}{1-\epsilon}, \quad \epsilon = \left[1 + 2 \frac{\nu^2 + Q^2}{Q^2} \tan^2 \frac{\theta}{2} \right]^{-1} \quad (17.88b)$$

Using $d^3 k' = dk' dk_T^2/\pi = \omega(1-\gamma)d\gamma dk_T^2/\pi$ and Eqs. (17.61), the above expression becomes in the infinite momentum frame

$$\begin{aligned} d\sigma(ep \rightarrow e'X) &= [\sigma_T + \epsilon(\gamma)\sigma_L] \Gamma d\Omega d\omega' \xrightarrow{p^\infty \text{ frame}} \sigma_T dP(\gamma, p_T^2) d\gamma \\ dP(\gamma, p_T^2) d\gamma &= \frac{\alpha}{2\pi} P_{\gamma e}(\gamma) \frac{dp_T^2}{p_T^2} d\gamma, \quad p_T^2 = \left(2k' \sin \frac{\theta}{2} \right)^2 \\ P_{\gamma e}(\gamma) &= \left[\frac{1 + (1-\gamma)^2}{\gamma} \right] \end{aligned} \quad (17.89)$$

One sees that the quantity $dP(\gamma, p_T^2)d\gamma$ can be interpreted as the photon flux prepared by the electron. Equations (17.89) are called the Weizsäcker–Williams formula [78, 96, 381, 390].

Problem 17.11

Derive Eqs. (17.88).

Problem 17.12

Derive Eqs. (17.89).

The expression diverges for $p_T^2 \rightarrow 0$, which is the familiar Coulomb force singularity at $r = 0$. Integration over p_T^2 can be carried out as follows. The denominator p_T^2 is really Q^2 in the original expression, which is expressed in the infinite momentum frame as

$$Q^2 = -(k - k')^2 \simeq \frac{p_T^2}{1-\gamma} + \frac{m_e^2 \gamma^2}{1-\gamma} \quad (17.90)$$

Therefore, we can set the lower limit of p_T^2 integration to not 0 but $m_e^2 \gamma^2$ if the divergence is to be avoided. The upper limit of the integration is given by $(p_T^2)_{\max} \sim (k' \theta_{\max})^2 \sim \omega^2 (1-\gamma)^2 \theta_{\max}^2$. Neglecting $\ln \theta_{\max}^2$ compared to $\ln(\omega^2/m_e^2)$, we have

$$\int d \ln p_T^2 \simeq 2 \ln (\omega / m_e) \quad (17.91)$$

which gives the integral form

$$\sigma(eN \rightarrow e'N') = \frac{\alpha}{\pi} \ln \frac{\omega}{m_e} \int \frac{1 + (1-\gamma)^2}{\gamma} \sigma(\gamma N \rightarrow N'; \gamma) d\gamma \quad (17.92)$$

Equation (17.92) is also referred to as the Weizsäcker–Williams formula in integral form.

The Weizsäcker–Williams formula considers electron–proton scattering as a two-step process: the electron–photon and photon–proton reactions. The electron’s role is simply to provide the photon flux. The formula is valid only in the infinite momentum frame but it becomes a very useful tool in QCD, where such frames are routinely provided.

17.7

How to Do Neutrino Experiments

Deep inelastic scattering by the electron clarified the parton structure of the nucleon. We now want to do the same experiments with the neutrino to take advantage of its ability to distinguish the quark flavor. It took 25 years before the neutrino was detected after Pauli’s prediction in 1930 and another 30 years before high-intensity neutrino beams could be used for systematic investigations. The neutrino is the only particle that participates exclusively in the weak interaction.⁴⁾ However, once it became possible to do experiments with intense neutrino beams, they provided an ideal probe to investigate the characteristics of the weak interaction, which under normal circumstances is hidden behind the strong and/or electromagnetic interactions. The neutrino is electrically neutral, and because of its weak reaction rate with matter, neutrino reactions do not happen easily. To do neutrino experiments by actively detecting it, construction of a large accelerator and commensurate development of technology were indispensable. We describe briefly the essence of how to do neutrino experiments.

17.7.1

Neutrino Beams

High-energy ($\gg 10$ MeV) and high-intensity neutrino $\nu_\mu(\bar{\nu}_\mu)$ beams are mainly obtained using

$$\begin{aligned}\pi^+(\pi^-) &\rightarrow \mu^+(\mu^-) + \nu_\mu(\bar{\nu}_\mu) \\ K^+(K^-) &\rightarrow \mu^+(\mu^-) + \nu_\mu(\bar{\nu}_\mu)\end{aligned}\tag{17.93}$$

To obtain $\nu_e(\bar{\nu}_e)$

$$\begin{aligned}\mu^+(\mu^-) &\rightarrow e^+(e^-) + \nu_e(\bar{\nu}_e) + \bar{\nu}_\mu(\nu_\mu) \\ K^+(K^-) &\rightarrow \pi^0 + e^+(e^-) + \nu_e(\bar{\nu}_e)\end{aligned}\tag{17.94}$$

are used. For low-energy $\bar{\nu}_e$ neutrinos (1–4 MeV), β decays of fission material produced in atomic reactors are also used.

4) Except for gravity, which affects everything that has energy. Gravity becomes important when cosmic neutrinos are discussed.

In early experiments before about 1970, the requirement of a higher intensity beam was much stronger than its quality. For this purpose, it was necessary to focus pions and kaons produced by the primary proton beam in the forward direction. It is impossible to focus all of them but a horn nearly achieves it. For realistic explanations we take examples from Fermilab and CERN in the 1970s, which supplied primary protons with energy ~ 400 GeV. The horn is shaped as shown in Fig. 17.12a, with a cross-sectional view through its center shown in (b). When a large current (say $\sim 10^4$ A) is supplied to the horn, a circular magnetic field outside and surrounding the horn is generated. Such a large current can only be made by storing charge in a condenser (capacitor) bank and discharging it instantaneously by shorting the circuit. Because of this, the beam has to be condensed in time (fast extraction), which makes the operation of neutrino experiments incompatible with others, because most of them require slow extraction to use time coincidence techniques to select reactions. In this sense, neutrino experiments are costly.

The target is placed at the mouth of the horn, where it is irradiated by the proton beam. Produced secondary pions (and kaons) go forward, but those emitted at a large angle enter the magnetic field and are bent forward. The horn is designed to focus secondary particles emitted at almost any angle within a limit. The particles then go through a long (~ 350 m) tunnel, where they decay to produce neutrinos. Pions that did not decay are stopped by a thick shield at the end of the tunnel. Since hadrons interact strongly, a concrete or iron shield of a few meters is enough to absorb all of them. But muons interact only electromagnetically and lose typically only ~ 1.8 MeV/g. To filter muons of, for example, 200 GeV, it takes $\sim 10^5$ g = 100 kg = 120 m of iron. If cost matters, earth may be substituted, but its density is

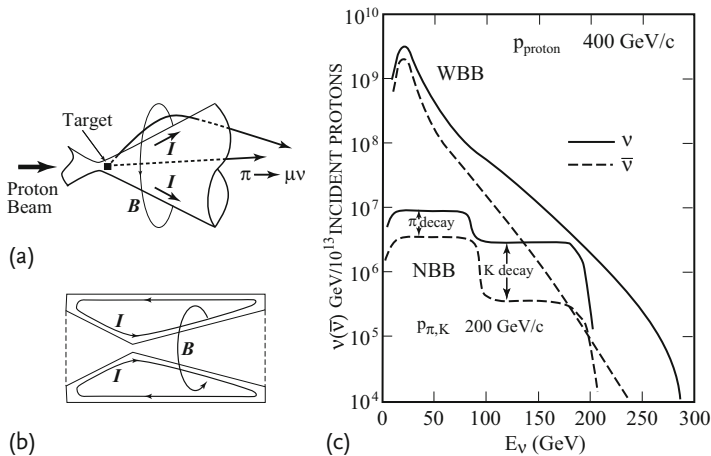


Figure 17.12 (a) Principle of horn focusing. When current I is supplied to the cylindrical horn, a circular magnetic field is produced, which focuses one sign of charged pions in the forward direction. (b) Cross-sectional view

of the horn. (c) WBB (wide band beam) and NBB (narrow band beam) spectra of the neutrino beam. NBB is produced by monochromatic secondary pions and kaons [352].

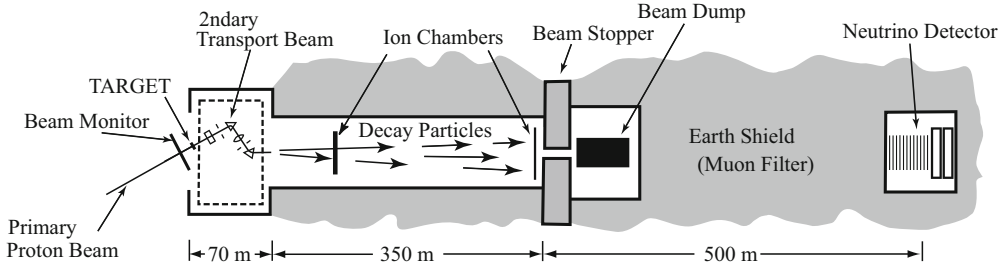


Figure 17.13 Layout of the dichromatic neutrino beam line at Fermilab circa 1970. The primary proton beam hits the target and produces pions (and kaons). The pions are made monochromatic by a beam focusing system consisting of dipole and quadrupole magnets,

then they are injected into the decay pipe, where they decay to muons and neutrinos. Undecayed pions are absorbed by a beam dump and muons are stopped by muon filters made of earth (or iron).

small ($2\text{--}3 \text{ gr/cm}^{-3}$), and the necessary length would be ~ 3 times, which causes 10 times intensity loss [$\propto 1/(\text{distance})^2$]. The neutrino beam obtained in this way has a wide spectrum and is called a wide band neutrino beam (WBB).

A typical spectrum is shown in Fig. 17.12c. Oppositely charged particles are bent in opposite directions and diverge, which can be used to separate $\bar{\nu}$ from ν somewhat, but particles in the forward direction inside the horn are not separated, which is the reason for a large mixture of $\bar{\nu}$ in ν . Figure 17.13 shows the layout of a Fermilab neutrino beam. A dichromatic beam is described below but the overall layout is similar to that of a WBB beam, except that the horn is replaced by a set of dipole and quadrupole magnets.

As the technology advanced, one wanted to have a beam of good quality with good control of energy and intensity. For this purpose narrow band beams (NBBs) began to be used. The neutrino being neutral, its energy cannot be measured directly, but taking advantage of two-body decay [$\pi(K) \rightarrow \mu \nu$] in producing it, one can determine its energy kinematically. If we mark variables in the CM frame with $*$, the neutrino energy in the LAB frame is connected to those in the CM frame by Lorentz transformation:

$$E_{\nu \text{ LAB}} = \gamma(E_{\nu}^* + \beta p_{\nu}^* \cos \theta^*) \quad (17.95a)$$

$$p_{\nu}^* = E_{\nu}^* = (M^2 - m_{\mu}^2)/2M \quad (17.95b)$$

where M is the parent's (π or K) mass and β, γ are Lorentz factors that are expressed in terms of the parent's energy (E) and momentum (P) as

$$\gamma = E/M, \quad \beta\gamma = P/E \quad (17.96)$$

As the parent is spinless, the decay distribution in the CM frame is isotropic. Therefore, the distribution of particles in the LAB frame is

$$\frac{dN}{N} = \frac{1}{4\pi} d\Omega^* = \frac{1}{2} d\cos \theta^* = \frac{dE_{\text{LAB}}}{2\beta\gamma E_{\nu}^*} \sim \text{const. } dE_{\text{LAB}} \quad (17.97)$$

Namely, the energy spectrum of the neutrino is flat between the minimum and the maximum of its energy

$$\min(E_{\nu \text{ LAB}}) = E_{\nu}^* \gamma (1 - \beta) \simeq 0 \quad (17.98a)$$

$$\max(E_{\nu \text{ LAB}}) = E_{\nu}^* \gamma (1 + \beta) \simeq 2\gamma \frac{M^2 - m_{\mu}^2}{2M} = E \left[1 - \frac{m_{\mu}^2}{M^2} \right] \quad (17.98b)$$

$$= \begin{cases} 0.427 E & \pi \text{ decay} \\ 0.954 E & K \text{ decay} \end{cases} \quad (17.98c)$$

Examples of such flat spectra are shown in Fig. 17.12c. One sees that the maximum neutrino energy from K decay is about two times that from π decay. However, the pion production cross section is much higher than that of kaons. As the decay angle θ_{ν} and the energy E_{ν} of the neutrino in the LAB frame are related by (Fig. 17.14a)

$$E_{\nu} = \frac{M^2 - m_{\mu}^2}{2(E - P \cos \theta_{\nu})} \quad (17.99)$$

if there is independent information about the neutrino decay angle θ_{ν} , its energy can be determined. Using $E \gg M, \cos \theta \simeq 1 - \theta^2/2$, we obtain

$$E_{\nu} \simeq E_{\nu \text{ max}} \frac{\theta_0^2}{\theta^2 + \theta_0^2}, \quad \theta_0 = M/P \quad (17.100)$$

The distribution curve is plotted in Fig. 17.14b. θ_{ν} can be determined from the distance of the interaction point at the detector from the beam center. Figure 17.14c shows a scatter plot of θ_{ν} versus E_{ν} . The two bands correspond to neutrinos from K and π . In this way, at fixed angle one can obtain a dichromatic beam with two energy peaks. The low-lying flat spectra of Fig. 17.12c are integrated spectra of the dichromatic beam.

Recent Developments More recently, high intensity monochromatic beams of varying energy have become available [362]. This apparatus uses horns with a secondary beam set at an off-axis angle from the primary beam, taking advantage of the differing spectra of the secondaries depending on the production angle. The primary beam intensity is also increased. The so-called superbeam uses a proton beam having 1–4 MW power.

With future neutrino oscillation experiments in mind, the concepts of “neutrino factory” and “beta beam” are being developed [42]. The former uses an intense tertiary muon beam that is phase rotated and cooled⁵⁾ for better optical properties, accelerated, stored and allowed to decay. In this way 10^{20} muons/year are envisaged. The beta beam accelerates β -active unstable isotopes, stores them and

5) Phase rotation is an accelerator technique that uses a magnetic field to make the energy spread narrow at the expense of making time spread wide or vice versa. Cooling is another technique to reduce transverse divergence of the beam. Those techniques are necessary because the secondary and tertiary beams spread widely both in energy and direction.

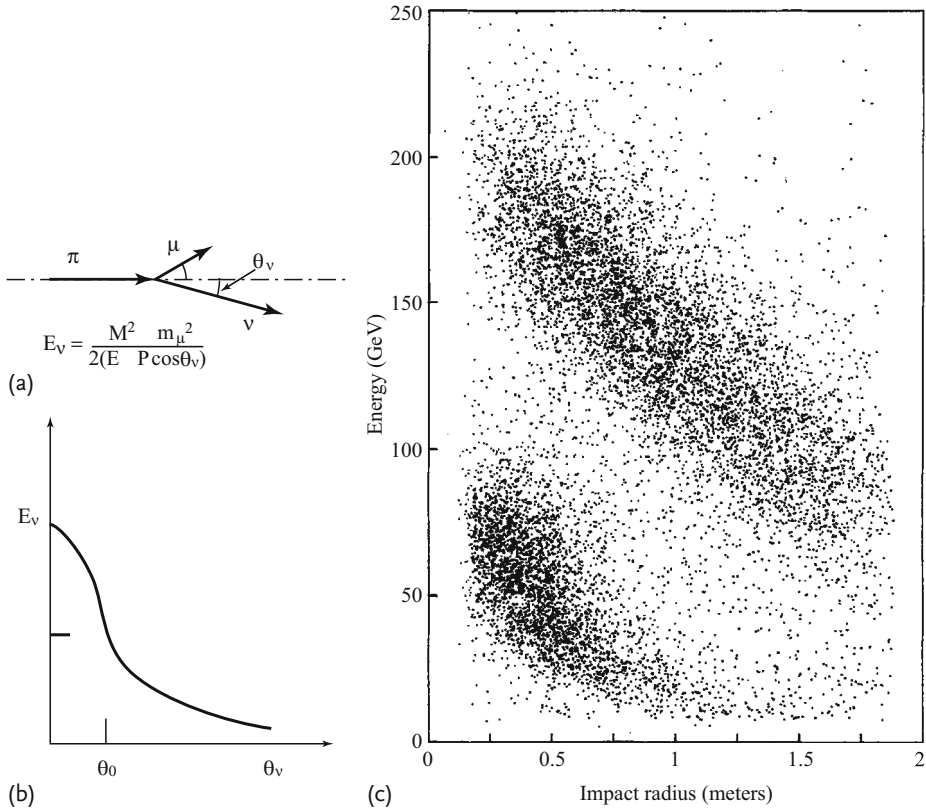


Figure 17.14 (a) If the parent particle is monochromatic, the neutrino energy E_ν and the decay angle θ_v are related. (b) When $E \gg M_{\pi, K}$, $E_\nu \simeq E_{\nu, \max} \times \theta_0^2 / (\theta^2 + \theta_0^2)$. (c) Scatter plot of measured neutrino ener-

gy (muon + hadron energy) versus detector radius for a sample of events. The separate bands are due to neutrinos from pion and kaon decay [116].

lets them decay. Since β decay has very low Q value in the isotope's rest frame, neutrinos from accelerated isotopes have good quality. They are monochromatic, collimated and with virtually no background.

17.7.2

Neutrino Detectors

Event rates of neutrino experiments are given by $N = N_\nu \sigma_{\text{tot}} N_{\text{target}}$. Since σ_{tot} is the subject of the investigation, it is beyond the control of the experimenters. However, the fact that the cross section increases in proportion to the neutrino energy ($\sigma_{\text{tot}} \propto E_\nu$) can be used to the experimenters' advantage by raising the energy. Making N_ν and N_{target} large is where the experimenters' skill comes in.

In neutrino experiments, one makes a gigantic target to increase N_{target} , but in order not to lose generated events the target itself has to serve as a detector. A bubble chamber is an ideal target cum detector capable of reconstructing all the kinematic variables. Its operating principle was briefly described in Fig. 14.2. It can often record a single event that definitively detects a new phenomenon, depending on what one is pursuing. Good examples are the discovery of the Ω particle to test flavor $SU(3)$ (Fig. 14.6) and a neutral current event to confirm the Glashow–Weinberg–Salam electroweak theory [199, 200].

For precision neutrino experiments, however, one wants to have a more massive target/detector to gain high statistics. The size of the bubble chamber is inherently constrained by its use in a magnetic field. Gargamelle and BEBC (Big European Bubble Chamber) at CERN and the 15 foot bubble chamber at Fermilab are made with the highest technology available using superconducting magnets and yet limited to ~ 10 tons at most because of the mentioned constraint [1]. More suitable is an electronic counter detector, the mass of which can be made arbitrarily large to gain high statistics.

Generated particles in the neutrino interactions other than e and μ are almost all pions and nucleons, which are hadronized remnants of a quark jet (a phenomenon in which many hadrons are ejected in the same direction). Detailed identification of each hadron's motion is not so important as the collective behavior of the jet. The bubble chamber is overqualified for such a purpose. Eliminating unnecessary information drastically and concentrating on detecting the leptons and the total energy of the hadrons, the detector structure can be simpler but massive. Figure 17.15 shows a typical detector, that of the CDHS (CERN–Dortmund–Heidelberg–Saclay) group at CERN [59, 116]. Basically, they are made of iron modules (toroidal and magnetized) with drift chambers and scintillation counters interspersed among the modules to measure muon tracks and hadronic energy. The whole detector is divided into 19 supermodules and each supermodule is further subdivided into 5 or 10 cm thick iron plates. The combination of iron/scintillator modules constitutes a sampling calorimeter for hadrons. The total weight is 1235 ton, and yet the detector is electronics personified, equipped with a great number of readout modules. Information that the detector can supply is:

1. Coordinates of the interaction point: from the transverse position the energy (ω) of the incoming neutrino can be determined using Eq. (17.100).
2. Scattering angle θ of the muon and its energy (ω'): the energy (momentum to be exact) is determined by the curvature of the muon tracks going through the magnetized iron.
3. Hadronic total energy (E_{had}): measured by the hadron calorimeter.

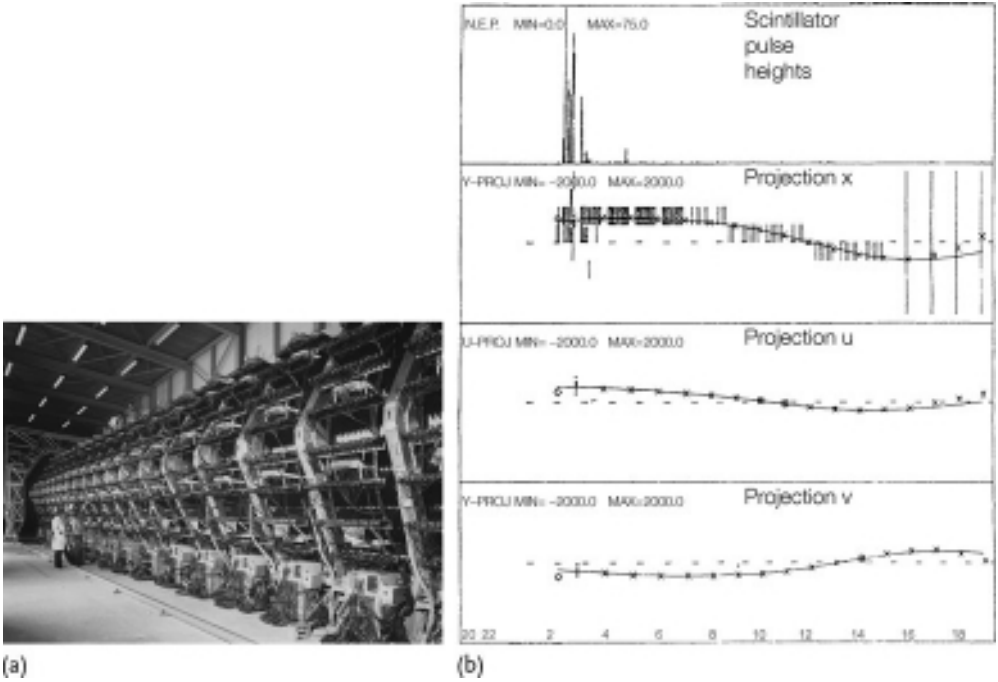


Figure 17.15 (a) View along the 19 modules of the CDHS electronic detector at SPS. The black lightguides and phototubes, which are used to measure the hadron energy, can be seen sticking out of the magnetized iron modules. The hexagonal structures are the drift chambers that measure the muon trajectories [352]. PHOTO CERN. (b) Computer reconstruction of a charged-current neutrino interaction in the detector. The neutrino en-

ters from the left. The top row gives the pulse height in the scintillator planes along the detector. This is a measure of energy deposited to determine the total hadron energy. The panels below show the measured coordinates in the three drift-chamber projections together with the fit curve for the muon reconstruction. The curvature tells the momentum of the muon [3].

From the above information, s , x , y of charged current interactions ($\nu_\mu + N \rightarrow \mu + \text{hadrons}$) can be determined as well as a self-consistency check made ($E_{\text{had}} = \omega - \omega'$).⁶⁾

CDHS used thick iron to gain mass, which is good for detecting muons and hadronic showers, the former because of its long range and the latter because of the collective behavior, but then the electron, which makes a short electromagnetic shower, is absorbed by the iron and undetectable. Charged current interactions induced by $\bar{\nu}_e$ cannot be distinguished from neutral current interactions, which makes active detection of, say, the $\nu_\mu \rightarrow \nu_e$ transition impossible, an important phenomenon known as neutrino oscillation. In this case only the disappearance

6) Neutral current phenomena ($\nu_\mu + N \rightarrow \nu_\mu + \text{hadrons}$), where no muons are observed, can be measured also, but in this case the lepton (neutrino) information is missing and no x information is obtained.

of the ν_μ can be measured by comparing fluxes at different places, which is much more difficult. If one wants to identify electrons and reconstruct their kinematics, one needs fine-grained sampling, which uses more readouts and reduces the target density, hence the total mass. This type would be the next generation of detectors.

17.8

ν - p Deep Inelastic Scattering

17.8.1

Cross Sections and Structure Functions

Feynman diagrams describing neutrino inelastic scattering are shown in Fig. 17.16b. In the low-energy region that we are dealing with ($E_\nu < 100$ GeV), the W mass can be set as infinity and the interaction is well described by the four-Fermi Lagrangian. But to make the comparison with e - p scattering clear we draw the W line in the diagram. Both processes are considered as two-body scattering of quarks by leptons in the parton model. Setting the electron and neutrino kinematics identical and fixing the incoming energy, we see their scattering is a function of x and y only. If the scaling rule holds, x represents the parton's initial momentum, the distribution of which is given by $q(x)$ irrespective of incoming particle species. Denoting leptonic and hadronic weak (charged) current as l^μ and J^μ , respectively, we can express the Lagrangian $\mathcal{L}_F$ as

$$\begin{aligned}
 -\mathcal{L}_F &= \frac{G_F}{\sqrt{2}} (l^{-\mu} J_\mu^+ + l^{+\mu} J_\mu^-) \\
 l^{-\mu} &= \bar{\mu} \gamma^\mu (1 - \gamma^5) \nu, \quad l^{+\mu} = \bar{\nu} \gamma^\mu (1 - \gamma^5) \mu = (l^{-\mu})^\dagger \\
 J_\mu^+ &= \bar{u} \gamma_\mu (1 - \gamma^5) d, \quad J_\mu^- = \bar{d} \gamma_\mu (1 - \gamma^5) u = (J_\mu^+)^\dagger
 \end{aligned} \tag{17.101}$$

The first term in the first line contributes to ν and the second to $\bar{\nu}$ scattering. Defining structure functions for ν , $\bar{\nu}$ separately but in a similar form to e - p scattering,

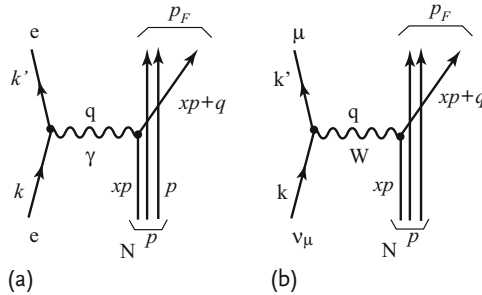


Figure 17.16 (a) e - p inelastic scattering. (b) ν - N inelastic scattering in the quark model.

we have

$$\begin{aligned}
 W_{\rho\sigma}^{\nu,\bar{\nu}} &= \frac{(2\pi)^4}{4\pi M} \sum_F \delta^4(p + q - P_F) \left(\frac{1}{2} \right) \sum_{\text{spin}} \langle p | J_\rho^\mp | P_F \rangle \langle P_F | J_\sigma^\pm | p \rangle \\
 &= \left(-g_{\rho\sigma} + \frac{q_\rho q_\sigma}{q^2} \right) W_1(Q^2, \nu) \\
 &\quad + \left(p_\rho - \frac{(p \cdot q) q_\rho}{q^2} \right) \left(p_\sigma - \frac{(p \cdot q) q_\sigma}{q^2} \right) \frac{W_2(Q^2, \nu)}{M^2} \\
 &\quad - i \epsilon_{\rho\sigma\alpha\beta} p^\alpha q^\beta \frac{W_3(Q^2, \nu)}{2M^2} \\
 &\quad + [\sim (q_\rho q_\sigma)] + [\sim (p_\rho q_\sigma + q_\rho p_\sigma)] + [\sim i(p_\rho q_\sigma - q_\rho p_\sigma)]
 \end{aligned} \tag{17.102}$$

where P_F is sum of momenta of final particles. The weak current contains the axial current, which is not conserved. The constraint due to the current conservation is absent and $W^{\mu\nu}$ cannot be simplified like the electromagnetic current. However, the last term in Eq. (17.102) breaks T invariance, and the three terms in the last line are proportional to lepton mass (divided by incident neutrino energy) when multiplied by the lepton current and can be neglected compared to the first three terms. The third term is a new term, which is introduced by parity violation. The cross section can be calculated to give

$$\begin{aligned}
 \frac{d^2 \sigma^{\nu,\bar{\nu}}}{d\Omega d\omega'} &= \frac{(G_F/\sqrt{2})^2}{4(k \cdot p)} 4K_L^{\rho\sigma} \cdot 4\pi M W_{\rho\sigma}^{\nu,\bar{\nu}} \frac{d^3 k'}{(2\pi)^3 2\omega'} \\
 K_L^{\rho\sigma} &= \frac{1}{4} \sum_{\text{spin}} \bar{u}(k') \gamma^\rho (1 - \gamma^5) u(k) \bar{u}(k) \gamma^\sigma (1 - \gamma^5) u(k')
 \end{aligned} \tag{17.103a}$$

$$= 2 \left[k'^\rho k^\sigma + k'^\sigma k^\rho - (k \cdot k') g^{\rho\sigma} + i \epsilon^{\rho\sigma\alpha\beta} k_\alpha k'_\beta \right] \tag{17.103b}$$

The neutrino being left-handed, a factor 1/2 for spin average is not necessary. Calculating Eq. (17.103), we obtain cross sections for $\nu, \bar{\nu}$ - N scattering:

$$\begin{aligned}
 \frac{d^2 \sigma^{\nu,\bar{\nu}}}{d\Omega d\omega'} &= \frac{G_F^2}{2\pi^2} \omega'^2 \left[2 W_1^{\nu,\bar{\nu}}(Q^2, \nu) \sin^2 \frac{\theta}{2} + W_2^{\nu,\bar{\nu}}(Q^2, \nu) \cos^2 \frac{\theta}{2} \right. \\
 &\quad \left. \mp W_3^{\nu,\bar{\nu}}(Q^2, \nu) \frac{\omega + \omega'}{M} \sin^2 \frac{\theta}{2} \right]
 \end{aligned} \tag{17.104}$$

Changing variables to x, y , we find the structure constants become in the scaling limit

$$\begin{aligned}
 M W_1(Q^2, \nu) &\rightarrow F_1(x), \quad \nu W_2(Q^2, \nu) \rightarrow F_2(x), \\
 \nu W_3(Q^2, \nu) &\rightarrow -F_3(x)
 \end{aligned} \tag{17.105}$$

The minus sign in front of F_3 is for convenience to make $F_3 > 0$. The cross section

is expressed as

$$\frac{d^2\sigma^{\nu,\bar{\nu}}}{dx d\gamma} = \frac{G_F^2}{2\pi} s \left[(1-\gamma) F_2^{\nu,\bar{\nu}}(x) + \left(\frac{\gamma^2}{2}\right) 2x F_1^{\nu,\bar{\nu}}(x) \pm \gamma \left(1 - \frac{\gamma}{2}\right) x F_3^{\nu,\bar{\nu}}(x) \right] \quad (17.106)$$

Problem 17.13

Derive Eq. (17.104).

As the charge symmetry operation ($e^{-i\pi I_y}$) changes $p \rightarrow -n$ and $J^\pm \rightarrow J^\mp$, the following equalities hold:

$$F_i^{\nu p} = F_i^{\bar{\nu} n}, \quad F_i^{\bar{\nu} p} = F_i^{\nu n}, \quad i = 1, 2, 3 \quad (17.107)$$

Then for an isoscalar target ($N_p = N_n$)

$$F_i^\nu = (F_i^{\nu p} + F_i^{\nu n})/2 = (F_i^{\bar{\nu} p} + F_i^{\bar{\nu} n})/2 = F_i^{\bar{\nu}} \quad (17.108)$$

and the difference between σ^ν and $\sigma^{\bar{\nu}}$ is only the sign of F_3 . In conclusion, the $\nu, \bar{\nu}$ - N cross section on an isoscalar target becomes

$$\frac{d^2\sigma^{\nu,\bar{\nu}}}{dx d\gamma} = \frac{G_F^2}{2\pi} s \left[(1-\gamma) F_2^\nu(x) + \left(\frac{\gamma^2}{2}\right) 2x F_1^\nu(x) \pm \gamma \left(1 - \frac{\gamma}{2}\right) x F_3^\nu(x) \right] \quad (17.109)$$

Comparing with Eq. (17.70), we see that, apart from the F_3 term, the neutrino cross section can be formally obtained from the e - p cross section by making the change

$$\frac{4\pi\alpha^2}{Q^4} \rightarrow \frac{G_F^2}{2\pi} \quad (17.110)$$

This is natural if we consider the weak Lagrangian and the equivalent second-order electromagnetic Lagrangian. If the parton spin is $1/2$ and the Callan–Gross equality ($2xF_1 = F_2$) holds, Eq. (17.109) can be rewritten as

$$\frac{d^2\sigma^\nu}{dx d\gamma} = \frac{G_F^2}{2\pi} s \left[\frac{(F_2 + xF_3)}{2} + \frac{(F_2 - xF_3)}{2} (1-\gamma)^2 \right] \quad (17.111a)$$

$$\frac{d^2\sigma^{\bar{\nu}}}{dx d\gamma} = \frac{G_F^2}{2\pi} s \left[\frac{(F_2 + xF_3)}{2} (1-\gamma)^2 + \frac{(F_2 - xF_3)}{2} \right] \quad (17.111b)$$

Note, the structure functions F_i in the ν - N cross section and those in the e - p cross sections differ in general. However, if the parton model is right, the structure functions are the partons' distribution functions in the nucleon and they should be related simply. Let us see how they are related in the quark model.

17.8.2

 $\nu, \bar{\nu}$ - q Scattering

To calculate $\nu(\bar{\nu})$ - q scattering cross sections, we make maximum use of the helicity-tagged e - μ scattering cross section we derived in Chap. 7. As the weak Lagrangian is given by

$$\begin{aligned} -\mathcal{L}_W &= (G_F/\sqrt{2}) [\bar{u}\gamma_\mu(1-\gamma^5)d] [\bar{\mu}\gamma^\mu(1-\gamma^5)\nu] \\ &= \frac{4G_F}{\sqrt{2}} (\bar{u}_L\gamma_\mu d_L)(\bar{\mu}_L\gamma^\mu \nu_L) \end{aligned} \quad (17.112)$$

the expressions for the e - μ scattering cross section decomposed in terms of helicity [Eqs. (7.31)–(7.33)] are directly applicable. We substitute

$$\frac{e^2}{t} \rightarrow 4 \frac{G_F}{\sqrt{2}} \quad (17.113)$$

in the equations for the e - μ cross sections

$$\left. \frac{d\sigma_{LL}}{d\Omega} \right|_{\text{CM}} = \frac{\alpha^2}{t^2} s, \quad \left. \frac{d\sigma_{LR}}{d\Omega} \right|_{\text{CM}} = \frac{\alpha^2}{t^2} s \left(\frac{1 + \cos \theta^*}{2} \right)^2 \quad (17.114)$$

and use $\gamma = (1 - \cos \theta^*)/2$. Since the difference between the neutrino and antineutrino appears only in its helicity as far as the cross section is concerned, conversion from $e\mu$ to νq can be easily made:

$$\frac{d\sigma}{d\gamma} = \frac{G_F^2}{\pi} s : \quad (\nu q, \bar{\nu} \bar{q} \text{ scattering}) \quad (17.115a)$$

$$\frac{d\sigma}{d\gamma} = \frac{G_F^2}{\pi} s (1 - \gamma)^2 : \quad (\nu \bar{q}, \bar{\nu} q \text{ scattering}) \quad (17.115b)$$

Note that we do not take the spin average for neutrino scattering.

When we deal with quarks in the nucleon, we replace s by $x s$ and multiply by the distribution function $q_i(x)$, just as we did in the e - p inelastic scattering. As there are also antiquarks in the nucleon, whose distribution we denote by $\bar{q}_i(x)$, we have

$$\frac{d^2\sigma^\nu}{dx d\gamma} = \frac{G_F^2}{2\pi} s \sum 2x [q_i(x) + \bar{q}_i(x)(1-\gamma)^2] \quad (17.116a)$$

$$\frac{d^2\sigma^{\bar{\nu}}}{dx d\gamma} = \frac{G_F^2}{2\pi} s \sum 2x [q_i(x)(1-\gamma)^2 + \bar{q}_i(x)] \quad (17.116b)$$

Comparing with Eq. (17.111), we obtain

$$F_2^\nu = \sum 2x [q_i(x) + \bar{q}_i(x)] \quad (17.117a)$$

$$xF_3^\nu = \sum 2x [q_i(x) - \bar{q}_i(x)] \quad (17.117b)$$

17.8.3

Valence Quarks and Sea Quarks

There are two kinds of quarks in the nucleon, valence quarks, which define the quantum number of the nucleon (uud in the proton or udd in the neutron), and sea quarks, whose quantum numbers are the same as those of the vacuum. As the sea quarks consist of $q_i \bar{q}_i$ and for them $q_i(x) = \bar{q}_i(x)$, Eq. (17.117) means F_2 represents all the quarks and antiquarks and F_3 the valence quarks only. If we denote the distribution functions of u, d, s quarks in the proton as $u(x), d(x), s(x)$, charge symmetry requires those in the neutron to be $d(x), u(x), s(x)$. If we neglect s quarks in the nucleon and c quark production, charge and lepton number conservation requires that reactions between neutrinos and quarks are limited to

$$\nu_\mu + d \rightarrow \mu^- + u, \quad \nu_\mu + \bar{u} \rightarrow \mu^- + \bar{d} \quad (17.118a)$$

$$\bar{\nu}_\mu + u \rightarrow \mu^+ + d, \quad \bar{\nu}_\mu + \bar{d} \rightarrow \mu^+ + \bar{u} \quad (17.118b)$$

From this we can write the structure functions of the proton and neutron as

$$F_2^{\nu p}(x) = 2x[d(x) + \bar{u}(x)] \quad (17.119a)$$

$$F_2^{\nu n}(x) = 2x[u(x) + \bar{d}(x)] \quad (17.119b)$$

Therefore for isoscalar targets

$$F_2^\nu = (F_2^{\nu p} + F_2^{\nu n})/2 = x[u + d + \bar{u} + \bar{d}] \quad (17.120a)$$

$$xF_3^\nu = x(F_3^{\nu p} + F_3^{\nu n})/2 = x[u + d - \bar{u} - \bar{d}] \equiv x[u_v + d_v] \quad (17.120b)$$

We have marked the distribution function of valence quarks with the subscript v . Equation (17.116) is rewritten as

$$\frac{d^2 \sigma^\nu}{dx d\gamma} = \frac{G_F^2}{2\pi} s x [q(x) + \bar{q}(x)(1 - \gamma)^2] \quad (17.121a)$$

$$\frac{d^2 \sigma^{\bar{\nu}}}{dx d\gamma} = \frac{G_F^2}{2\pi} s x [q(x)(1 - \gamma)^2 + \bar{q}(x)] \quad (17.121b)$$

But this time $q(x) = u(x) + d(x)$, $\bar{q}(x) = \bar{u}(x) + \bar{d}(x)$. For e - N scattering the cross section is proportional to the charge squared and we obtain from Eq. (17.72) the following expressions:

$$F_2^{ep} = x \left[\frac{4}{9}(u + \bar{u}) + \frac{1}{9}(d + \bar{d}) \right] \quad (17.122a)$$

$$F_2^{en} = x \left[\frac{4}{9}(d + \bar{d}) + \frac{1}{9}(u + \bar{u}) \right] \quad (17.122b)$$

Therefore for an isoscalar target

$$F_2^e = \frac{5}{18} x [u + d + \bar{u} + \bar{d}] = \frac{5}{18} x [q(x) + \bar{q}(x)] \quad (17.123)$$

Therefore, if the quark model is right, the structure function for the isoscalar target has to satisfy

$$F_2^\nu = \frac{18}{5} F_2^e \quad (17.124)$$

17.8.4

Comparisons with Experimental Data

Total Cross Section Integrating Eq. (17.121) over x and y

$$\sigma_{\text{tot}}^\nu = \sigma_0(Q + \overline{Q}/3) \quad (17.125a)$$

$$\sigma_{\text{tot}}^{\overline{\nu}} = \sigma_0(Q/3 + \overline{Q}) \quad (17.125b)$$

$$\sigma_0 = \frac{G_F^2}{2\pi} s, \quad Q = \int dx x q(x), \quad \overline{Q} = \int dx x \overline{q}(x) \quad (17.125c)$$

This means that as a result of Bjorken scaling, the total cross section of the neutrino increases in proportion to the incident neutrino energy

$$\sigma_{\text{tot}}^{\nu, \overline{\nu}} \propto s = 2M\omega \quad (17.126)$$

The compilation of neutrino total cross section given in Fig. 17.17 confirms the above conclusion well.

If the rule $\sigma_{\text{tot}} \propto \omega$ breaks down, it may be due to either 1) the effect of the W mass or a higher order effect of the weak interaction or 2) the parton model is bad (needs QCD correction). To date, the slope of the neutrino cross section for $\omega < 350$ GeV is constant and no sign of breakdown is seen.

Antiquarks in the Nucleon If we set $\beta = \overline{Q}/(Q + \overline{Q})$, which is the fraction of the antiquark's momentum, β can be calculated from the total cross-section ratio:

$$R \equiv \frac{\sigma_{\text{tot}}^{\overline{\nu}}}{\sigma_{\text{tot}}^\nu} = \frac{Q + 3\overline{Q}}{3Q + \overline{Q}} = \frac{\sigma_{\text{tot}}^{\overline{\nu}}/E}{\sigma_{\text{tot}}^\nu/E} = 0.5 \quad (17.127)$$

which gives $\beta = 0.175$.⁷⁾

$d\sigma/dy$ Integrating Eq. (17.121) over x , we obtain the y distribution

$$\frac{d\sigma^\nu}{dy} = \sigma_0[Q + \overline{Q}(1-y)^2] \quad (17.128a)$$

$$\frac{d\sigma^{\overline{\nu}}}{dy} = \sigma_0[Q(1-y)^2 + \overline{Q}] \quad (17.128b)$$

7) Rigorously, the distribution functions are not functions of x only but contain a Q^2 -dependent correction (QCD effect), which makes β energy dependent. We ignore its effect here.

In the limit $Q \gg \bar{Q}$,

$$\frac{d\sigma^{\nu}}{d\gamma} \sim 1, \quad \frac{d\sigma^{\bar{\nu}}}{d\gamma} \sim (1 - \gamma)^2 \quad (17.129)$$

Remembering that the cross sections of ν , $\bar{\nu}$ with a point particle are given by Eq. (17.115) and has the shapes $d\sigma/d\gamma(\nu q) \sim 1$, $s\sigma/d\gamma(\bar{\nu}q) \sim (1 - \gamma)^2$, we see that the cross sections of ν with nucleons directly reflects those of point particles. We present actual data [3] in Fig. 17.18. The assumption of no antiquarks in the nucleon Eq. (17.129) is almost valid, but the effect of a small mixture of antiquarks is seen as a deviation at $\gamma \approx 1$.

Electric Charge of the Quark According to the quark/parton model, the structure functions represent momentum distributions in the nucleon. They are given by Eqs. (17.120) and (17.123) and represent the same function apart from their coefficients. Thier ratio should be 5/18 if the quark model is right. Figure 17.18b shows a comparison of neutrino data with $e-p$ and $\mu-p$ data [3]. The agreement with predictions is excellent and supports the idea of the parton being the quark.

Later observations on various processes, including the $R = \sigma(e^-e^+ \rightarrow \text{all hadrons})/\sigma(e^-e^+ \rightarrow \mu^-\mu^+)$ annihilation process and $\pi^0 \rightarrow 2\gamma$, show good agreement with the quark model (Sect. 14.3). It was also shown that no parton models with integer charge can explain all the data, but the quark model can give

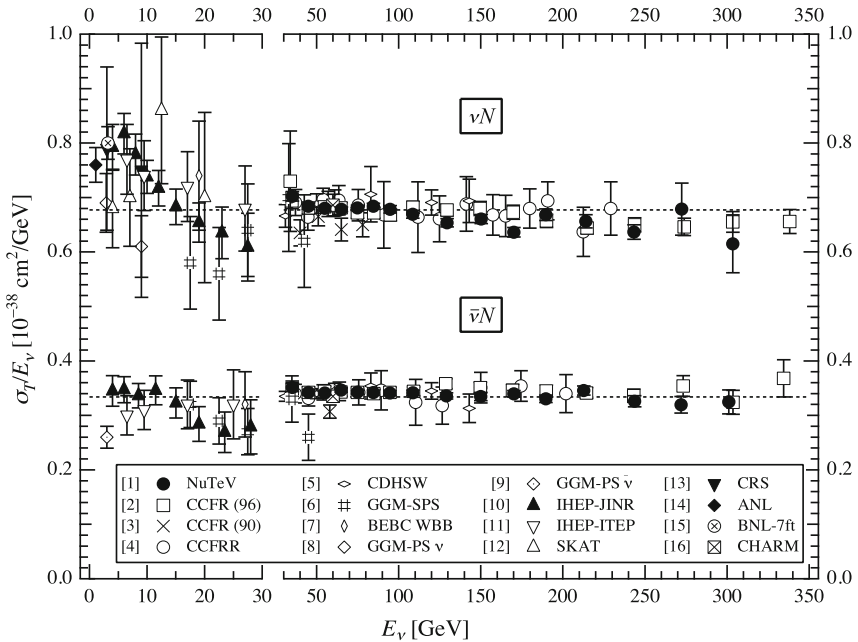


Figure 17.17 Slope of the total cross section of the neutrino $\sigma_{\text{tot}}/E_\nu$ [311].

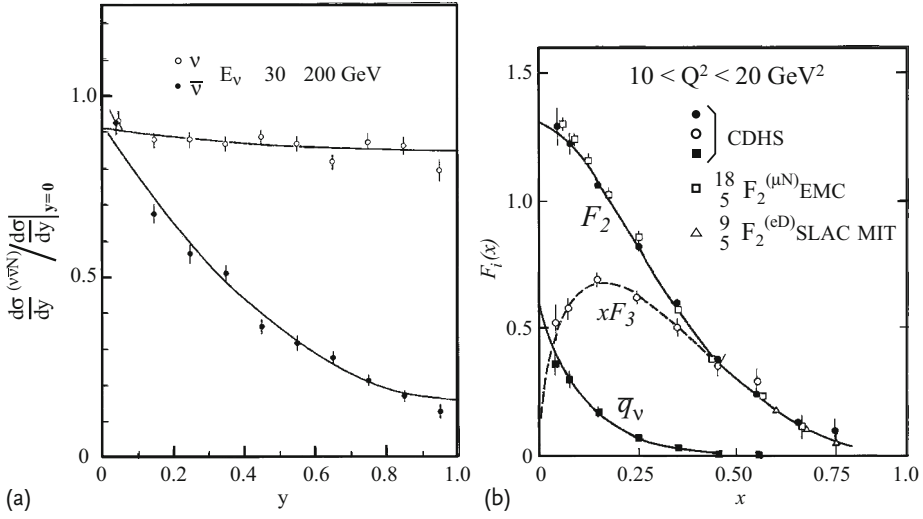


Figure 17.18 (a) Except for a small shift due to a small mixture of antiquarks, the parton model formulae $d\sigma^\nu/dy \sim 1$, $d\sigma^{\bar{\nu}}/dy \sim (1-y)^2$ agree well with the data. (b) Com-

parison of structure functions due to ν , e , μ . $F_2(x) = q(x) + \bar{q}(x)$, $xF_3(x) = q(x) - \bar{q}(x)$ and $\bar{q}(x)$ are plotted. $\bullet$, $\circ$, $\blacksquare$ are ν data, $\square = \mu p$, $\triangle = eD$ data [3].

satisfactory explanations for all of them [246]. Thus it was proved without doubt that the parton is nothing but the quark itself.

17.8.5

Sum Rules

Expressing the quantum numbers of the nucleon in terms of quarks, we have the following sum rules. Here we include s quarks.

$$\text{Baryon number: } B = 1 = \int dx \left(\frac{1}{3} \right) [u - \bar{u} + d - \bar{d} + s - \bar{s}] \quad (17.130a)$$

$$\text{Isospin: } I_3 = 1/2 = \int dx \left(\frac{1}{2} \right) [u - d - \bar{u} + \bar{d}] = \int dx \left(\frac{1}{2} \right) [u_v - d_v] \quad (17.130b)$$

$$\begin{aligned} \text{Electric charge: } Q = 1 &= \int dx \left[\left(\frac{2}{3} \right) (u - \bar{u}) - \left(\frac{1}{3} \right) (d - \bar{d} + s - \bar{s}) \right] \\ &= \int dx \left[\left(\frac{2}{3} \right) u_v - \left(\frac{1}{3} \right) (d_v + s - \bar{s}) \right] \end{aligned} \quad (17.130c)$$

$$\text{Strangeness: } S = 0 = - \int dx (s - \bar{s}) \quad (17.130d)$$

From Eq. (17.122) we also have

$$\int dx \frac{F_2^{ep} - F_2^{en}}{x} = \frac{1}{3} \int dx [(u_v - d_v) + 2(\bar{u} - \bar{d})] \quad (17.131)$$

Combining the $SU(3)$ symmetry assumption $\bar{u}(x) = \bar{d}(x) = \bar{s}(x) (= s(x))$ with Eq. (17.130b), we have another equality

$$U \equiv \int dx \frac{F_2^{ep} - F_2^{en}}{x} = \frac{1}{3} \quad (17.132)$$

They are not all independent but are related by Nishijima–Gell-Mann’s law.

Equation (17.130a) is referred to as the Gross–Llewellyn Smith (GLS) sum rule [188], Eq. (17.130c) as Adler’s sum rule [6] and Eq. (17.132) as Gottfried’s sum rule [182]. Verification of these sum rules is not easy, as accurate data in the region $x \sim 0$ are necessary. Experimentally, data are available for $0.003 \leq x \leq 0.8$ and $B \sim 0.9 \pm 0.2$, $U \sim 0.227 \pm 0.016$. We may say the GLS and Gottfried’s sum rules are roughly valid.

Existence of the Gluon As the total momentum of the quarks and gluons should add up to that of the proton,

$$1 = \int dx \left[\sum x q_i(x) + x g(x) \right] \quad (17.133)$$

where $g(x)$ denotes the gluon momentum distribution function. However, since the gluon does not interact with the electron or the neutrino, the total momentum P_q of quarks that appear in Eq. (17.133) is given by

$$P_q = \int dx \sum x q_i(x) = 1 - \int dx x g(x) < 1 \quad (17.134)$$

The total momentum of quarks and antiquarks is given by $Q + \bar{Q}$. From Eq. (17.125) and the experimental value of the total cross section, we have the total momentum

$$Q + \bar{Q} = \frac{3\pi}{4G_F^2 \omega M} (\sigma_{\text{tot}}^\nu + \sigma_{\text{tot}}^{\bar{\nu}}) = 0.5 \pm 0.04 \quad (17.135)$$

Since sum of momenta obtained from e – p scattering must give an identical value, there are particles which carry the rest of the momentum but do not interact weakly or electromagnetically. This was the first indication of the gluon’s existence. The gluon carries half of the nucleon momentum, which means it exists in the nucleon not instantaneously but permanently.

In summary, we have shown that the parton consists of the quark and the gluon. Direct evidence of the gluon was later presented by three-jet production in the electron–positron annihilation process $e^-e^+ \rightarrow q + \bar{q} + g$ [43, 291] (Fig. 17.19), and that it has spin 1 was also confirmed [363].

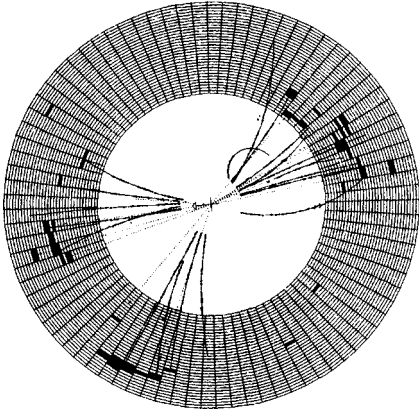


Figure 17.19 Direct production of the gluon. A three-jet event by the reaction $e^- + e^+ \rightarrow q + \bar{q} + g$ in the JADE experiment [291].

17.9

Parton Model in Hadron–Hadron Collisions

17.9.1

Drell–Yan Process

The Drell–Yan process is a way of producing a (large mass) muon pair in hadronic reactions. It is a good test bench for applying the parton model adopted in lepton deep inelastic scattering to hadronic processes. We have learned that the nucleon, a representative hadron, can be treated as a collection of partons in the infinite momentum limit. If the high-energy hadron reactions, which produce a great number of many other hadrons, can be interpreted as elastic or quasielastic parton scattering, interpretations of hadronic processes would be greatly simplified. If an antiparton in one of the hadrons annihilates with another parton in the other hadron, producing a lepton pair via electromagnetic interaction, it would provide an easily recognizable object, especially if it is hard, i.e. has a large invariant mass. Let us consider the reaction

$$p \text{ (or } \pi, \bar{p}) + p \rightarrow \mu^- + \mu^+ + X \quad (17.136)$$

where X means anything. In the parton model, the process can be considered as the one depicted in Fig. 17.20. In the parton model, a quark q having momentum $\hat{p}_1 = x_1 P_1$ pair-annihilates with an antiquark $\bar{q}$ having $\hat{p}_2 = x_2 P_2$ to become a muon pair. We denote the kinematic variables as

$$q(\hat{p}_1) + \bar{q}(\hat{p}_2) \rightarrow \mu^-(k_1) + \mu^+(k_2) \quad (17.137)$$

Writing the invariant mass and momentum of the muon pair as m, q^μ and neglecting the parton mass, we have in the center of mass (CM) frame of the colliding protons,

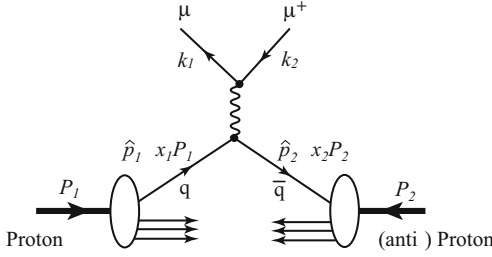


Figure 17.20 In the parton model, $p + p \rightarrow \mu^- \mu^+ + X$ is replaced by $q\bar{q} \rightarrow \gamma^* \rightarrow \mu^- \mu^+$.

$$\hat{p}_1 = x_1 P_1 \simeq x_1(E, 0, 0, P) \quad (17.138a)$$

$$\hat{p}_2 = x_2 P_2 \simeq x_2(E, 0, 0, -P) \quad (17.138b)$$

$$q = x_1 P_1 + x_2 P_2 = [(x_1 + x_2)E, 0, 0, (x_1 - x_2)P] \quad (17.138c)$$

where we have neglected the proton mass and set $E \simeq P$. From this we can define the invariant mass of the muon pair

$$m^2 = q^2 = (x_1 + x_2)^2 E^2 - (x_1 - x_2)^2 P^2 \simeq 4x_1 x_2 E^2 = x_1 x_2 s \equiv \tau s \quad (17.139)$$

Note, here, $q^2 > 0$, namely it is time-like, in contrast to $q^2 < 0$ in deep inelastic scattering. The cross section for $q\bar{q} \rightarrow \mu^- \mu^+$ can be obtained from that of $e\bar{e} \rightarrow \mu\bar{\mu}$ [Eq. (7.38)]

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} (e^- e^+ \rightarrow \mu^- \mu^+) = \frac{4\pi\alpha^2}{3s} \left[\frac{3}{16\pi} (1 + \cos^2 \theta) \right] \quad (17.140a)$$

$$\sigma_{\text{tot}}(e^- e^+ \rightarrow \mu^- \mu^+) = \frac{4\pi\alpha^2}{3s} \quad (17.140b)$$

by replacing $e^- e^+$ by $q\bar{q}$. Namely, we replace s by m^2 and correct for the quark charge λ_q and color degrees of freedom. As there are three colors R, G, B for the quarks but the muon pair is colorless, only combinations of $R\bar{R}, G\bar{G}$ and $B\bar{B}$ are allowed. Since each color exists with probability $1/3$, the color factor is $(1/3)^2 \times 3$. Setting the probability of the quarks having momentum x_1, x_2 as $q(x_1)\bar{q}(x_2)dx_1 dx_2$, we obtain

$$d^2\sigma(p p \rightarrow \mu\bar{\mu} + X) = \frac{4\pi\alpha^2}{9m^2} \sum_i \lambda_i^2 [q_i(x_1)\bar{q}_i(x_2) + q_i(x_2)\bar{q}_i(x_1)] dx_1 dx_2 \quad (17.141)$$

We introduce a directly observable variable referred to as the “Feynman x ”

$$x_F \equiv x_1 - x_2 = \pm 2|q|/\sqrt{s} \quad (17.142)$$

where q is the momentum of the muon pair in the pp CM frame. Using the Feynman x ,⁸⁾ we can rewrite

$$dx_1 dx_2 = \frac{dm^2 dx_F}{(x_1 + x_2)s} \quad (17.143a)$$

$$\frac{d^2\sigma}{dm^2 dx_F} = \frac{4\pi\alpha^2}{9m^2 s} \sum \lambda_i^2 \frac{[q_i(x_1)\bar{q}_i(x_2) + q_i(x_2)\bar{q}_i(x_1)]}{(x_1 + x_2)} \quad (17.143b)$$

$$= \frac{4\pi\alpha^2}{9\tau^2 s^2} \frac{x_1 x_2}{(x_1 + x_2)} \sum \lambda_i^2 [q_i(x_1)\bar{q}_i(x_2) + q_i(x_2)\bar{q}_i(x_1)] \quad (17.143c)$$

x_1, x_2 can be calculated from two observables τ, x_F :

$$x_1 = \frac{1}{2} \left(\sqrt{4\tau + x_F^2} + x_F \right) \quad (17.144a)$$

$$x_2 = \frac{1}{2} \left(\sqrt{4\tau + x_F^2} - x_F \right) \quad (17.144b)$$

The significance of Eq. (17.143c) lies in the fact that we can use parton distribution functions determined by deep inelastic scattering. Note, however, that the kinematical region of its applicability has to be chosen with care. The parton model is valid only in hard reactions, namely, energy and momentum transfer associated with reactions have to be large. The parton's momentum itself has to be larger than 1 GeV. From experience in $e(\nu)$ – N scattering, it would be safe if $x \gtrsim 1/E$. Assuming the conditions are satisfied, we examine Eq. (17.143b) at $x_F = 0$ ($x_1 = x_2 = \sqrt{\tau}$). It is expressed as

$$m^4 \left. \frac{d^2\sigma}{dm^2 dx_F} \right|_{x_F=0} = \frac{4\pi\alpha^2}{9} \sqrt{\tau} \sum \lambda_i^2 q_i(\sqrt{\tau}) \bar{q}_i(\sqrt{\tau}) \quad (17.145)$$

This means we can observe the product of distribution functions and check the validity of the parton model. To obtain $d\sigma/dm^2$, we may integrate Eq. (17.143b), but it is more convenient to go back to Eq. (17.141) and use the relation

$$1 = \delta(m^2 - x_1 x_2 s) dm^2 \quad (17.146)$$

After changing the order of integration we obtain

$$m^4 \frac{d\sigma}{dm^2} = \frac{4\pi\alpha^2}{9} \tau \frac{dL}{d\tau}(\tau) \quad (17.147a)$$

$$\frac{dL}{d\tau}(\tau) = \int dx_1 dx_2 \delta(x_1 x_2 - \tau) \sum \lambda_i^2 [q_i(x_1)\bar{q}_i(x_2) + q_i(x_2)\bar{q}_i(x_1)] \quad (17.147b)$$

$dL/d\tau(\tau)$ can be interpreted as the luminosity of the parton collision. According to Eq. (17.147), the lhs does not depend on energy if $\tau = m^2/s$ is kept constant, in other words, a scaling law holds for the Drell–Yan cross section. Figure 17.21a [79] shows the Drell–Yan cross sections at 200, 300 and 400 GeV. Scaling is confirmed and we have also shown the validity of the parton model in hadronic reactions.

8) The “ x ” of deep inelastic scattering is referred to as the Bjorken x .

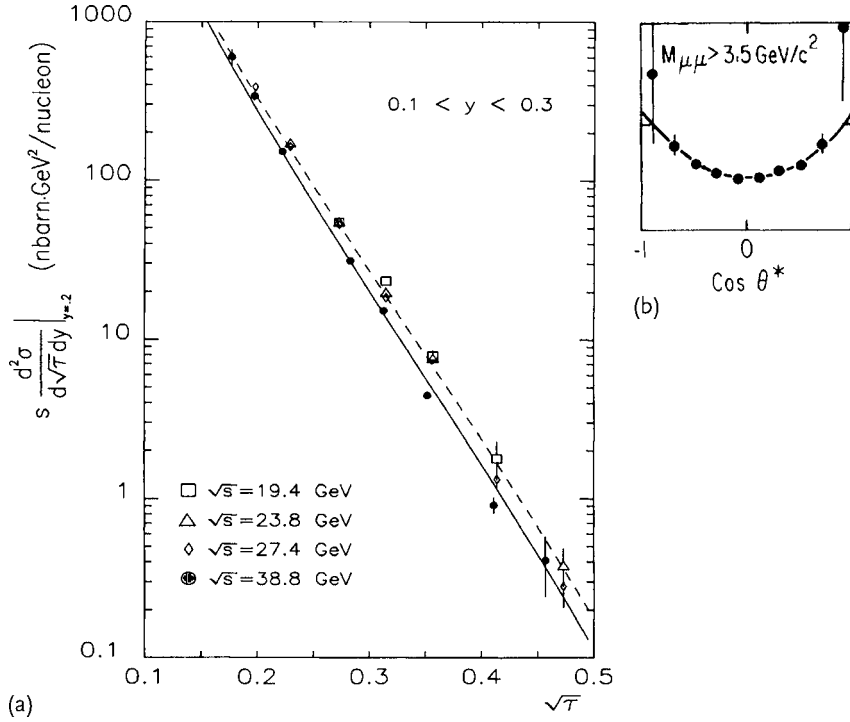


Figure 17.21 $s(d^2\sigma/d\sqrt{\tau}dy)$ vs $\sqrt{\tau}$. (a) The dashed (solid) line corresponds to order α_s QCD prediction at $\sqrt{s} = 19.4(38.8)$ GeV. Data are for $\square$ 19.4, $\triangle$ 23.8, $\diamond$ 27.4, $\bullet$ 38.8 GeV. (b) Angular distribution of muon pairs showing $1 + \cos^2 \theta$ characteristic of $q\bar{q} \rightarrow \mu\bar{\mu}$ [79, 207].

Note, however, that although momentum and angular distributions are well reproduced, the absolute cross section exhibits a small discrepancy; the difference is referred to as the K factor

$$K = \frac{(d\sigma/dm^2)_{\text{exp}}}{(d\sigma/dm^2)_{\text{theor}}} \quad (17.148)$$

Observation [79] shows that $K \sim 1.6$ in this energy region and is considered to arise from QCD higher order corrections.

17.9.2

Other Hadronic Processes

The formalism we developed for the Drell–Yan process can easily be extended to other hadronic reactions. Let us consider a process where hadron A and hadron B collide and emit a parton or a lepton, which we denote generically as c :

$$A + B \rightarrow c + X \quad (17.149)$$

The reaction cross section of the process can be expressed in parton language as that of parton a in hadron A and b in $B [= \hat{\sigma}(ab \rightarrow cX)]$ multiplied by parton distribution functions. If we use $f_A^a(x_a)dx_a$ to denote the probability of parton a being found in hadron A with momentum x_a , the hadronic cross section is expressed as

$$\sigma(AB \rightarrow cX) = \int dx_a dx_b \hat{\sigma}(ab \rightarrow cX) \times [f_A^a(x_a)f_B^b(x_b) + (A \leftrightarrow B \text{ if } a \neq b)] \quad (17.150)$$

Here $\hat{\sigma}$ denotes the elementary process by partons and is to be understood as the average of the color degrees of freedom for the initial state and the sum over the final states. Writing the total energy of the partons a and b as $\hat{s}$, we have

$$\hat{s} = x_a x_b s = \tau s \quad (17.151)$$

The formula corresponding to Eqs. (17.147) can be written as

$$\frac{d\sigma}{d\tau}(AB \rightarrow cX) = \frac{dL_{ab}}{d\tau}(\tau) \hat{\sigma}(ab \rightarrow cX; \hat{s} = \tau s) \quad (17.152a)$$

$$\frac{dL_{ab}}{d\tau}(\tau) = \int dx_a dx_b \delta(x_a x_b - \tau) [f_A^a(x_a)f_B^b(x_b) + (A \leftrightarrow B \text{ if } a \neq b)] \quad (17.152b)$$

As for observables, rapidity η is commonly used in addition to Feynman x , where

$$E_{ab} = x_a E + x_b E = (x_a + x_b)E \quad (17.153a)$$

$$P_{ab} = x_a P + x_b(-P) \simeq (x_a - x_b)E \quad (17.153b)$$

$$\eta = \frac{1}{2} \ln \frac{(E_{ab} + P_{ab\parallel})}{(E_{ab} - P_{ab\parallel})} = \frac{1}{2} \ln \frac{x_a}{x_b} \quad (17.153c)$$

x_a, x_b can be obtained using Eq. (17.144) if τ and x_F are known, but when τ and η are known, they can be obtained from

$$x_a = \sqrt{\tau} e^{+\eta}, \quad x_b = \sqrt{\tau} e^{-\eta} \quad (17.154)$$

In terms of rapidity, the cross section is expressed as

$$\frac{d^2\sigma}{d\eta d\tau} = \frac{d^2\sigma}{dx_a dx_b} = [f_A^a(x_a)f_B^b(x_b) + (A \leftrightarrow B \text{ if } a \neq b)] \hat{\sigma} \quad (17.155)$$

These formulae are used for expressing cross sections for production of W, Z and jets in hadronic reactions.

17.10

A Glimpse of QCD's Power

We have learned that partons are quarks and gluons and that gluons carry a fair amount of their parents' momentum. The dynamics of quarks and gluons have

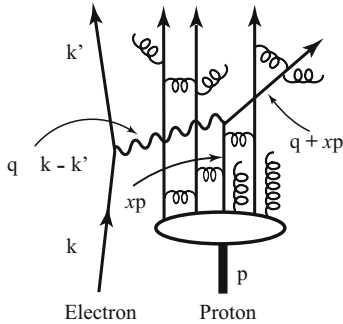


Figure 17.22 Deep inelastic scattering in QCD. Partons are made of quarks and gluons. Quarks and gluons are constantly interacting with each other. For a complete QCD treatment, the contribution of the gluon has to be included.

been investigated successfully using QCD, a color $SU(3)$ gauge theory. In its framework, the deep inelastic scattering process can be thought of as the electron interacting with one of the quarks (Fig. 17.22). In the parton model, partons are treated as mutually independent and the whole reaction can be thought of as a sum of independent e - q elastic scatterings. In terms of the equivalent photon scattering, the virtual photon carrying momentum q and mass $Q^2 = -q^2$ is absorbed by a quark. In QCD, the quarks are accompanied by gluon clouds, just like the electron or the proton is surrounded by photon or pion clouds. When the probe used to look into the inner structure of the target is soft (namely when the momentum transfer it carries is small), it does not have much resolving power. The virtual photon is simply absorbed by a parton carrying fractional momentum x of the target proton

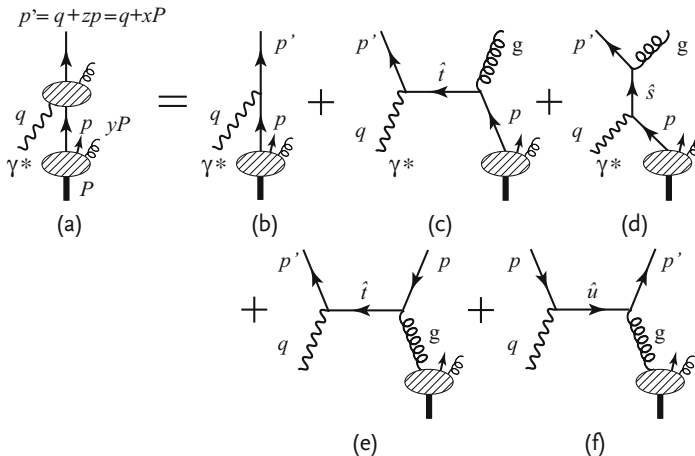


Figure 17.23 Deep inelastic scattering process in QCD. (a) As the probe becomes harder, it begins to see a structure in the quark. The quark is split from the gluon cloud. (b)

The parton model is equal to the Born term. (c,d) The next higher order corrections are Compton-like processes.

changing its momentum from $p = xP$ to $p' = q + xP$. In QCD the process can be considered as the Born term depicted in Fig. 17.23b. But as the probe becomes harder and penetrates deep into the object, it begins to resolve the core from the cloud.

When the photon sees a magnified quark, the probability of finding a quark in the quark–gluon cloud composite is not 100%. In the infinite momentum frame, we may consider that the quark in the cloud carries a fraction z of the total momentum, and the virtual photon interacts with it. The process can be expressed pictorially as in Fig. 17.23a. The proton carrying momentum P supplies a parton with fractional momentum $p = \gamma P$. Subsequently the photon sees a magnified quark interacting with the quark in the cloud and picks up fraction z of $p = \gamma P$. The parton with $p = \gamma P$ comes out as one carrying $q + zp = q + xP$, $x = z\gamma$.

Just as the parton model treated the proton as a collection of partons, the photon with larger Q^2 sees a parton split into a collection of partons. The parton model is then the Born term (Fig. 17.23b) of the whole reaction (a), and one can think of the next-order QCD process as depicted in Fig. 17.23c–f. In Fig. 17.23c,d, the photon

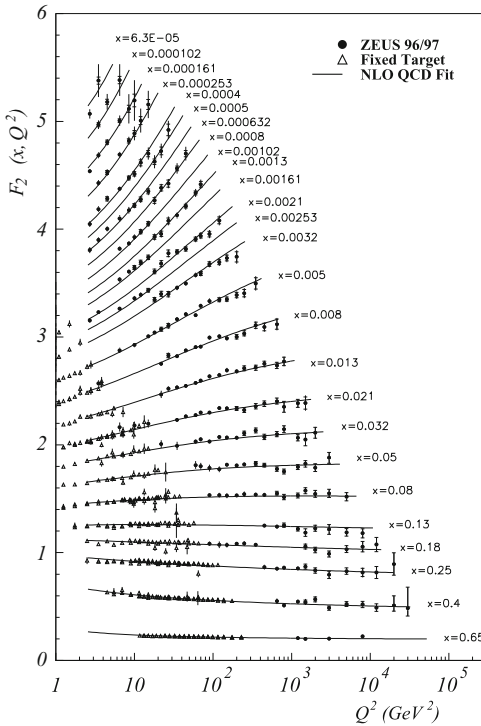


Figure 17.24 $F_2(x, Q^2)$ plotted as a function of Q^2 for different values of x . The data are a compilation of those obtained from e – P and μ – P scattering and include those of

the ZEUS [399], NMC [295], BCDMS [52] and E665 [4] groups. The lines are NLO (next to leading order) QCD fitting.

sees a quark split into a quark and a gluon and in Fig. 17.23e,f, a gluon split into a $q\bar{q}$ pair.

The whole process is formulated as the DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) evolution equation [17, 18, 123, 187]. We do not go into the details of its calculation here, but let us catch a glimpse of its power. Figure 17.24 shows an example of the DGLAP formula fitted to the structure function. The data clearly show variation of the cross section as a function of Q^2 i.e. the breakdown of the scaling, hence the naive parton model. The higher order QCD effect is at work, as shown by the various slopes of the distribution of the data at fixed x , and can reproduce the data correctly for a vast range of kinematic variables. The qualitative features of the data can be understood as follows. As Q^2 grows, the virtual photon sees more and more gluons, which in turn create $q\bar{q}$ pairs. This means the number of quarks in the large x region decreases while that in the small x region increases. This sort of behavior originates from multiple emission of gluons, which is common to all gauge field theories. But this amazing data reproducibility together with other evidence we mentioned in describing the quark model in Chap. 14 firmly established the validity of QCD as the correct theory of quarks and gluons.

18

Gauge Theories

18.1

Historical Prelude

The development of particle physics during the period beginning in the 1970s completely changed our view of matter. For the first time in history, we began to think we could answer the question about what rules control the world of matter. Some may feel it exaggerated to claim that much, but the success of gauge theory and the resultant changes of our view on matter were so revolutionary as to justify these words. It is an epoch-making phenomenon comparable only to the establishing of relativity and quantum mechanics. The gauge principle is now considered a fundamental cosmic principle in conjunction with relativity and quantum mechanics. Yet the gauge principle thus established was by no means a new idea. It was well known that the Maxwell equations satisfy it, and it was in 1918, immediately after the proposal of general relativity, that Weyl published the first gauge theory [387, 388]. Why did it take such a long time to recognize it as one of the most fundamental principles governing nature? The reader may refer to [304] for a history of how gauge theory developed.

First of all, a long time was spent establishing self-consistent quantum field theories. It was in the late 1940s that QED, the first quantum field theory based on the gauge principle, was established. Divergence, an inherent difficulty of quantum field theory, was treated by renormalization. Gauge theory could have made more progress triggered by this opportunity, but there were several factors hindering it. First, renormalization formalism depends on a recipe in which a finite term is extracted by subtracting infinity from infinity, and this was not considered a refined theory mathematically. Second, field theoretical treatment of strong interaction phenomena faced the difficulty that the coupling constant was large (~ 10) and perturbation approaches did not work well. Other methods were no better. Above all, in the 1960s the idea of a particle democracy, in particular the bootstrap model, according to which all the particles are made of other particles, gained considerable support. The notion of fundamental particles and confidence in field theory itself were weak.

Furthermore, the weak and strong interaction phenomena did not reveal the characteristics that any gauge theory should respect. First, the force has to be long

range, like the Coulomb force, but both were short range. Today, we know this characteristic is hidden for different reasons, the weak force due to spontaneous symmetry breaking and the strong force due to confinement. We were finally able to understand their true nature as a result of our perception that quarks and leptons are the fundamental constituents of matter. Based on the quark model, it was possible to sort out and organize essential features from a proliferation of hadronic and weak interaction phenomena, thus enabling us to see through this framework how the fundamental forces operate.

The second characteristic of gauge theory is its universality. The strength of gravity/electromagnetism depends only on the mass/electric charge and not on matter species or its form. General relativity is also a sort of gauge theory, but because of its mathematical complexity it took a long time to fully appreciate this [202, 230, 376].

The third characteristic is that the source of the force in gauge theory must be a conserved quantity. Energy momentum/electric charge, the sources of the gravitational/electromagnetic force, were known to be strictly conserved quantities. On the other hand, there were no such conserved quantities in strong and weak interaction. At least they were invisible to our eyes. The reason is the same as for their short-range character.

The fourth characteristic is that in strong and weak interactions the action of the force changes the identity of the particle (exchange force). For instance, the proton and neutron interchange with each other under the action of the nuclear force and the neutrino is transformed to the charged lepton. It was only after Yang and Mills's mathematical formulation [396] that extended the applicability of gauge theory to exchange forces that they were reconsidered in its framework.

Despite its great success, the gauge principle is not an easy concept and is presented in a formal way in most textbooks. Here, we try to explain it more physically, using known concepts and analogies, for an intuitive understanding but with some loss of mathematical rigor. The explanation will take a somewhat roundabout route, so let us list the logical steps we follow from now on.

1. First we define the gauge principle, mathematically and formally.
2. Then we try to apply a geometrical interpretation to the mathematical formula and attach physical meanings, making correspondence with gravity (general relativity).
3. We introduce an experiment to demonstrate the Aharonov–Bohm effect as proof that the gauge potential is an observable in quantum mechanics and plays a more fundamental role than the electromagnetic field strength.
4. We extend gauge symmetry to nonabelian groups using $SU(2)$ (isospin) symmetry. We show, applying the geometrical interpretation to gauge theory, that extension to nonabelian groups is straightforward and their formulae are easily derived in a natural and logical way.
5. QCD is a direct application of $SU(3)$ gauge theory to quark colors. We derive the asymptotic freedom qualitatively, whose discovery elevated QCD to the theory of the strong interaction historically. We also discuss confinement in a qualitative way.

6. Establishing the weak interaction theory is more complicated in two ways. First, it mixes with the electromagnetic interaction and the symmetry has to be slightly modified to $SU(2)_L \times U(1)_Y$. Glashow's scheme to combine the two forces is discussed first.
7. Second, the weak force carrier, the $W^\pm$ boson, has a large mass. If the weak force respects the gauge principle, the symmetry is broken and as a result the theory is unrenormalizable. To rescue the situation, the concept of spontaneous symmetry breaking is introduced. Within its framework, a symmetry-breaking phenomenon can be realized without breaking the symmetry mathematically. Then the Higgs mechanism to attach finite mass to particles is discussed.
8. Combining (6) and (7) the GWS (Glashow–Weinberg–Salam) model, also known as electroweak theory, is constructed.

QCD and the GWS model are collectively dubbed gauge theories based on $SU(3)_{\text{color}} \times SU(2)_L \times U(1)_Y$ and referred to as the “Standard Model” of particle physics.

18.2

Gauge Principle

18.2.1

Formal Definition

We review QED as the model of the gauge theory to start from. From the Lagrangian's invariance under global (or first-kind) gauge transformation for a fermion field ψ

$$\psi(x) \rightarrow \psi'(x) = e^{-i\alpha} \psi(x), \quad \bar{\psi}(x) \rightarrow \bar{\psi}'(x) = \bar{\psi}(x) e^{i\alpha} \quad (18.1)$$

Noether's theorem guarantees the existence of a conserved current

$$j_{\text{EM}}^\mu = q \bar{\psi} \gamma^\mu \psi, \quad \partial_\mu j_{\text{EM}}^\mu = 0 \quad (18.2)$$

where q is the signed electric charge in units of proton charge. A conserved charge operator Q is defined as the integral of the time component of the conserved current, which is a generator of the gauge transformation (see arguments in Chap. 5)

$$Q = q \int d^3x \bar{\psi} \gamma^0 \psi, \quad e^{iQ\alpha} \psi e^{-iQ\alpha} = e^{-iq\alpha} \psi \quad (18.3)$$

where the factor q is extracted from the phase considering the different charge that each fermionic field carries. The gauge field (photon) is said to couple to the conserved current when the interaction Lagrangian is expressed as

$$\mathcal{L}_{\text{EM,INT}} = -q \bar{\psi} \gamma^\mu \psi A_\mu \quad (18.4)$$

The photon field satisfies the Maxwell equations

$$\partial_\mu F^{\mu\nu} = q j^\nu \quad (18.5a)$$

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \quad (18.5b)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (18.5c)$$

which are invariant under the gauge transformation of the photon field

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + q^{-1} \partial_\mu \alpha(x) \quad (18.6)$$

Here α is an arbitrary well-behaved continuous function of x , not related to α in Eq. (18.1) at this stage. The field ψ satisfies the free Dirac equation, but if it is gauge transformed with $\alpha(x)$, which depends on x (local or second kind), according to

$$\psi(x) \rightarrow \psi'(x) = e^{-i\alpha(x)} \psi(x), \quad \bar{\psi}(x) \rightarrow \bar{\psi}'(x) = \bar{\psi}(x) e^{i\alpha(x)} \quad (18.7)$$

the transformed field ψ' no longer satisfies the Dirac equation because the differential operator brings $\partial_\mu \alpha(x)$ into the equation of motion. However, a Dirac field coupled with the gauge field has Lagrangian

$$\mathcal{L}_{\text{EM}} = \bar{\psi} [i \gamma^\mu (\partial_\mu + i q A_\mu) - m] \psi \quad (18.8)$$

which is invariant under simultaneous local gauge transformations of both ψ and A_μ if one uses the same $\alpha(x)$. In other words, if the force field is inserted in the Lagrangian in such a form that the differential operator ∂_μ is replaced by the covariant differential operator D_μ

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i q A_\mu \quad (18.9)$$

the equation of motion is invariant under the local gauge transformations provided both the fermion and the force fields are transformed simultaneously. This fact has been known for quite a while under the name “minimum interaction”. Gauge theory has promoted it to a principle.

Gauge Principle

“The equation of motion is invariant under a local gauge transformation.”

It may also be stated as follows. When the Lagrangian we deal with is invariant under local gauge transformations, a conserved object, which may be called electric charge (or more generally hypercharge), is the source of the force field. The interaction Lagrangian that couples the matter (fermion) field with the force field is obtained by replacing the differential operator in it by the covariant differential. The force field introduced this way is called the gauge field. If we make an interacting system from a free field Lagrangian, which respects a global gauge symmetry, by replacing the derivative by the covariant derivative, we say the system is gauged.

Conserved quantities can be obtained by symmetry transformations, but not all of them are sources of a force. Nowadays, all the fundamental forces (strong, electromagnetic, weak and gravity) are considered as based on the gauge principle, and only those conserved objects based on the gauge principle are strictly conserved. Those not based on them (including lepton and baryon conservation) are considered ineligible for the name. In fact, in GUTs (grand unified theories), which deal with very high energy phenomenon that could only be realized at the time of big bang, they are broken.

In nonabelian gauge symmetry the phase transformation is extended to multi-component fields in matrix representations. Examples are those of $SU(2)$ (isospin) and $SU(3)$ (unitary spin) we met in Chap. 14.

18.2.2

Gravity as a Geometry

According to Einstein's theory of general relativity, a particle always moves on a straight line (geodesic line to be exact) in space-time. According to it, space-time is not flat as we perceive intuitively but curved. How it is curved is determined by the presence of matter (energy-momentum tensor). As it is hard to imagine the deformation of space-time in the four dimensions in which we live, we consider a two-dimensional space. Then, figuratively speaking, the space behaves like a stretched rubber sheet. When it is empty without matter, it is flat (Euclidean space) and Cartesian coordinates can be set up (Fig. 18.1a). A mass point or particle A on it moves on a straight-line trajectory. Its motion of equation is Newton's first law (law of inertia), or its extension to special relativity satisfying Lorentz invariance $d\mathbf{p}/d\tau = m d\mathbf{v}/d\tau = 0$, where τ is proper time. When there is a mass B, it deforms the space as shown in Fig. 18.1b. The space is no longer flat, but particle A's motion still follows a geodesic line (i.e. a straight line in the curved space). However, projected on Cartesian coordinates it looks like a curved line. Then it is interpreted as particle A's motion being affected by B's force bending its trajectory.

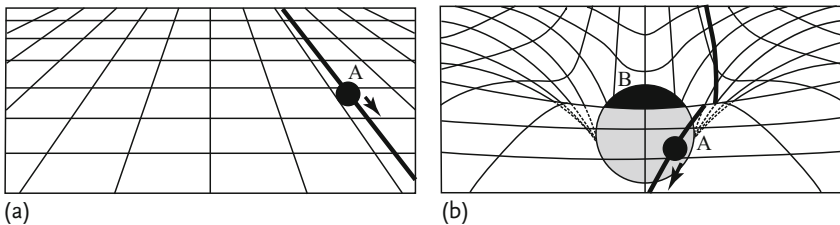


Figure 18.1 Geometrical interpretation of gravity. (a) Space-time is like a stretched rubber sheet. If there is no matter, it is flat and Cartesian coordinates can be set up. A point mass particle A moves on a straight trajectory obeying the law of inertia. (b) When there is a big mass B, the space is deformed (curved)

and no Cartesian coordinates can be set up. A's trajectory is still a geodesic line, namely a straight line in the curved space, but projected on Cartesian coordinates it is curved and is interpreted as dynamical bending of the trajectory of A attracted by B's gravity.

ry. Gravity as a dynamical phenomenon between two massive objects, in Newton's concept, has been replaced by a geometry of space-time, in general relativity.

To describe the equation of motion, we need 1) to know how space-time is deformed by the presence of matter and 2) to define geodesic lines in a curved space.

We will see that the gauge force can also be interpreted geometrically. We consider components of complex matter fields as vector components in a coordinate frame in an internal space and gauge transformation as a rotation (or more generally coordinate transformation) in the internal space. Then the force can be interpreted as the effect of a curved hyperspace, the space-time and the internal space combined. The source of the gauge force is the (electric) charge, whose presence deforms hyperspace just as a mass deforms space-time. We interpret the motion of the particle as following a geodesic in hyperspace, which is curved when projected on space-time.

18.2.3

Parallel Transport and Connection

In order to construct mathematical formulae for geodesic lines in a curved space, we need to understand the notion of parallel transport of a vector. Parallel transport is the action of moving a vector from one place to another without changing its magnitude or direction. Then a geodesic line is defined as a trajectory that can be obtained by parallel transport of its tangential vector. In space-time, where its metric (equivalent of length in Euclidean space)¹⁾ is defined, it can be shown that the geodesic line is also the shortest path connecting the two points. Therefore, it is an extension of the concept of a straight line in Euclidean space. Let us start by defining a straight line locally. If a line is represented by its direction (tangential vector), the definition of a straight line is that when its tangential vector is parallel transported by an infinitesimal distance from $\mathbf{x} = (x, y)$ to $\mathbf{x} + d\mathbf{x} = (x + dx, y + dy)$, it does not change its direction or, equivalently, components of the vector do not change. Translated to a particle's motion, it can be described by Newton's first law, that is, the particle does not change its direction when it has moved an infinitesimal distance in time dt . Thus a straight line is defined by the statement that the time derivative of the particle's spatial coordinates (velocity) stays constant (Fig. 18.2a):

$$\text{Straight line in Cartesian coordinates: } \frac{d\mathbf{v}}{dt} = 0 \quad \text{or} \quad \frac{d\mathbf{p}}{dt} = 0 \quad (18.11)$$

However, if we use curvilinear coordinates, the mathematical expression Eq. (18.11) has to be modified. In order to see this, let us view the same line in polar coordinates. This is a frame where two base vectors ($\mathbf{e}_r, \mathbf{e}_\theta$) to define its local coordinates are directed in the r (radial) and θ (orthogonal to radial) directions

- 1) In Minkowski space, the metric ds^2 is defined as

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = (cdt)^2 - dx^2 - dy^2 - dz^2 \quad (18.11)$$

$\eta_{\mu\nu}$ is referred to as the metric tensor. In general relativity, the metric tensor is a function of space-time point x and denoted as $g_{\mu\nu}(x)$.

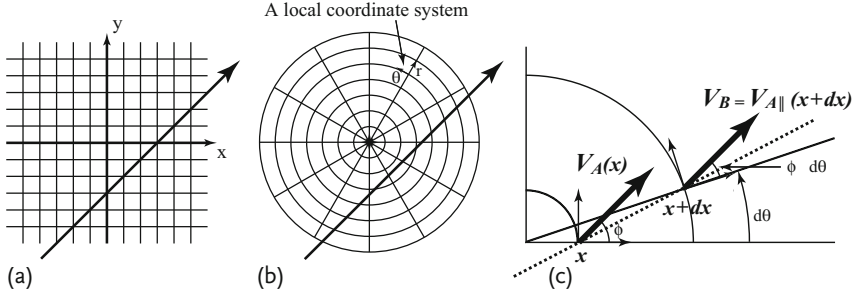


Figure 18.2 (a) A tangential vector pointing in the direction of a straight line in Cartesian coordinates has the same components everywhere. (b) In polar coordinates, while the direction of the vector stays unchanged, its direction relative to the local coordinate axis

is different everywhere. Hence the vector's components are functions of the space point. (c) A vector $V_A(x)$ at point $A = x$ is parallel transported to a point $B = x + dx$ without changing its magnitude and direction. It is denoted as $V_{A||}(x + dx)$.

(Fig. 18.2b). The directions of the coordinate frame set up at space point $x \equiv (r, \theta)$ point away from the origin (e_r) or are rotated by $\pi/2$ (e_θ), and are inclined relative to the Cartesian setting. To define a straight line locally, the change of base vectors has to be compensated.

Let us find an expression for vector parallel transport. A vector $V_A(x)$ shown in Fig. 18.2c at point $x \equiv (r, \theta)$ has components $V_A = (V_r, V_\theta) = (V \cos \phi, V \sin \phi)$ in polar coordinates as well as in Cartesian coordinates. Note that we have deliberately set the point A at $\theta = 0$ to make their components identical. When it is parallel transported to point B , $x + dx = (r + dr, \theta + d\theta)$, we denote it as $V_B = V_{A||}(x + dx)$. Here, the base vectors of the polar coordinate at $x + dx$ is inclined by an angle $d\theta$ relative to those at x . Components of $V_{A||}$ are the same everywhere as those of $V_A(x)$ in Cartesian coordinates, but in the polar coordinate frame correction terms to compensate the directional change of the base vectors have to be included:

$$\begin{aligned} V_{A||r}(x + dx) &\equiv V_{Br} = V \cos(\phi - d\theta) \simeq V \cos \phi + d\theta V \sin \phi \\ &= V_r + d\theta V_\theta \end{aligned} \quad (18.12a)$$

$$\begin{aligned} V_{A||\theta}(x + dx) &\equiv V_{B\theta} = V \sin(\phi - d\theta) \simeq V \sin \phi - d\theta V \cos \phi \\ &= V_\theta - d\theta V_r \end{aligned} \quad (18.12b)$$

This example clearly shows that vector components in the polar coordinates change their values when parallel transported. In differential geometry, the transported vector is generically expressed as

$$V_{||}(x + dx) = V(x) - \mathbf{F}_\nu(x) V(x) dx^\nu \quad (18.13)$$

The correction term is assumed to be a linear transformation of the original vector and proportional to the displacement if it is small. The assumption is justified in our example Eq. (18.12). $\mathbf{F}_\nu$ introduced in Eq. (18.13) is referred to as the “connection”. Note that $\mathbf{F}_\nu$ is a space-time vector, as indicated by its subscript ν , but it

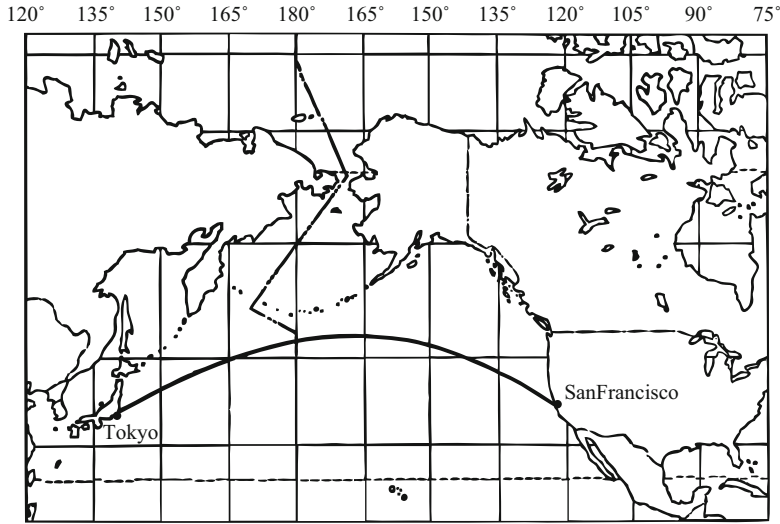


Figure 18.3 A great circle connecting Tokyo and San Francisco is a geodesic on the earth's surface (sphere). It looks curved when projected onto Cartesian coordinates in a flat space. In a local orthogonal coordinate system

(longitude, latitude) set up on the earth's surface, an airplane flying along the great circle is being parallel transported. However, its direction with respect to local coordinates changes incessantly.

is also a matrix acting on an object to be parallel transported. Since in Eq. (18.13) $\mathbf{V}$ itself is a space vector, $\mathbf{\Gamma}_\nu$ here is a n -component vector whose components are $n \times n$ matrix where n is the dimension of the space-time.

In differential geometry (and in general relativity), where a metric is given the connection is a derived quantity in terms of the derivatives of the metric tensor $g_{\mu\nu}$.²⁾ In Cartesian coordinates, all the coordinates are globally aligned in the same direction and components of the parallel transported vector are the same everywhere, namely, the connection identically vanishes everywhere. The polar coordinates we described above are set up on a flat plane, hence they can always be transformed back to the original Cartesian coordinates. In other words, there always exists a transformation that makes the connection vanish identically in a flat space by a suitable transformation. Since Cartesian coordinates can only be defined in a straight and flat space, the existence of Cartesian coordinates is equivalent to the space being Euclidean.

2) It is defined as

$$\Gamma_\mu \equiv [\Gamma_\mu]_\nu^\rho = \frac{1}{2} g^{\rho\lambda} (\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\lambda\mu} - \partial_\lambda g_{\mu\nu}) \quad (18.14)$$

and is referred to as the Christoffel connection. $g_{\mu\nu}$ is a metric tensor defining the metric of the space (see Footnote 1).

Note: In differential geometry, two distinct vectors, contravariant and covariant vectors, are defined. The Christoffel connection is used to relate their components. Vectors in our example used in Eq. (18.12) are related to but not identical to one of them because in differential geometry the base vectors change their magnitude as well as their direction.

Next, we consider the earth's surface, where the coordinates are expressed in terms of two angles (longitude, latitude). They are orthogonal coordinates everywhere and locally (i.e., in a limited region) are almost identical to Cartesian coordinates (with suitable rescaling), but globally they are not. A geodesic line (great circle) connecting Tokyo and San Francisco is depicted in Fig. 18.3. The line's direction changes continuously relative to local coordinates. Here, also, parallel transport can be defined using the connection. But this time, no Cartesian coordinates can be set up on the sphere. Therefore there is no transformation to convert the (longitude, latitude) system to Cartesian coordinates or, equivalently, there are no transformations to make the connection vanish everywhere. Obviously, the connection contains information on the space itself as well as the characteristics of the coordinate system. This is expected because the connection is derived from the metric tensor which defines the structure of the spacetime.

18.2.4

Rotation in Internal Space

We learned in Chap. 9 that in order to describe physical phenomena we have to set up a coordinate system but that the phenomena themselves do not depend on which coordinates are adopted. An experiment carried out in Japan or in Germany should produce the same result, which is referred to as translational symmetry. From it, Noether's theorem derives energy-momentum conservation. Similarly, whether the experimental setup is directed north or east, it should produce the same result. This is rotational symmetry, from which angular momentum conservation is derived. Contrary to these examples, charge and isospin conservation are derived from gauge symmetry (a phase transformation invariance), which has nothing to do with real space-time. But the above logic suggests that a coordinate system can be set up in internal space, too, and the origin of the conservation law can be considered as a result of the physical phenomena being independent of how we choose the internal coordinate system.

The simplest example is a charged particle that has two degrees of freedom, corresponding to positive and negative charge. Let us denote the field components as ψ_1, ψ_2 ; the field can be considered as a two-component vector in a certain internal space. Denoting the orthonormal base vectors in this space as e_1, e_2 , we can express the field as

$$\psi = \psi_1 e_1 + \psi_2 e_2 \quad (18.15)$$

If the physical phenomena do not depend on how we set up the coordinates, the laws of physics do not change if we rotate the coordinates. Denoting transformed ψ as ψ' , by rotating the coordinate axis by an angle α we can write

$$\begin{aligned} \psi'_1 &= \psi_1 \cos \alpha + \psi_2 \sin \alpha \\ \psi'_2 &= -\psi_1 \sin \alpha + \psi_2 \cos \alpha \end{aligned} \quad (18.16)$$

Introducing the complex field by

$$\psi = (\psi_1 + i\psi_2)/\sqrt{2}, \quad \psi^\dagger = (\psi_1 - i\psi_2)/\sqrt{2} \quad (18.17)$$

we have

$$\psi' = e^{-i\alpha}\psi, \quad \psi'^\dagger = e^{i\alpha}\psi^\dagger \quad (18.18)$$

The rotation of a real field in two-dimensional space reduces to a phase transformation in a one-dimensional complex space. Charge conservation results from the symmetry of the phase transformation, which is referred to as $U(1)$ gauge symmetry.

Now we imagine a hyperspace by combining (four-dimensional) space-time and internal space and consider the field ψ a vector in the hyperspace. If the coordinates in the hyperspace are equivalents of the curvilinear coordinates, their direction changes locally. If the rotation angle in the internal space depends on space-time coordinates $\alpha = \alpha(x)$, it is a local gauge transformation. Then in analogy to Eq. (18.13) a parallel transported field in the hyperspace can be expressed as

$$\psi_{||}(x + dx) = \psi(x) - iqA_\nu(x)\psi(x)dx^\nu \quad (18.19)$$

which introduces $A_\nu(x)$ as the connection. Here x is used to denote space-time variables in Minkowski space. As we will soon see, $A_\nu(x)$ is nothing but the gauge potential. We use the terms (gauge) potential for A_μ and field strength for its derivative, i.e. $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, which is the electromagnetic field in $U(1)$ gauge symmetry. When we refer to ψ we call it the matter field to distinguish it from the gauge field. We have introduced the gauge potential not as a force field but as a “connection”, which is a geometrical object relating vectors in different places. The equation is identical to Eq. (18.13) if we regard $\mathbf{F}_\nu = iqA_\nu$ as an operator acting on a vector (ψ) in an internal space. Unlike the Christoffel connection, which acts on a space-time vector (hence has two extra space-time indices), the connection A_ν has only one subscript in space-time and is a Lorentz vector because the field ψ in Eq. (18.19) is a scalar in space-time. The extraction of the q factor is for later convenience and the reason for introducing the imaginary number i is because the rotation in the internal space is a phase transformation.

Perhaps a simple example may lead to an intuitive understanding of the introduction of i . Consider the parallel transformation depicted in Fig. 18.2c as a vector defined in an internal space but moved from x to $x + dx$ in space-time. Then we can introduce a vector in an internal complex plane

$$V(x) = V_r + iV_\theta$$

$$V_{||}(x + dx) = V_{||r} + iV_{||\theta} \stackrel{(18.12)}{=} V - i d\theta V = [1 - iA_\nu(x)dx^\nu]V \quad (18.20a)$$

$$A_\nu(x) = \frac{\partial\theta}{\partial x^\nu} = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)^3 \quad (18.20b)$$

3) One may notice that, apart from its normalization, this happens to be the gauge potential that is realized outside a solenoidal coil. See Problem 18.1.

The interpretation of the gauge potential as the connection in a hyperspace means the gauge potential connects the locally rotating coordinates in the internal space and contains in itself the properties of the hyperspace. Note, however, no metric is defined in the internal space. Unlike the Christoffel connection, which is derived from the metric, A_ν as the connection is a starting point for the discussion that follows. The assumption is only justified by overall consistency.

18.2.5

Curvature of a Space

Considering the role of the connection, we realize that it should contain information about the characteristics (or “curvedness”) of a curved space. Let us consider how we can extract such information from the connection. Parallel transport of a vector in a flat space is easy to understand intuitively, but that in a curved space takes some getting used to. Let us consider parallel transport of vectors on a sphere (in real space). How a vector moves on the sphere can be estimated by painting many parallel arrows on a flat plane and copying them on the sphere by rolling the sphere on it. Figure 18.4 illustrates two examples of vector transport.

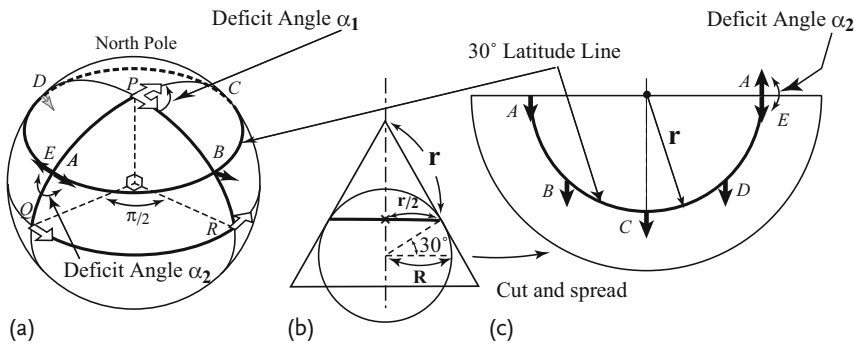


Figure 18.4 Deficit angle as a measure of curvature. (a) A vector denoted by an arrow of outline type starting from the north pole (P) is parallel transported along the lines $P \rightarrow Q \rightarrow R \rightarrow P$, where Q, R are points on the equator 90° apart. When it comes back to where it started, its direction is $\pi/2$ different, which is referred to as the deficit angle. The total curvature is defined as the deficit angle in radians and the average curvature by dividing it by the area enclosed by the loop, which is $(\pi/2)/(4\pi R^2/8) = 1/R^2$. Local curvature is obtained by making the closed loop infinitesimally small.

ABCDE is a line along latitude = 30° (points A and E are identical). It is not a great circle,

and the black vector at A, when parallel transported along the latitude 30° line, changes its direction relative to the local (longitude, latitude) coordinates as depicted. When it circles around and comes back its direction is opposite to the original direction. This can be explained by making a plane tangential to the line ABCDE (b) and spreading it out on a plane, which is shown in (c). The vector's direction is the same everywhere, but it is not the same relative to the line ABCDE. As the area of the sphere above the latitude 30° is πR^2 and the deficit angle is π , the curvature is again $1/R^2$.

A conspicuous characteristic of a curved space is that if a vector is parallel transported along a closed loop, its final direction is different from what it was initially. The difference is referred to as the “deficit angle”. As the bending of the space grows, the deficit angle becomes larger. The total curvature of an area is the deficit angle expressed in radians and the average curvature is defined as that divided by the area. Local curvature is defined as the average curvature with the closed loop taken to be infinitesimal.

As an example, let us consider a sphere with radius R . In Fig. 18.4a, a vector denoted by a white (outline typeface) arrow at the north pole is parallel transported to the equator $P \rightarrow Q$ along a longitude line, then moved along the equator to the east as far as 90° in longitude ($Q \rightarrow R$), and finally back to the north pole ($R \rightarrow P$). The vector direction is shifted by $\pi/2$ relative to the original. The area surrounded by the closed line $P \rightarrow Q \rightarrow R \rightarrow P$ is $\pi R^2/2$. Therefore, the average curvature is $(\pi/2)/(\pi R^2/2) = 1/R^2$.

For the case of parallel transport along a line at latitude 30° , one can see the vector rotates relative to the local coordinates as it moves. How it moves can be understood by spreading the tangential plane on a flat plane. As the plane tangential to the 30° latitude line is a cone, one can see that the parallel arrows on the plane rotate relative to the latitude line (Fig. 18.4c). Points A and E are identical. As is clear from the figure, the two vectors at A and E are parallel on the spread-out sheet, but in opposite directions on the sphere, hence the deficit angle is π .

In a similar manner, we can define curvature in a hyperspace by parallel transporting the field ψ along a closed loop and by calculating the deficit angle. Let us consider the case when the field is transformed $\psi \rightarrow \psi + \delta\psi$ by circulating along a loop defined as the periphery of an infinitesimal surface element $dS = dx^\mu dx^\nu$ (Fig. 18.5):

$$\begin{aligned}\delta\psi &= -iq \oint A_\mu \psi dx^\mu \\ &= -iq \int_{dS} [\nabla \times A]_{n \perp dS} \psi dx^\mu dx^\nu \simeq -iq F_{\mu\nu} \psi dx^\mu dx^\nu\end{aligned}\quad (18.21a)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (18.21b)$$

where Stokes' theorem was used in obtaining the third equality. The rotation angle $\Delta\alpha$ is given by

$$e^{-i\Delta\alpha}\psi = \psi + \delta\psi \simeq (1 - i\Delta\alpha)\psi, \quad \therefore \quad \Delta\alpha = i\delta\psi/\psi \quad (18.22)$$

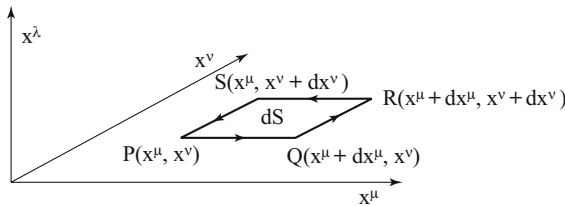


Figure 18.5 Integral path for calculating covariant derivatives $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

Therefore the curvature in the hyperspace is given by

$$\frac{\Delta\alpha}{dx^\mu dx^\nu} = \frac{i\delta\psi/\psi}{dx^\mu dx^\nu} = qF_{\mu\nu} \quad (18.23)$$

This shows that the electromagnetic field is nothing but the curvature of a hyperspace. The Maxwell equation [see Eq. (3.49)]

$$\partial_\mu F^{\mu\nu} = qj^\nu \quad (18.24)$$

can be interpreted as defining equations of curvature.

The Electromagnetic Field is the Curvature of a Hyperspace

The existence of the electric charge (and its current) deforms the hyperspace and its curvature is defined by the Maxwell equation.

Let us compare it with the equation of general relativity (Einstein's equation):

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G_N T_{\mu\nu} \quad (18.25)$$

where G_N is the gravitational constant and $R_{\mu\nu}$ and $R = (g^{\mu\nu} R_{\mu\nu})$ on the left hand side are referred to as Ricci's tensor. It is derived by contracting the Riemann tensor, the curvature of space-time.⁴⁾ This, in turn, is derived from the Christoffel connection by the same procedure as we used to derive the electromagnetic field. We remind the reader that the exact mathematical formula in differential geometry is not important, rather, tracing the logical flow and trying to understand the cause and result are. As $T_{\mu\nu}$ on the rhs is the energy-momentum tensor, Einstein's equation says that the energy distribution in space-time determines the curvature of space-time. One recognizes complete parallelism in the formalism of electromagnetic and gravity equations.

18.2.6

Covariant Derivative

A free-field equation is generally a second-rank differential equation, which is generically expressed as

$$\mathcal{A}(\partial_\mu)\psi(x) = 0 \quad (18.27)$$

- 4) Riemann's and Ricci's curvature tensors are expressed as

$$\begin{aligned} R_{\sigma\mu\nu}^\rho &= -[\partial_\mu \Gamma_\nu^\rho - \partial_\nu \Gamma_\mu^\rho + (\Gamma_\mu^\rho \Gamma_\nu^\sigma - \Gamma_\nu^\rho \Gamma_\mu^\sigma)]_\sigma^\rho \\ R_{\mu\nu} &= R_{\mu\rho\nu}^\rho, \quad R^{\mu\nu} = g^{\mu\rho} g^{\nu\sigma} R_{\rho\sigma}, \\ R &= R^\mu{}_\mu = g^{\mu\rho} R_{\rho\mu} \end{aligned} \quad (18.26)$$

The third term in the square brackets, which is nonlinear in Γ , stems from the nonabelian nature of the Christoffel connection. Notice it has exactly the same form as the nonabelian gauge field [see Eq. (18.57) and discussions in Sect. 18.4.3].

where $\mathcal{A}(a)$ is a polynomial of a up to quadratic terms. We need to define the differential operator in a general coordinate frame. When a field $\psi(x)$ is defined at all space-time points, its derivative is normally defined as

$$\frac{\partial \psi(x)}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta \psi}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\psi(x + \Delta x) - \psi(x)}{\Delta x} \quad (18.28)$$

For a vector field, the above formula applies to each component. The difference for a vector field as compared to a scalar field is that, when making difference $\Delta \psi$ between $\psi(x + \Delta x)$ and $\psi(x)$, the feet of both vectors have to be at the same space-time point. In the Cartesian coordinate system, one need not worry about this, because the direction of a vector or equivalently all components of the vector are the same anywhere in space. This is not true in curvilinear coordinates, because the same vector has different components at different space points. In this case, the vector at x has to be parallel transported to $x + \Delta x$ without changing its magnitude and direction (Fig. 18.6). We already know the expression for a vector parallel transported from x to $x + dx$. It is given in Eq. (18.13). To differentiate ψ , one has to make a difference $\Delta \psi$ by subtracting $\psi_{\parallel}(x + \Delta x)$ from $\psi(x + \Delta x)$ as shown in Fig. 18.6.

$$D_{\mu} \psi(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta \psi(x)}{\Delta x^{\mu}} = \partial_{\mu} \psi(x) + \mathbf{F}_{\mu}(x) \psi(x) \quad (18.29a)$$

which is referred as covariant derivative. Here, we illustrated ψ as a vector in the spacetime and used Eq. (18.13) to define the derivative as given in the differential geometry. Similarly, to differentiate ψ in the curved hyperspace, we use Eq. (18.19). Then its derivative can be defined as

$$\begin{aligned} \frac{D \psi(x)}{dx^{\mu}} &\equiv \lim_{\Delta x^{\mu} \rightarrow 0} \frac{\psi(x^{\mu} + \Delta x^{\mu}) - \psi_{\parallel}(x^{\mu} + \Delta x^{\mu})}{\Delta x^{\mu}} \\ &= \frac{d \psi(x)}{dx^{\mu}} + i q A(x)_{,\mu}(x) \psi(x) \end{aligned} \quad (18.29b)$$

which is precisely the “covariant derivative” we defined in Eq. (18.9). The covariant derivative contains the connection, which contains the characteristics of the coordinate system as well as the structure of the space.

According to differential geometry, in curvilinear coordinates the same equation holds anywhere in the space if one uses covariant derivatives instead of ordinary differential operators. Phrased in a different way, an equation expressed in terms of covariant derivatives keeps the same form under any general coordinate transformation provided a suitable connection is used. This is covariance under general coordinate transformation or the principle of general relativity. The principle itself is a matter of geometry and has no physical substance. The physics lies in identifying the space-time nature has chosen and the equation of motion in that space. The former is realized by Einstein’s equation and the equation of motion is obtained from the principle of equivalence. We already have the equivalent of the Einstein equation in the hyperspace. Therefore, our next task is to translate the equivalence principle to the language of the gauge theory.

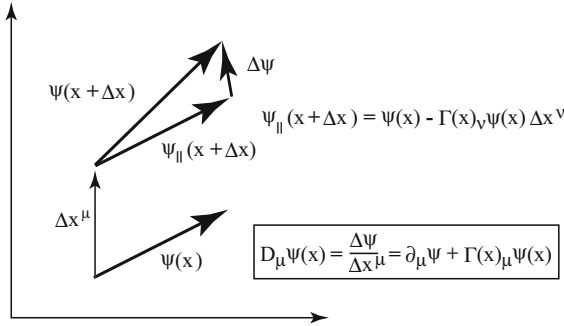


Figure 18.6 Parallel transport and covariant derivative. Parallel transport of a vector in curvilinear coordinates is defined by using the “connection Γ_μ ”. $\psi(x)$ parallel transported to

point $x + \Delta x$ is denoted as $\psi_{||}(x + \Delta x)$. A derivative in curvilinear coordinates is defined as $\lim_{\Delta x \rightarrow 0} [\psi(x + \Delta x) - \psi_{||}(x + \Delta x)] / \Delta x$.

18.2.7

Principle of Equivalence

The principle of equivalence identifies gravity as the inertial force in an accelerating system. In an inertial frame, there is no gravity and the equation of motion is described by Newton’s law ($dp/dt = 0$) or its extension to special relativity ($dp/d\tau = 0$, τ = proper time). Since the equation expressed in terms of covariant derivatives keeps the same form if transformed to an accelerating frame, the equation of motion under the gravitational field is expressed as

Principle of Equivalence

Eq. of motion in inertial frame	$\rightarrow$	Eq. of motion in curved space
$\frac{dp^\mu}{d\tau} = 0$		$\frac{Dp^\mu}{d\tau} = 0$
		(18.30)

This means that the trajectory that a point particle follows is a geodesic line. We have now presented the mathematical formula of what we described pictorially in Fig. 18.1. A particle A under the effect of another mass B follows the trajectory of a geodesic line. In order for the above equation to hold, one has to verify the existence of the inertial frame in an accelerating frame. Thinking physically, it is always possible by getting in an elevator and cutting the suspension wire (Fig. 18.7a). The elevator goes into free fall and no gravity exists inside it. If one feels bad about getting in a falling object, a satellite would provide the same inertial frame. In mathematical language, there is always a transformation that makes the connection vanish at a point. The vanishing of the connection means the existence

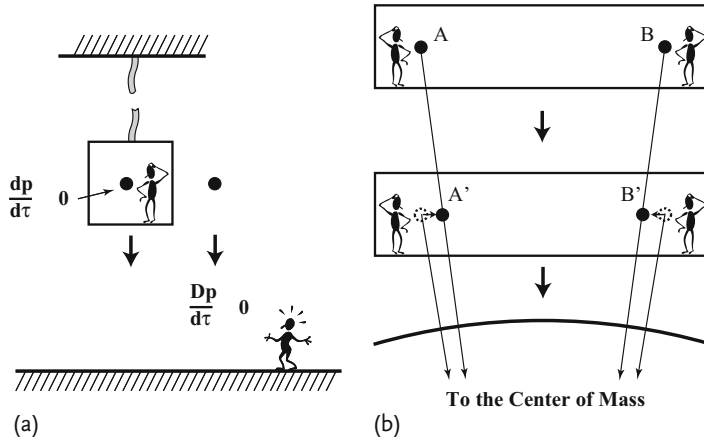


Figure 18.7 Principle of equivalence. (a) In a gravity field, space-time is curved, but there always exists a coordinate transformation where “the connection” vanishes at a point. Physically, it can be realized by entering an elevator and cutting its suspension wire. The equation of motion in the elevator is that of the inertial frame, i.e. $dp/d\tau = 0$, where τ is proper time. Outside the elevator it is obtained by replac-

ing the derivative with its covariant version and transforming back to the original accelerating frame. (b) An illustration of the fact that the connection vanishes only locally. In a wide area containing A and B attracted by the earth (a spherical gravitational center), one sees two mass points are attracted to each other without any force, which means the area is not in an inertial frame.

of Cartesian (Minkowskian in the case of space-time) coordinates, which, in turn, means an inertial frame. Notice that a local inertial frame means the vanishing of the connection at a point mathematically, but physically it includes its immediate neighborhood as long as the higher order terms in the Taylor expansion can be neglected. If one considers a wide area such as that depicted in Fig. 18.7b, one sees two mass points A and B come closer as time elapses, which means the area is not an inertial frame.

What is meant by “local”? For radius < 10 m, a circular accelerator can be built assuming the earth is flat. But when the radius of the acceleration circle is on the order of kilometers, the curvature of the earth’s surface can no longer be neglected in aligning magnets. Here, locality is limited only to ~ 100 m. Do the dynamics of galaxy formation or even the formation of a cluster of galaxies need to take into account cosmic curvature effect? Here locality applies to a region as large as millions of light years. Locality changes depending on the kind of observables we are talking about.

Going back to particle field theories, in the absence of an interaction, the matter field equation can be derived from free Lagrangians. We already know them, the Klein–Gordon, Dirac or Maxwell (or Proca if massive) equations for spin 0, 1/2 or 1 fields. They are the equivalent of the law of inertia in gravity. General relativity means covariance of the equation under general coordinate transformation. In gauge theory, it means covariance under local gauge transformation. In both cases it is achieved if the equation is written in terms of the covariant derivatives.

The principle of equivalence tells us that by replacing the derivatives in the free Lagrangian by the covariant derivatives, the Lagrangian in curved space, namely, that of interacting fields, is obtained. This is precisely the process of gauging we mentioned before.

Principle of Equivalence in Gauge Theory

$$\begin{array}{ccc} \text{Eq. of free field} & & \text{Eq. of interacting field} \\ \mathcal{A}(\partial_\mu)\psi(x) = 0 & \rightarrow & \mathcal{A}(D_\mu)\psi(x) = 0 \end{array} \quad (18.31)$$

The existence of an inertial frame in general relativity is translated to gauge theory as follows. No matter what internal coordinate [gauge = phase $\alpha(x)$] is used, there always exists a coordinate (gauge) transformation such that the connection (gauge potential) vanishes at at least one point. The requirement is expressed as

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu + \partial_\mu \alpha(x) = 0 \quad (18.32)$$

Let us see if there is a gauge transformation that makes the gauge field vanish at every space-time point. The condition that Eq. (18.32) can be solved is for α to be integrable, namely

$$\partial_\mu \partial_\nu \alpha(x) = \partial_\nu \partial_\mu \alpha(x) \quad (18.33)$$

Combining this with Eq. (18.21b) gives

$$F_{\mu\nu} = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) = 0 \quad (18.34)$$

Namely, the nonexistence of the electromagnetic field is the condition for the vanishing of the gauge potential everywhere. Since, in hyperspace, the electromagnetic field represents its curvature, and the vanishing of the gauge potential everywhere means a flat hyperspace, it is logical to require the electromagnetic field to vanish identically.

18.2.8

General Relativity and Gauge Theory

Let us rephrase the theory of general relativity in the language of gauge theories.

In an inertial frame, Lorentz (or more generally Poincaré) invariance holds. The Lorentz transformation [Eq. (3.76)] has the form of a phase transformation, and contains 6 (10) parameters, the velocity and rotation angle (and displacement). They are constants and the transformation could be dubbed a

global gauge transformation. If the transformation is made local, i.e. if the parameters are functions of space-time points, it is a general coordinate transformation. The principle of general relativity is rephrased as invariance under local Poincaré transformation. It requires that the equation of motion does not change its form. The principle of general relativity is nothing but invariance under local gauge transformation.

A big difference between gravitational theory and gauge theory is that, in gauge theory, parallel transport is activated in an internal space and the connection has only one subscript as far as space-time components are concerned. This is why the gauge field is a Lorentz vector. Hence the gauge particle has spin $s = 1$. If the

Table 18.1 Correspondence between gauge theory and general relativity.

	frame	Gauge Theory internal space	General Relativity space-time
Global Symmetry	coordinate trf.	gauge trf.	Poincaré trf.
	invariants	charge	energy
	eq. motion	KG/Dirac eq. ^a	inertial law
		$\mathcal{A}(\partial_\mu)\psi = 0$	$\frac{dp^\mu}{d\tau} = 0$ ^b
	coordinate trf.	local gauge trf.	local Poincaré trf.
			or general coordinate trf.
	connection	gauge potential	Christoffel connection
Local Symmetry (curved space)		$iqA_\mu(x)$	$\Gamma_\mu^\lambda = \Gamma^\lambda_{\nu\mu}$
	parallel trpt.	$\psi - iqA_\mu\psi dx^\mu$	$p^\lambda - \Gamma^\lambda_{\mu\nu}p^\mu dx^\nu$
	cov. derivative	$D_\mu = \partial_\mu + iqA_\mu$	$D_\mu p^\lambda = \partial_\mu p^\lambda + \Gamma^\lambda_{\mu\nu}p^\nu$
	curvature	electromag. field	Riemann's curvature
		$F_{\mu\nu}$	$R^\lambda_{\mu\nu\rho}$
	origin of curv.	charge	energy
	field equation	Maxwell eq.	Einstein eq.
		$\partial_\mu F^{\mu\nu} = qj^\nu$	$R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu} = 8\pi G_N T^{\mu\nu}$
	eq. motion	$\mathcal{A}(D_\mu)\psi = 0$	$\frac{Dp^\mu}{d\tau} = 0$
	equivalence principle	potential can vanish locally	Γ_μ can vanish locally (existence of local inertial frame)
	total curv. = 0 (flat space)	$A_\mu = 0$ globally	existence of Cartesian coordinates $\Gamma_\mu = 0$ globally

^a KG = Klein–Gordon

^b τ is proper time, $d\tau = dt\sqrt{1-\beta^2}$

hyperspace is curved, the Lorentz invariance is preserved. But in gravity, application of a general coordinate transformation breaks Lorentz invariance, which means the space-time is curved. Since the object to be parallel transported is a Lorentz vector, the (Christoffel) connection has three indices for space-time components; one as a vector in spacetime, two as a matrix acting on spacetime vector to be parallel transported. Since the connection for gravity is derived as derivatives of the metric tensor of rank 2 that gives the gravitational potential, the graviton, the quantum of the gravitational field, has spin 2. As a summary of what we have described so far, we list correspondences between gauge and gravity fields in Table 18.1.

Are All the Fundamental Forces Geometrical?

The geometrical interpretation of the gauge force we have presented is not rigorous but by no means accidental. In 1956, Utiyama proved that general relativity is a kind of gauge theory [376]. There is also Kaluza–Klein theory, which was proposed in 1921. Working in five-dimensional space-time, they showed that the electromagnetic force can be derived from general relativity as the force produced by the energy-momentum in the fifth dimension and projected in four-dimensional space-time. In their model, the fifth dimension is not infinitely long but curled to form a cylinder, making the space translation in the fifth dimension periodic, which is essentially an act of gauge transformation. The electromagnetic potential appears as the metric tensor connecting the fifth dimension with ordinary four-dimensional space-time. The electric charge is proportional to the fifth component of the momentum. Just as all three forces – weak, electromagnetic and strong – are described by local gauge theories based on $SU(3)_{\text{color}} \times SU(2)_L \times U(1)_Y$ symmetry, it is known that quantum gravity can be obtained by gauging the supersymmetry that connects fermions and bosons. This is referred to as supergravity. However, a self-consistent method of quantizing the gravitational field is not yet known. One idea that is considered promising is to regard elementary particles not as points but as strings. Superstring theory tries to unify all four forces by applying the gauge principle to supersymmetric strings in ten-dimensional space-time. It is also a generalized Kaluza–Klein theory, extending general relativity to larger dimensions. Whether the extra six-dimensional space is curled up as Kaluza and Klein suggested or whether we live in a brane world floating in a vast ten-dimensional space is a hot topic currently being discussed. A possible way of detecting the extra dimension has been suggested and could be tested by the Large Hadron Collider (LHC) experiment.

18.3

Aharonov–Bohm Effect

The gauge potential has a freedom of gauge transformation, which casts doubt on it being an observable. In classical electromagnetism, it is merely a mathematical entity, useful but of no physical significance. The electromagnetic field $F^{\mu\nu}$, on the other hand, is gauge invariant and physical. However, in quantum mechanics, in describing the electromagnetic interaction in the Schrödinger equation, we use $A^\mu = (\phi/c, \mathbf{A})$, which suggests the potential plays a fundamental role. In fact, in quantum mechanics, A_μ is an observable and plays a more fundamental role than the electromagnetic field. Its role was suggested in the geometrical interpretation of the gauge fields, but is most clearly demonstrated by the Aharonov–Bohm effect [8].⁵⁾ It is the effect that the gauge potential causes in a well-known double slit experiment with electrons (see Fig. 18.8).

We first describe the electron wave's behavior without an electromagnetic field. The wave function at point R can be obtained from that at P by the following procedure. Starting from P, electrons reach R either through slit Q or through slit S. Denote the two paths as C_1, C_2 , the two wave functions at R as ψ_1, ψ_2 and the phases that ψ_1, ψ_2 have acquired as α_1, α_2 . The particle distribution at R is given by

$$|\psi|^2 = |\psi_1 + \psi_2|^2 = |\psi_2|^2 |1 + \exp\{-i(\alpha_1 - \alpha_2)\}|^2 |\psi_1/\psi_2|^2 \quad (18.35)$$

From the expression, one sees that the absolute value of the phase is not an observable but its difference is. Next, we place a confined solenoidal magnetic field (perpendicular to the plane) at a place far away from the electron path. If the solenoid is made slim and long enough that no magnetic field leaks out, the electrons never feel the field. Then classically the solenoid should have no effect on the electron path. However, observation shows that the interference pattern changes when the field is turned on. The effect is referred to as the Aharonov–Bohm effect.

The reason is as follows. Even if the field strength vanishes, the potential need not. The condition is $\mathbf{B} = \nabla \times \mathbf{A}(x) = 0$, which restricts the allowed functional form of the potential outside the solenoid as

$$\mathbf{A}(x) = \frac{1}{q} \nabla f(x) \quad (18.36a)$$

Turning on the magnetic field produces the gauge potential, which transforms the wave function to

$$\psi \rightarrow e^{-if_i(x)} \psi_i(x) \quad (18.36b)$$

Here, we attached suffix i to differentiate phases picked up by electrons in the two paths. One should note that this is the requirement of the gauge principle.

5) It was first discovered by Ehrenberg and Siday in 1949 [127] and rediscovered by Aharonov and Bohm.

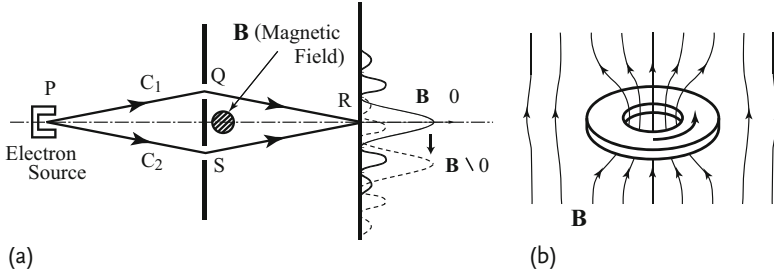


Figure 18.8 (a) Aharonov–Bohm effect. A double-slit experiment can demonstrate clearly the wave nature of the electron beam. If a confined solenoidal magnetic field is placed far away from the electron paths, the interference pattern changes. Although the magnetic field vanishes outside the solenoid, the poten-

tial does not. This is clear proof of the gauge potential affecting the electron's motion. (b) Magnetic flux quantization. When there is a superconducting ring current, the magnetic field is repelled from the superconductor and the flux inside the ring is quantized, showing another role of the gauge potential.

Turning on the magnetic field means the gauge potential is produced by a gauge transformation that accompanies the corresponding phase transformation in the matter field. Compare Eqs. (18.6) and (18.7). Solving Eq. (18.36a) for $f_i(x)$ we can express f_i as a line integral

$$f_i(x) = q \int_{P, C_i}^R A(x) \cdot dx \quad (18.36c)$$

Substituting Eqs. (18.36b) and (18.36c) into Eq. (18.35), we have a modified electron distribution with the magnetic field on:

$$|\psi_B|^2 = |\psi_2|^2 \left| 1 + \exp \left\{ -iq \oint A \cdot dx - i(\alpha_1 - \alpha_2) \right\} \right| \left| \frac{\psi_1}{\psi_2} \right|^2 \quad (18.36d)$$

$$\oint A \cdot dx = \left(\int_{C_1} - \int_{C_2} \right) A \cdot dx \quad (18.36e)$$

$\oint$ is a line integral along a closed loop made by C_1 and C_2 . If there is no magnetic field inside the loop, the integral vanishes and the interference pattern does not change. But if there is a field somewhere inside the loop, the magnetic flux in the loop does not vanish. Namely,

$$\Phi = \int B_n dS = \oint A \cdot dx \neq 0 \quad (18.37)$$

where Stokes' theorem was used in going from the second to the third equation. Despite the fact that the electron never passes a region where the magnetic field exists, it perceives the field and the interference pattern changes. The prediction of Eq. (18.36d) has been experimentally confirmed [303, 306]. The Aharonov–Bohm effect with an electric field is also observed [377]. Since the electron never sees the magnetic field, it must be the gauge field that is acting on the electron. Thus the

Aharonov–Bohm effect has clearly established that the gauge potential is real and physical and that the electromagnetic field $F_{\mu\nu}$ alone is not enough to describe the electromagnetic phenomena. Since $F_{\mu\nu}$ can be derived from the potential by differentiation, the potential is a physical object more fundamental than the field.

Note, this shows that the gauge potential itself is not an observable, but its difference is. When it manifests itself as an observable, the gauge invariance is maintained. Notice that the line integral defined in Eq. (18.37) is gauge invariant if the potential itself can be varied locally.

Another interesting example related to the Aharonov–Bohm effect is the superconducting ring current in the presence of a magnetic field [Fig. 18.8]. Owing to the Meissner effect, the magnetic field cannot exist in the superconductor. However, the electron wave function acquires a phase given by Eq. (18.37) every time it goes around the ring. The requirement that the wave function be single valued forces the magnetic flux to be quantized:

$$\Phi = 2\pi \frac{\hbar}{2e} n = 2.067833636 \times 10^{-15} \text{ Wb} \times n, \quad n = 0, \pm 1, \pm 2, \dots \quad (18.38)$$

This is also experimentally confirmed [118, 218] and has wide applications, for example in the Josephson junction. The factor two in the denominator is related to the fact that the supercurrent is produced by condensed Cooper pairs, which carry charge $-2e$.⁶⁾

In the geometrical interpretation of the gauge potential, the Aharonov–Bohm effect is easy to understand. The magnetic flux is the total curvature of the closed area. Consider a cone with its apex rounded (see Fig. 18.9a). The round area is where the magnetic field exists and the space (two-dimensional plane in the figure) is curved, but the other area is flat in the sense that the local curvature vanishes. This is easily seen by cutting the cone and spreading it out on a flat plane (Fig. 18.9b). The side plane of the cone is tangential to the flat plane and local curvature vanishes everywhere. Both paths ABC and ADE are straight lines in the flat space locally and the direction keeps its original value. Namely, the electron on the path never experiences the magnetic field, but at point A(= E) where the two paths meet, there is a deficit angle. Geometrically, the local curvature always vanishes, but the total curvature does not if the loop surrounds the apex area. If the loop does not contain the apex area, the total curvature, hence the deficit angle, vanishes. The electron perceives the total curvature, which is equivalent to feeling the electromagnetic force, as the global behavior of the space geometry.

Notice that the gauge potential is nonzero outside the solenoid. As we stated in Sect. 18.2.7, the potential can be made to vanish everywhere only when the electromagnetic field vanishes identically. If it exists somewhere, no matter how far sep-

6) The Cooper pair consists of two electrons with opposite momentum and spin. They attract each other through phonon (the quantum of lattice vibration) exchange and make a bound state, which is the Cooper pair. Normally, thermal motion destroys the bound states, but below a critical

temperature a great number of Cooper pairs are produced (Bose–Einstein condensation) and act coherently, making a quantum fluid of electrons. It is an example of spontaneous symmetry breaking, which will be discussed later in this chapter.

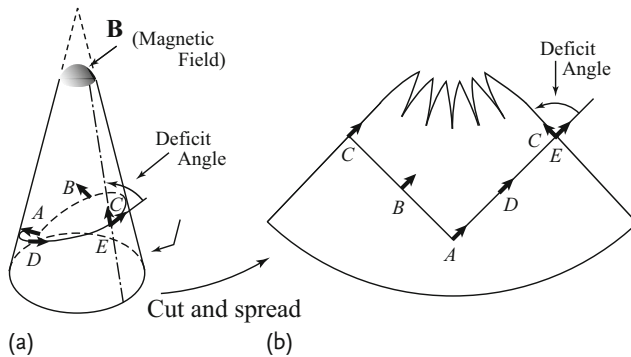


Figure 18.9 An example where the effect of the gauge potential A_μ is manifest. The electromagnetic field is interpreted as the curvature of a hyperspace. A particle passing where no magnetic field exists (a flat plane) feels its

effect when the space is globally curved like the surface of a cone. The two particle's trajectories are ABC and ADE , and when they meet at $C (= E)$, their direction is different.

arated, it deforms the space and no Cartesian coordinates can exist. The geodesic line in the curved space looks bent if projected on the Cartesian coordinates, meaning the electron perceives the electromagnetic force.

Considering the role of the gauge potential as the connection, which is locally defined but also contains global information about the entire space, the connection's controlling power must prevail throughout the whole space. Therefore, its quantum must have vanishing mass, otherwise its control would not reach far beyond its Compton wavelength. Thus, the geometrical interpretation of gauge theory provides a natural explanation of the vanishing mass of the force quantum.

Problem 18.1

Derive the vector potential outside a solenoid of radius a .

An Experiment to Prove the Aharonov–Bohm Effect The Aharonov–Bohm effect touches fundamental issues in quantum mechanics. Experimental tests of the Aharonov–Bohm effect were carried out immediately after their proposal (for a review see [303]) and confirmed it. However, with the emergence of the gauge principle as a successful axiom underlying all the forces, many subtle theoretical as well as experimental questions were raised as to the validity of the effect. Theoretically, locality, reality of the gauge potential and the single valuedness of the wave function, etc. were at issue. A major complaint on the experimental side was that the effect can be explained by leakage of the magnetic field.

Here, we introduce an experiment to prove the Aharonov–Bohm effect that responded to the criticism [306, 367, 368]. A solenoid or magnetized whiskers, however long and confined, have an outer field in order to close the flow of the magnetic flux. Consequently, the experiment was carried out on a toroidal magnet using electron holography. A schematic layout for the hologram formation and its reconstruction method are depicted in Fig. 18.10a,b. A specimen (toroidal magnet)

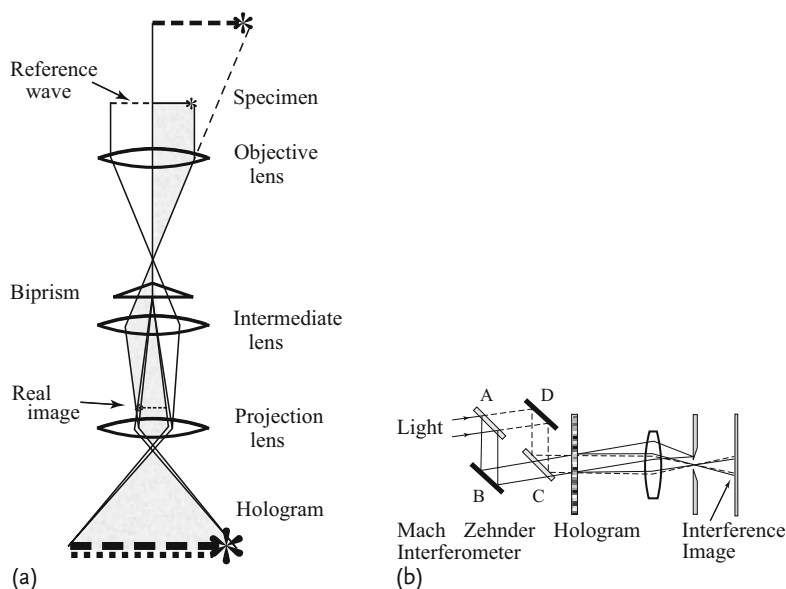


Figure 18.10 (a) Schematic diagram of the electron optical system for hologram formation. An image of the specimen is formed on a photographic film. A biprism is employed to cause the objective wave and reference wave to overlap and form a hologram. (b) Schematic diagram of hologram reconstruction. Col-

limited laser light is split into two beams and both illuminate the hologram. One (ABC path) reconstructs the image, which is superposed with the other transmitted through the ADC path (comparison beam) to produce an interference micrograph. Figures adapted from [306].

was illuminated by an electron beam. Its image was magnified and recorded on a photographic film. Only half of the beam was used. The other half was used as a reference beam, which was overlapped with the image by a biprism to form the hologram. To reconstruct the original image from the hologram, two laser beams separated by a half-silvered mirror were used. One, which goes along the path ABC in the figure, reconstructs the image, focused on the observation plane. Note that the beam splits into three after going through the hologram, one is transmitted straight forward and the two others are the image beam and its conjugate emitted at angles tilted symmetrically to the transmitted one. The other beam (path ADC) was used as the comparison beam. It also splits into three, only the transmitted beam is superposed on the observation plane to form the interference images. Two types of interference images were used: a contour map and an interferogram. The former is obtained by setting the reconstructed and transmitted waves parallel, as illustrated in Fig. 18.10b. The latter, i.e. the interferogram, is obtained by tilting the two waves to each other ($\sim 10^{-3}$ rad).

Figure 18.11a is a contour map illustrating magnetic flux lines. There, a constant flux flow of h/e lies between two adjacent lines. An interferogram showing wave fronts and a schematic illustration are given in Fig. 18.11b and c. The photograph reveals that phase differences exist between two electron beams that have passed

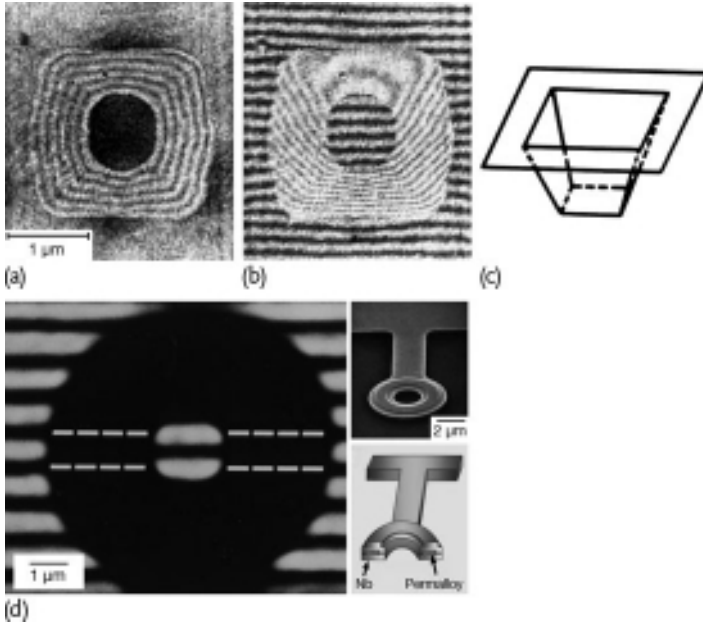


Figure 18.11 (a–c) An electron beam illuminated the toroid from above. (a) A contour map showing magnetic flux lines. (b) Interference pattern showing the retarded phase of the electrons going through the hole. (c) Illustration of the wave front. (d) Interference

pattern made with a Permalloy toroid wrapped with superconducting Nb to prevent magnetic flux leakage. The bridge stretching upward is a thermal contact to control the temperature of the toroid [367, 368].

through the inner and outer spaces of a toroidal magnet, where there are no magnetic fields. In addition, from (b), the phase displacement was counted as 5.5λ , agreeing with the theoretical value of 6λ to within 20%. However, this interferogram is still taken with part of the electrons going through the magnetic field.

The third picture was taken with a toroid completely shielded by superconducting material. The toroid is made of ferromagnetic Permalloy sandwiched between two SiO_2 substrates and surrounded by niobium (Nb). Below the transition temperature T_c , the niobium becomes superconducting and repels the magnetic field completely by the Meissner effect, thus eliminating it inside the toroid. Its presence, moreover, quantizes the magnetic flux $\Phi = n(h/2e)$, $n = \text{integer}$. Shielding makes the interference stripes invisible on the toroid, and only relative displacements of wave fronts modulo the wavelength can be detected. For the superconducting phase ($T < T_c$), the phase shift is exactly 0 for n even and π for n odd. The picture shows a case for n odd (π). One can see the inside stripe is exactly in the middle of two outside stripes.

Thus, the experiment has established the Aharonov–Bohm effect firmly and consequently the physical role of the gauge potential. If it is not accepted, nonlocality of the magnetic field has to be argued. The existence of quantized flux touches

another fundamental issue in quantum mechanics, i.e., single valuedness of wave functions.

18.4

Nonabelian Gauge Theories

18.4.1

Isospin Operator

Up to now, we have regarded the internal space as a one-dimensional complex space. Our arguments in Sect. 18.2 can be extended directly to an n -dimensional complex space. The major difference is that for $n \geq 2$ two consecutive coordinate transformations ($C = AB$) are different if the order is reversed ($C' = BA \neq AB$). Because of this, the field strength and its equation become nonlinear, making mathematical manipulations complicated. Extension of the formalism to $n = 2$ complex space is to introduce $SU(2)$ (isospin) symmetry. In Chap. 9, we introduced isospin of the strong interaction by considering $u \leftrightarrow d$ symmetry or, more precisely, by requiring the equation of motion to be invariant for unitary transformations:

$$\psi \equiv \begin{bmatrix} u \\ d \end{bmatrix} \rightarrow \begin{bmatrix} u' \\ d' \end{bmatrix} = \begin{bmatrix} U \end{bmatrix} \begin{bmatrix} u \\ d \end{bmatrix} \quad (18.39)$$

Without loss of generality, the unitary matrix can be expressed as

$$U = \exp \left(-i \sum_{a=1}^3 \alpha_a t_a \right) = \exp(-i \boldsymbol{\alpha} \cdot \mathbf{t}) \quad (18.40)$$

which has the form of phase transformations. In general, $SU(n)$ transformation can be expressed in this form using $n^2 - 1$ traceless hermitian matrices called the generators of the group. When t_a are noncommutative, the representation matrices U are noncommutative, too, and the group is referred to as nonabelian. The group characteristics are essentially⁷⁾ determined by their commutation relation (Lie algebra), which is expressed as

$$[t_a, t_b] = \sum_{c=1}^{n^2-1} i f_{abc} t_c \equiv i f_{abc} t_c \quad (18.41)$$

where f_{abc} are completely antisymmetric constants referred to as structure constants of the group. For the case of $SU(2)$, $t_a = \tau_a/2$ holds, where τ_a are Pauli matrices and

$$f_{abc} = \epsilon_{ijk} \quad SU(2) \quad (18.42)$$

7) Meaning those of local nature. Those of global nature, such as connectedness, may be different. For instance, $SU(2)$ and $SO(3)$ share the same structure constants but there is 1:2 correspondence between the two (see Sect. A.2 in Appendix A).

In an N -dimensional representation space spanned by N -component vectors, U are $N \times N$ unitary matrices. When $N = n$, the matrices are referred to as the fundamental representation, and if $N = n^2 - 1$ they are referred to as the adjoint representation. In the following, we limit our discussion to $SU(2)$, but conclusions are applicable to any $SU(n)$.

If the Lagrangian is invariant under $SU(2)$ transformation of the form Eq. (18.40), where α_a 's ($a = 1 \sim 3$) are constants, Noether's theorem predicts the existence of a conserved isospin current for the matter field with isospin I :

$$J_a^\mu(x) = g \bar{\psi} \gamma^\mu t_a \psi$$

$$\psi = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_N \end{bmatrix} \quad (18.43)$$

where $N = 2I + 1$ and ψ is a vector in isospin space. The generators t_a are $N \times N$ matrices. The space integral of Eq. (18.43)

$$\hat{I}_a = \int d^3x J_a^0(x) \quad (18.44)$$

defines a conserved operator, which is referred to as isospin. In the following, we sometimes refer to it as isospin charge, meaning it generates the force field just as the electric charge, which is a conserved observable, generates the electromagnetic force field. Note that a field having $I = 0$ does not generate a force field just as the neutral charge does not generate the electric field. The isospins are operators to generate isospin rotations ($SU(2)$ phase transformations), i.e.

$$e^{i\alpha \cdot \hat{I}} \psi e^{-i\alpha \cdot \hat{I}} = \exp[-i\alpha \cdot \mathbf{t}] \psi \quad (18.45)$$

which is the global gauge transformation of $SU(2)$. The formalism we developed in the previous section can be directly applied. Since $SU(2)$ transformation is mathematically almost identical to three-dimensional space rotation, our analogy of the gauge transformation to a rotation in the internal space is intuitively more appealing. By gauging the $SU(2)$ transformation, or equivalently replacing the derivatives with the covariant derivatives, we can obtain the Lagrangian of interacting fields.

18.4.2

Gauge Potential

From now on, we consider matter fields belonging to $I = 1/2$ or the fundamental representations of $SU(2)$, as is the case for the weak interaction of quarks and leptons, though the following discussions are valid for any isospin value. The gauge potential of $SU(2)$ can be introduced by directly applying Eq. (18.19). Since there are three ($n^2 - 1$ in general) parameters in Eq. (18.40) for gauge transformation of the matter field, we need three gauge potentials to keep the local gauge transformation invariant. This requirement restricts the form of the gauge potential. All

we have to do is to replace $A_\mu(x)$ by

$$A_\mu(x) \rightarrow [\mathcal{A}_\mu(x)]_{ij} = \sum_a A_{a\mu}(x) [t_a]_{ij} = \mathbf{A}_\mu \cdot [\mathbf{t}]_{ij} \quad (18.46)$$

which are 2×2 matrices acting on the two-component matter field ψ . The infinitesimal parallel transport of the matter field is defined as

$$\begin{aligned} \psi_\parallel(x + dx) &= \psi(x) - ig\mathcal{A}_\mu\psi = \psi - ig\mathbf{A}_\mu \cdot \mathbf{t}\psi \\ \text{or in components} \quad \psi_{\parallel i}(x + dx) &= \psi_i(x) - ig\mathbf{A}_\mu(x) \cdot [\mathbf{t}]_{ij} \psi_j(x) \end{aligned} \quad (18.47)$$

where g is the coupling constant corresponding to e in QED. The corresponding covariant derivative can be defined as

$$[D_\mu]_{ij} = [\partial_\mu + ig\mathcal{A}_\mu]_{ij} = \delta_{ij}\partial_\mu + ig\mathbf{A}_\mu \cdot \mathbf{t}_{ij} \quad (18.48)$$

The gauge transformation of the potential can be obtained by the requirement that it keeps the equation of motion invariant when the matter field undergoes a gauge transformation

$$\psi(x) \rightarrow \psi'(x) = U(x)\psi(x) = \exp[-ig\boldsymbol{\alpha}(x) \cdot \mathbf{t}]\psi(x) \quad (18.49)$$

where we have extracted the factor g from the phase for later convenience. If we denote the gauge-transformed covariant derivative of D_μ as D'_μ , the Lagrangian is kept invariant if

$$\begin{aligned} (D_\mu\psi)' &= D'_\mu\psi' = D'_\mu(U\psi) = U(D_\mu\psi) \\ \therefore D'_\mu &= UD_\mu U^{-1} \end{aligned} \quad (18.50)$$

which defines the transformation property of the gauge field:

$$\mathcal{A}'_\mu = U\mathcal{A}_\mu U^{-1} - \frac{i}{g} U\partial_\mu U^{-1} \quad (18.51)$$

Setting U as an infinitesimal transformation (α small), we have

$$A'_{a\mu} = A_{a\mu} + f_{abc}\alpha_b A_{c\mu} + \partial_\mu\alpha_a(x) + \dots \quad (18.52)$$

which does not depend on the representation matrices t_a . Adopting vector notation in the internal space, we can express the above relation also as

$$\begin{aligned} \mathbf{A}'_\mu &= \mathbf{A}_\mu + \boldsymbol{\alpha} \times \mathbf{A}_\mu + \partial_\mu\boldsymbol{\alpha}(x) + \dots \\ (\boldsymbol{\alpha} \times \mathbf{A}_\mu)_a &= \sum_{b,c} f_{abc}\alpha_b A_{c\mu} \end{aligned} \quad (18.53)$$

Since, in quantized field theory, the gauge potential is a field on an equal footing with matter fields, we refer to the gauge potential as the gauge field here. The

meaning of the equation is that the gauge field is rotated in the internal space, which means that the gauge field can be phase transformed by the isospin operator given in Eq. (18.44). This is possible only when the gauge field carries isospin, hence the gauge field itself is the source of the force. This feature did not exist in $U(1)$ gauge symmetry. In other words, while the photon is neutral and cannot generate the electromagnetic field by itself, the $SU(2)$ gauge field carries isospin charge and can be the source of the force field. The isospin of the gauge field is $I = 1$ since there are three components. As the action of the force is represented by coupling to the gauge field in the Lagrangian, the interaction of the $SU(2)$ gauge field is nonlinear, which is the conspicuous characteristic of nonabelian gauge theory.

18.4.3

Isospin Force Field and Equation of Motion

The curvature of the isospin hyperspace can be obtained by the same procedure as for $F_{\mu\nu}$ in QED. But this time the gauge field are matrices and noncommutative, which means the order of parallel transport has to be chosen carefully. Let us move the matter field in the $x = x^\mu, y = x^\nu$ plane and take the same infinitesimally small loop as was done for the electromagnetic field (Fig. 18.5).

Omitting variables other than x, y , we have

1. Parallel transport from $P \rightarrow Q$:

$$\psi_{\parallel}(x + dx, y) = [1 - ig\mathcal{A}_x(x, y)dx]\psi(x, y) \quad (18.54a)$$

2. Parallel transport from $Q \rightarrow R$:

$$\psi_{\parallel}(x + dx, y + dy) = [1 - ig\mathcal{A}_y(x + dx, y)dy]\psi_{\parallel}(x + dx, y) \quad (18.54b)$$

Denoting the parallel-transported ψ via route $P \rightarrow Q \rightarrow R$ as $\psi_{\parallel}^{\text{PQR}}$, we find

$$\begin{aligned} \psi_{\parallel}^{\text{PQR}} &= [1 - ig\mathcal{A}_y(x + dx, y)dy][1 - ig\mathcal{A}_x(x, y)dx]\psi(x, y) \\ &\simeq [1 - ig\{\mathcal{A}_y(x, y)dy + \mathcal{A}_x(x, y)dx + \partial_x \mathcal{A}_y(x, y)dx dy\} \\ &\quad - g^2 \mathcal{A}_y(x, y)\mathcal{A}_x(x, y)dx dy]\psi(x, y) \end{aligned} \quad (18.55)$$

The transport $R \rightarrow S \rightarrow P$ can be obtained from the transport $P \rightarrow S \rightarrow R$ by changing the sign of the path and the transport $P \rightarrow S \rightarrow R$ by exchanging $x \leftrightarrow y$ in $\psi_{\parallel}^{\text{PQR}}$ in Eq. (18.55). Namely,

$$\begin{aligned} \oint_{\text{PQRS}} \psi_{\parallel}(x, y) &\simeq \psi_{\parallel}^{\text{PQR}} - \psi_{\parallel}^{\text{PSR}} \\ &= -ig[\partial_x \mathcal{A}_y - \partial_y \mathcal{A}_x + ig(\mathcal{A}_x \mathcal{A}_y - \mathcal{A}_y \mathcal{A}_x)]\psi(x, y)dx dy \end{aligned} \quad (18.56)$$

Comparing this with the defining equation of the curvature Eq. (18.21a), we obtain the field strength

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu + ig[\mathcal{A}_\mu, \mathcal{A}_\nu] = \frac{1}{ig}[D_\mu, D_\nu] \quad (18.57)$$

The second equality can be proved by actually performing the arithmetic of the commutator. It is an operator relation, but one sees that it does not contain any derivative after taking the commutator. For an abelian group $[\mathcal{A}_\mu, \mathcal{A}_\nu] = 0$ and Eq. (18.57) reduces to the electromagnetic field.

Problem 18.2

Derive the second equality of Eq. (18.57).

Using

$$[\mathcal{A}_\mu, \mathcal{A}_\nu] = [A_{a\mu}t_a, A_{b\nu}t_b] = i f_{abc} A_{a\mu} A_{b\nu} t_c = i(\mathbf{A}_\mu \times \mathbf{A}_\nu) \cdot \mathbf{t} \quad (18.58)$$

$\mathcal{F}_{uv}$ can also be expanded in terms of t_a :

$$\mathcal{F}_{\mu\nu} = F_{a\mu\nu}t_a = F_{\mu\nu} \cdot \mathbf{t} \quad (18.59a)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - g A_\mu \times A_\nu \quad (18.59b)$$

The gauge transformation of $\mathcal{F}_{\mu\nu}$ can be obtained from Eq. (18.57)

$$\mathcal{F}'_{\mu\nu} = \frac{1}{ig} [D'_\mu, D'_\nu] = \frac{1}{ig} U [D_\mu, D_\nu] U^{-1} = U \mathcal{F}_{\mu\nu} U^{-1} \quad (18.60)$$

As $\mathcal{F}_{\mu\nu}$ is a matrix, $\mathcal{F}'_{\mu\nu} \neq \mathcal{F}_{\mu\nu}$, i.e. the nonabelian field strength $\mathcal{F}_{\mu\nu}$ is not gauge invariant. When one wants to construct a gauge-invariant Lagrangian, one can use

$$\text{Tr}[\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}] = \frac{1}{2}F_{\mu\nu} \cdot F^{\mu\nu} \quad (18.61)$$

To obtain the derivative $D_\lambda \mathcal{F}_{\mu\nu}$, one can choose the path depicted in Fig. 18.12. The result is

$$D_\lambda \mathcal{F}_{\mu\nu} = \partial_\lambda \mathcal{F}_{\mu\nu} + ig[\mathcal{A}_\lambda, \mathcal{F}_{\mu\nu}] = [D_\lambda, \mathcal{F}_{\mu\nu}] \quad (18.62)$$

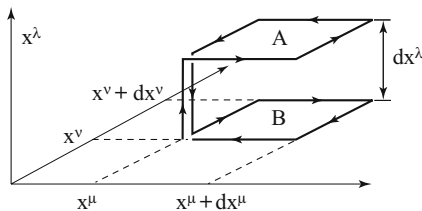


Figure 18.12 Integral path for calculating covariant derivatives $D_\lambda \mathcal{F}_{\mu\nu}$.

Then the equivalent of the Maxwell equation becomes

$$\begin{aligned} D_\mu \mathcal{F}^{\mu\nu} &= \partial_\mu \mathcal{F}^{\mu\nu} + ig[\mathcal{A}_\mu, \mathcal{F}^{\mu\nu}] = g\mathcal{J}^\nu \\ \mathcal{J}^\nu &= j_a^\nu t_a = \mathbf{j}^\nu \cdot \mathbf{t} \end{aligned} \quad (18.63)$$

or, separating the representation matrix,

$$\partial_\mu \mathbf{F}^{\mu\nu} - g\mathbf{A}_\mu \times \mathbf{F}^{\mu\nu} = g\mathbf{j}^\nu \quad (18.64)$$

which can also be obtained from the Lagrangian

$$\mathcal{L}_{SU(2)} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu} \cdot \mathbf{F}^{\mu\nu} \quad (18.65)$$

Problem 18.3

Derive Eq. (18.62).

Problem 18.4

Derive Eq. (18.64) from the Lagrangian in Eq. (18.65).

Substituting Eq. (18.62) in the Jacobi identity

$$[D_\mu, [D_\nu, D_\lambda]] + [D_\nu, [D_\lambda, D_\mu]] + [D_\lambda, [D_\mu, D_\nu]] = 0 \quad (18.66)$$

we have the $SU(2)$ equivalent of the sourceless Maxwell equation (3.49b):

$$D_\mu \mathcal{F}_{\nu\lambda} + D_\nu \mathcal{F}_{\lambda\mu} + D_\lambda \mathcal{F}_{\mu\nu} = 0 \quad (18.67)$$

Thus, we have derived a $SU(2)$ gauge-invariant Lagrangian and the associated equations of motion. The Lagrangian Eq. (18.65) has exactly the same form as that of QED, except that the matter field ψ is an $I = 1/2$ doublet and the gauge field $\mathcal{A}_\mu$ is 2×2 matrices acting on ψ and an $I = 1$ triplet. The major difference lies in that $F_{\mu\nu}$ contains nonlinear (quadratic) term and hence the Lagrangian contains up to quartic terms in the gauge fields.

So far we have discussed only $SU(2)$ gauge theory. Extension to general $SU(n)$ is straightforward. The only difference is that there exist $n^2 - 1$ generators and the same number of gauge fields and that the structure constant ϵ_{abc} has to be replaced by f_{abc} . We summarize the equations of motion for the $I = 1/2$ matter field and the $I = 1$ gauge field, which respect $SU(2)$ gauge invariance. Expecting applications to the weak interaction, we denote the $SU(2)$ gauge field as W_a^μ and its coupling strength as g_W .

$SU(2)$ Lagrangian

$$\mathcal{L}_{SU(2)} = \bar{\psi} \left[i \gamma^\mu \left(\partial_\mu + i g_W \mathbf{W}_\mu \cdot \frac{\boldsymbol{\tau}}{2} \right) - m \right] \psi - \frac{1}{4} \mathbf{F}_{\mu\nu} \cdot \mathbf{F}^{\mu\nu}$$

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}, \quad \mathbf{W}^\mu = (W_1^\mu, W_2^\mu, W_3^\mu) \quad (18.68)$$

$$F_a^{\mu\nu} = \partial^\mu W_a^\nu - \partial^\nu W_a^\mu - g_W \epsilon_{abc} W_b^\mu W_c^\nu$$

$$\left[\frac{\tau_a}{2}, \frac{\tau_b}{2} \right] = i \epsilon_{abc} \frac{\tau_c}{2} \quad (18.69)$$

The interaction Lagrangian can alternatively be expressed in terms of charged and neutral bosons as

$$-\mathcal{L}_{\text{INT}} = g_W \bar{\psi} \gamma^\mu \mathbf{W}_\mu \cdot \frac{\boldsymbol{\tau}}{2} \psi$$

$$= \frac{g_W}{\sqrt{2}} \bar{\psi} \gamma^\mu \left[\tau_+ W_\mu^+ + \tau_- W_\mu^- \right] \psi + g_W \bar{\psi} \gamma^\mu \left[\frac{\tau_3}{2} W_\mu^0 \right] \psi$$

$$W^\mp = \frac{1}{\sqrt{2}} (W_1 \pm i W_2), \quad W^0 = W_3 \quad (18.70)$$

The equations of motion are

$$(i \gamma^\mu \partial_\mu - m) \psi = g_W \mathbf{W} \cdot (\boldsymbol{\tau}/2) \psi \quad (18.71a)$$

$$\partial_\mu \mathbf{F}^{\mu\nu} - g \mathbf{W}_\mu \times \mathbf{F}^{\mu\nu} = g_W \bar{\psi} \gamma^\nu (\boldsymbol{\tau}/2) \psi \quad (18.71b)$$

18.5**QCD**

QCD is an abbreviation for quantum chromodynamics, meaning the dynamics of color charge, and obeys the same mathematical framework as QED. The QCD Lagrangian can be obtained from that of $SU(2)$ in Eq. (18.68) by simple replacements:

$$\begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \rightarrow \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad \sum_{a=1}^3 W_{a\mu} \frac{\tau_a}{2} \rightarrow \sum_{a=1}^8 A_{a\mu} \frac{\lambda_a}{2}, \quad (18.72)$$

$$\epsilon_{abc} \rightarrow f_{abc}, \quad g_W \rightarrow g_S$$

where R, G, B are Dirac fields having color R, G, B , $A_{a\mu}$ eight gluon vector fields, λ_a Gell-Mann's 3×3 matrices and the structure constant ϵ_{ijk} is replaced by f_{ijk} (see Appendix G). Corresponding to the eight generators, there are eight gauge fields

$A_{a\mu}$ ($a = 1, \dots, 8$) and the covariant derivative is defined as

$$D_\mu = \partial_\mu + ig_s \mathbf{A}_\mu \cdot \mathbf{t} = \partial_\mu + ig_s \sum_{a=1}^8 A_{a\mu} \frac{\lambda_a}{2} \quad (18.73)$$

The QCD Lagrangian is given by Eq. (18.68) with the modifications described above.

Who invented QCD is not clear. Many people suggested the existence of color degrees of freedom, including Gell-Mann. In 1964–1965, dynamical models to ascribe the source of the strong force to color were proposed [185, 193], but their models had a somewhat complicated symmetry structure. A turning point was the discovery of asymptotic freedom in 1973 by Gross, Wilczek and Politzer [189, 190, 324]. Asymptotic freedom means that the effective quark–gluon coupling constant $\alpha_s = g_s^2/4\pi$ is a decreasing function of the momentum transfer Q^2 , or more specifically behaves like $(\ln Q^2)^{-1}$. This means $\lim_{Q^2 \rightarrow \infty} \alpha_s(Q^2) \rightarrow 0$, namely a perturbative approach can be used in the high-energy region. This provided a theoretical justification of the parton model and a means to improve it. As a result, QCD surfaced as the sole candidate for the strong interaction. Asymptotic freedom is a consequence of the nonlinear gluon coupling. Color confinement is considered another consequence. Investigations using lattice QCD⁸⁾ are suggestive of it.

Since QCD is a field theory very similar to QED except for the color degree of freedom and nonlinear coupling of the gluon field, the calculation of perturbation series, i.e. that of Feynman diagrams in the lowest order, tree approximation, is almost identical to that of QED except for color counting.⁹⁾ When perturbative corrections are made to the parton model, such as emission of gluons from the quark, the structure functions develop as a function of Q^2 , i.e. a shift of the parton model from scaling at large Q^2 [$> 20 \text{ (GeV/c)}^2$] is apparent. Comparison of those predictions with precision data from electron, muon and neutrino deep inelastic scattering, as well as phenomena in e^-e^+ , established the validity of QCD in the late 1970s.

Hadron–hadron interactions can also be interpreted by the parton model [Eq. (17.150)]. QCD gave its justification and provided a recipe for how to calculate higher order corrections. A notable feature of quarks and gluons is that they

8) Lattice QCD is a technique to calculate values of various dynamical variables numerically. Four-dimensional space-time is divided into a lattice of interval a , breaking Lorentz invariance but in such a way as to preserve the gauge invariance. In the limit $a \rightarrow 0$, the Lorentz invariance is recovered, but at the moment the size is limited by computing power. Lattice QCD provides a powerful method to treat the nonperturbative aspect of QCD.

9) If one formulates QCD in the covariant gauge, longitudinal as well as scalar gluons are quantized, too. They are unphysical and should not appear in the observables.

In QED, they compensate each other internally and never appear as real photons.

In QCD, because of its nonabelian nature, the longitudinal and scalar gluons do not compensate each other exactly. A ghost field, a scalar but obeying Fermi–Dirac statistics, has to be introduced to cancel unphysical gluons. The ghost is not a physical particle but an artifact arising from adopting the covariant gauge. If one works in the axial ($A^3 = 0$) or temporal ($A^0 = 0$) gauge, one does not need the ghost, but the calculation is generally easier in the covariant gauge. See Sect. 11.7 for the existence of the ghost field.

are confined inside hadrons and at high energies only observed as jets. By observing jet phenomena, it became possible to treat the interactions between quarks and gluons at high energies quantitatively with high precision.

On the other hand, at small $Q^2 \lesssim 1 \text{ (GeV/c)}^2$ asymptotic freedom does not apply and a nonperturbative treatment is necessary. Lattice QCD provides a way to calculate hadron masses, decay constants and other dynamical variables in which a nonperturbative treatment is necessary. Also chiral perturbation theory was developed based on QCD; it handles many low-energy phenomena, including decays of heavy quarks (c, b, t). QCD is now a standard tool in nuclear physics, too.

As a theory that can reproduce various aspects of hadronic phenomena and that can calculate strong force corrections to electroweak theory, QCD, in parallel with GWS electroweak theory, has established itself as the Standard Model of particle physics. It has also developed new phenomena of its own, including the instanton effect and the strong CP problem with its associated prediction of axions and their possible effect on observables.

18.5.1

Asymptotic Freedom

In Sect. 8.1.5, we calculated the QED coupling constant as a function of Q^2 in Eq. (8.63), which is also referred to as the running coupling constant when its dependence on Q^2 is emphasized:

$$\begin{aligned}\alpha(Q^2) &= \frac{\alpha_0}{1 + \Pi_\gamma(Q^2)} \\ \Pi_\gamma(Q^2) &= \Pi_\gamma(0) + \hat{\Pi}_\gamma(Q^2) \\ \hat{\Pi}_\gamma(0) &= 0, \quad \hat{\Pi}_\gamma(Q^2) = -\frac{\alpha_0}{3\pi} \ln(Q^2/m_e^2) \quad Q^2 \gg m_e^2\end{aligned}\quad (18.74)$$

where the coupling $\alpha_0 = e_0^2/4\pi$, which is defined in terms of the bare charge e_0 , is to be considered as a mere parameter. $Q^2 = -q^2$ is the 4-momentum squared of the virtual (space-like) photon and m_e is the electron mass. $\Pi_\gamma(0)$ is a divergent quantity and has to be treated with the renormalization recipe. However, $\hat{\Pi}_\gamma(Q^2)$ is finite and physically meaningful. Expression (18.74) is obtained by summing a chain of vacuum polarization diagrams (a)–(e), which are repetitions of the one-loop contribution. Expressed in a somewhat different form

$$\frac{1}{\alpha(Q^2)} = \frac{1}{\alpha_0} - B(Q^2), \quad B(Q^2) = -\frac{1}{\alpha_0} [\Pi_\gamma(0) + \hat{\Pi}_\gamma(Q^2)] \quad (18.75)$$

The renormalization procedure replaces the divergent term by an observable, the coupling constant $\alpha(0)$ measured at infinite distance ($Q^2 = 0$). Namely, by

$$\frac{1}{\alpha(0)} = \frac{1}{\alpha_0} - B(0), \quad \alpha(0) = 1/137 \dots \quad (18.76)$$

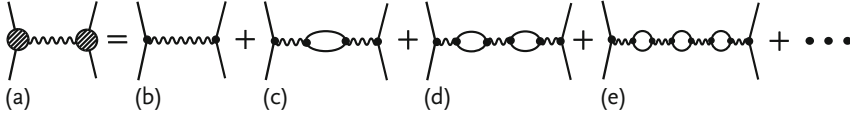


Figure 18.13 QED: Diagrams that are collected to calculate the running coupling constant: (a) is a symbolic expression for the sum of (b)–(e). (a) is called the leading log approximation (LLA).

Substituting Eq. (18.76) in Eq. (18.75), we have

$$\begin{aligned} \frac{1}{\alpha(Q^2)} &= \frac{1}{\alpha(0)} - [B(Q^2) - B(0)] = \frac{1}{\alpha(0)} - \beta(Q^2) \\ \beta(Q^2) &= \frac{1}{3\pi} \ln \frac{Q^2}{m_e^2} \end{aligned} \quad (18.77)$$

We now have a divergent-free (renormalized) expression for an effective coupling constant in terms of an observable quantity:

$$\alpha(Q^2) = \frac{\alpha(0)}{1 - \frac{\alpha(0)}{3\pi} \ln \frac{Q^2}{m_e^2}} \quad (18.78)$$

The above expression for the β function was obtained as the first-order perturbation solution, picking up only the dominant $\ln Q^2$ term in the limit $Q^2 \gg m_e^2$. However, the coupling constant obtained from the above expression actually contains the series of higher order corrections depicted in Fig. 18.13, as can be seen by expanding Eq. (18.78) in powers of $\alpha(0)$, and it is called the leading log approximation (LLA).

In QCD, we cannot subtract the coupling constant at $Q^2 = 0$, as it is infinite because of confinement. Instead, we choose some reference value $Q^2 = \mu^2$ to define the coupling constant. Denoting the QCD coupling constant as $\alpha_s(Q^2)$, we can write

$$\frac{1}{\alpha_s(Q^2)} = \frac{1}{\alpha_s(\mu^2)} - [B(Q^2) - B(\mu^2)] = \frac{1}{\alpha_s(\mu^2)} - \beta(Q^2) \quad (18.79)$$

Diagrams that contribute to the β function in the lowest order or LLA approximation to $\alpha_s(Q^2)$ are given in Fig. 18.14. The β function is expressed as¹⁰⁾

$$\beta(Q^2) = -\frac{\beta_0}{4\pi} \ln \frac{Q^2}{\mu^2} \quad (18.80a)$$

$$\beta_0 = -\left(\frac{2}{3}n_f + 5 - 16\right) \quad (18.80b)$$

10) The contribution of each diagram is gauge dependent. The physical interpretation is most clear in the physical gauge, where only two transverse gluons are quantized and the longitudinal gluon represents simply an instantaneous Coulomb potential.

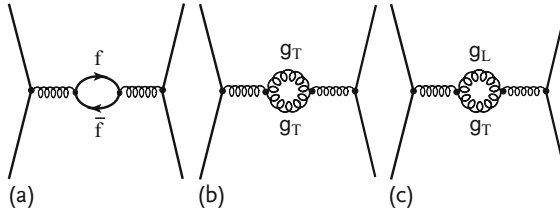


Figure 18.14 QCD: Diagrams that contribute to the quark–gluon running coupling constant. (a) is common to QED and QCD. (b,c) Gluon self-coupling diagrams. (b) Trans-

verse gluon pair. (c) Longitudinal gluon and transverse gluon pair. Repetition of types (a)–(c) are all added in LLA calculation.

The first term is the contribution of the fermion loop and n_f is the number of quark flavors.¹¹⁾ The factor (+5) in the second term in Eq. (18.80b) originates from pair creation of gluons having transverse polarizations. The third term (−16) is a contribution of one transverse gluon created together with a longitudinal gluon (Fig. 18.14c). The reason for both the first and second terms being positive is related to unitarity [Sect. (9.1.2)], i.e. $\text{Im}(T_{fi}) \sim \sum_n |T_{ni}|^2$. The amplitudes of the diagrams in Fig. 18.14 are proportional to $\sigma(q\bar{q} \rightarrow f\bar{f}, g\bar{g})$ for $Q^2 = -q^2 > 0$, as can be seen by cutting them across the loop. Namely, fermion pairs and transverse gluon pairs are physical and can be produced if energy-momentum conservation is satisfied. However, the longitudinal gluons are unphysical and do not contribute to the cross section. Therefore the β function can be negative, and in fact it is large and negative. One can see that as long as $(2/3)n_f < 11$, the β function is negative ($\beta < 0$). The QCD running coupling constant is now expressed as

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{4\pi} \left(11 - \frac{2}{3}n_f\right) \ln \frac{Q^2}{\mu^2}} \quad (18.81)$$

which approaches zero as $Q^2 \rightarrow \infty$ (asymptotic freedom). Asymptotic freedom justifies the use of perturbation theory at large Q^2 (phenomenologically at $Q \gtrsim 1 \text{ GeV}/c$) and the parton model, which treats quarks as point particles. The difference between QED and QCD is illustrated in Fig. 18.15. It is due to the color degree of freedom and nonlinear gluon couplings in QCD. The vacuum polarization effect due to physical fermions and gluons serves as a screening effect (similar to dielectricity in electromagnetism), but the effect of the unphysical gluon loop is opposite and acts as antiscreening. Note that the QED coupling constant is running, too, and when Q^2 is sufficiently large, the denominator of Eq. (18.78) could vanish formally, but if one calculates the value of Q necessary to achieve this, it is absurdly large and not attainable experimentally. In practice, the variation of $\alpha(Q^2)$ is very small in the range of Q^2 achievable today.

¹¹⁾ Note that in QED $n_f = 1$ and $\beta_0 = -2/3$, which agrees with Eq. (18.74) if we take into account the normalization difference of QCD, where what corresponds to e in QCD is actually $g/\sqrt{2}$ for historical reasons. (See the second line of Eq. (18.70) for why $g/\sqrt{2}$ is used.) n_f is the effective number of quark flavors (not always 6), whose mass satisfies the condition $m_f < Q$.

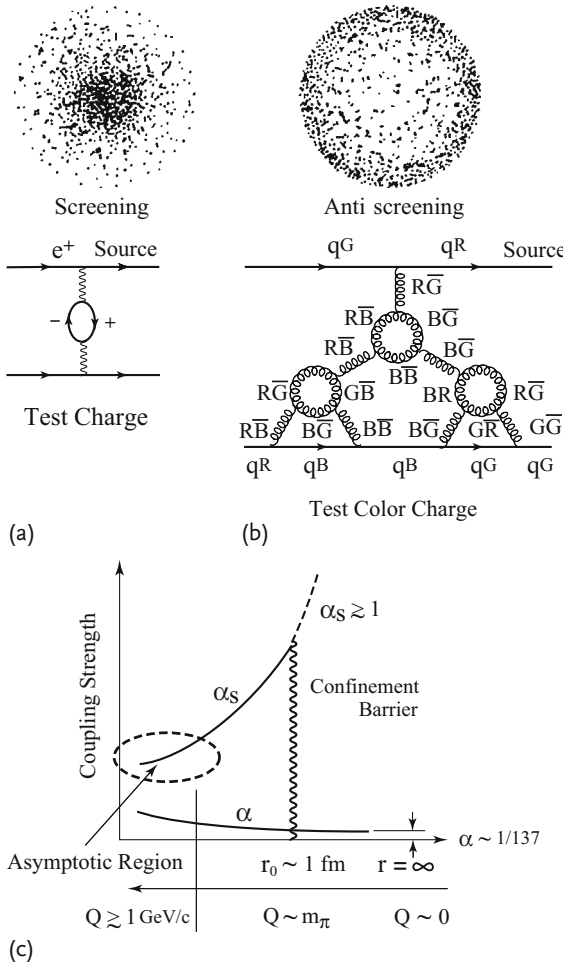


Figure 18.15 (a,b) Pictorial illustration of screening and antiscreening in QED/QCD. (a) The strength of QED coupling constant seen by a test charge decreases as it goes away from the source (dielectric effect of the vacuum). (b) The strength of α_s seen by a test color charge (another quark) increases due to the antiscreening effect of gluon self-coupling

and becomes the confining potential. (c) α , the QED coupling constant seen by a test charge, is finite at $r = \infty$ ($Q^2 = 0$) and increases at shorter distances as the screening is weaker. α_s is meaningless at large distances $r \gtrsim 10^{-13}$ cm because of confinement. It is small at short distances.

On the other hand, the variation of $\alpha_s(Q^2)$ within the technically achievable range is sizeable. Therefore, the treatment of the running coupling constant normalized at an arbitrary value of μ^2 takes a bit of getting used to. The observable quantities should not depend on which μ one chooses. The requirement takes the form of a renormalization group equation and constrains the behavior of the coupling constant. Here, we only show that regardless of μ^2 choice we can choose a

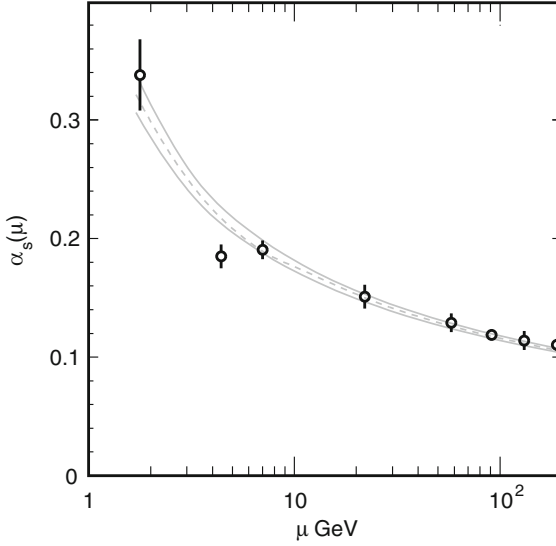


Figure 18.16 Measured α_s as a function of scale μ ; μ corresponds to Q in the text [311].

universal constant that is the fundamental constant of QCD. Equation (18.79) with Eq. (18.80) substituted gives

$$\frac{1}{\alpha_s(Q^2)} - \frac{\beta_0}{4\pi} \ln Q^2 = \frac{1}{\alpha_s(\mu^2)} - \frac{\beta_0}{4\pi} \ln \mu^2 \quad (18.82)$$

The lhs is a function of Q^2 and the rhs of μ^2 only. Consequently they are constant and we equate them with $-(\beta_0/4\pi) \ln \Lambda_{\text{QCD}}^2$:

$$\begin{aligned} \frac{1}{\alpha_s(Q^2)} - \frac{\beta_0}{4\pi} \ln Q^2 &= -\frac{\beta_0}{4\pi} \ln \Lambda_{\text{QCD}}^2 \\ \therefore \alpha_s(Q^2) &= \frac{4\pi}{\beta_0 \ln \frac{Q^2}{\Lambda_{\text{QCD}}^2}} = \frac{12\pi}{(33 - 2n_f) \ln \frac{Q^2}{\Lambda_{\text{QCD}}^2}} \end{aligned} \quad (18.83)$$

Figure 18.16 shows observed values of α_s at various Q^2 and indeed confirms the logarithmic dependence of the coupling constant. Note, however, that what is actually measured is $\alpha(Q^2)$ and Λ_{QCD} has to be calculated from it. The experimental value of Λ_{QCD} is 150–300 MeV. In principle, it should not depend on Q^2 , but Eq. (18.83) is an approximation using one-loop diagrams.¹²⁾ There still is great uncertainty as to the true value of Λ_{QCD} .

It is interesting to note that Eq. (18.83) becomes infinite at $Q = \Lambda_{\text{QCD}}$, which is on the order of $0.5 \text{ fm} = 0.5 \times 10^{-13} \text{ cm}$, roughly equal to half the hadron size. As the formula is valid only in the asymptotic region, Q has to be at least larger than $\sim 1 \text{ GeV}/c$, where $\alpha_s(1 \text{ GeV}/c) \sim 0.5$ for the formula (18.83) to be valid. It is

¹²⁾ For higher order expressions of $\alpha_s(Q^2)$ or Λ_{QCD} see [311].

conventional to express experimental values of α_s converted to that at $Q = m_Z = 91 \text{ GeV}/c$. The experimental value is [311]

$$\alpha_s(m_Z) = 0.1170 \pm 0.0012 \quad (18.84)$$

18.5.2

Confinement

Confinement is also considered to be a result of nonlinear gluon coupling. It is not rigorously proven yet, although lattice QCD calculations are indicative of it. Qualitatively, the reasoning for confinement is well understood. The argument closely follows that of the Meissner effect in superconductivity. It is a perfect diamagnetism induced by a Bose–Einstein condensate of Cooper pairs. They are bound states of an electron pair with opposite momentum and spin orientation attracted to each other through lattice vibrations (phonons). The attractive force, normally hidden by thermal motion, manifests itself at low temperature. When the magnetic field tries to penetrate any medium, eddy currents made by freely moving electrons are produced to compensate it, a familiar phenomenon in electromagnetism that is referred to as Lenz's effect. Normally the effect is small, but in the superconductor, the compensation works maximally and shows perfect diamagnetism. In the language of field theory, the photon has acquired a mass and the force becomes short ranged.

The invasion of the magnetic flux into the superconductor is limited by the Compton wavelength of the massive photon. If the strength of the magnetic field exceeds some critical value B_c , the magnetic energy can destroy the Cooper pairs and melt the condensate, and the superconductor is converted to a normal conductor. This is referred to as a type I superconductor (Fig. 18.17a). In a type II superconductor, the magnetic field can partially break the superconductivity for the field range $B_{c1} < B < B_{c2}$, and flux tubes that penetrate the superconductor are created (Fig. 18.17b). Within the narrow confined region of the tube, $B > B_{c2}$ and the medium is normal conducting. Outside the tube, it is superconducting and the magnetic flux is repelled from it by supercurrent circulating around the tubes. A simple explanation of superconductivity by Ginzburg and Landau is given in the boxed paragraph on p. 769 Confinement equation. 18.85.

Now imagine a colored quark in a medium in a sort of superconducting state, but with the roles of the magnetic and electric fields exchanged. The color electric field is created around the quark. In its vicinity, the field strength is strong enough that it melts the superconductivity of the medium, and the flux expands isotopically (Fig. 18.17c). The flux density decreases in proportion to the inverse of the distance squared and the field strength is of Coulomb type, just like the electric field in QED. At long distances, the field is not strong enough to destroy the superconducting state. As the total flux must be conserved unless it is absorbed by (an) other anticolored quark(s), the requirement of minimum energy state constrains the flux to be confined in a narrow tube with constant cross-sectional area, keeping its field strength above the critical value. This penetrates the medium (Fig. 18.17d) just like

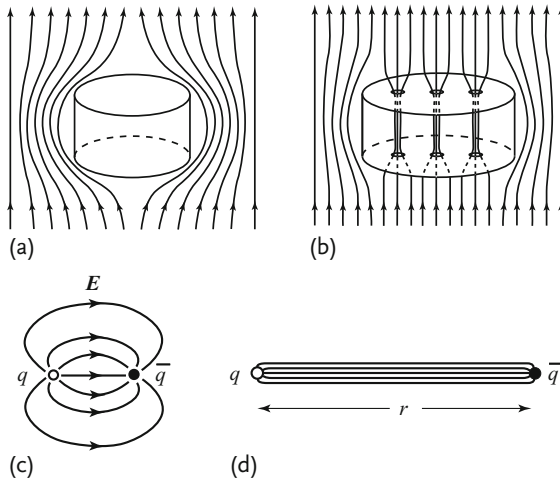


Figure 18.17 Meissner effect. (a) Type I superconductor. The magnetic flux is repelled from the superconductor for $B < B_c$. (b) Type II superconductor. For $B_{c1} < B < B_{c2}$ the magnetic flux is squeezed and flux tubes can penetrate the superconductor. Inside the tubes it is

a normal conductor. (c) The color electric field flux expands uniformly in the vicinity of the color charge and the force is Coulomb-like. (d) For large r , the flux is squeezed to form a tube with constant cross section. The flux density and the force strength stay constant.

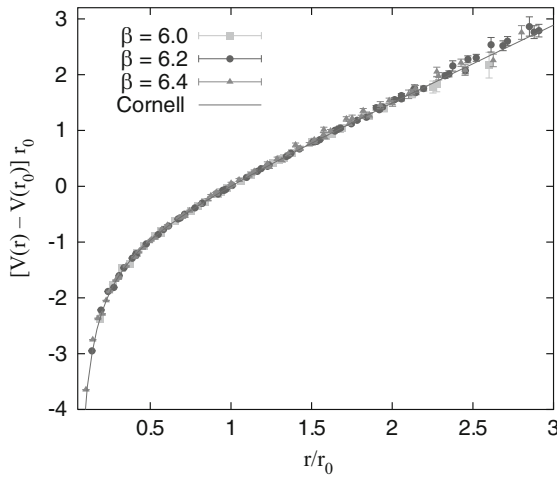


Figure 18.18 Confinement potential calculated by lattice QCD. The Cornell potential is given by $V(r) = \sigma r - e/r$, ($\sigma \sim 0.1 \text{ GeV} \times 10^{-13} \text{ cm}$, $e \sim 0.7 \text{ GeV}/10^{-13} \text{ cm}$) normalized to $V(r_0) = 0$. $\beta = 2N/g^2$ for $SU(N)$ [39].

the magnetic flux tube in the type II superconductor. As the flux per unit area in the tube is kept constant, the strength of the force is also kept constant. Its accumulated energy grows in proportion to the tube's length, in other words, it behaves like a stretched string with constant tension. The tube is terminated by another anticolor charge or the antiquark. Thus the potential energy between the two quarks is expressed by a combination of the Coulomb and linear potentials:

$$V(r) \sim \frac{a}{r} + br \quad (18.85)$$

It qualitatively reproduces the confining potential introduced to explain charmonium levels in Eq. (14.60). Figure 18.18 shows an example of a lattice QCD calculation of the potential.

Ginzburg–Landau Theory of Superconductivity

Superconductivity is characterized by its permanent electric current and perfect diamagnetism (Meissner effect [273]). When we write e^* ($= 2e$) and m^* ($= 2m$) for the electric charge and mass, respectively, of the supercurrent carrier (i.e. Cooper pair), its behavior in the magnetic field is described by the Ginzburg–Landau free energy [170]

$$F = F_n + \frac{1}{2}(\nabla \times \mathbf{A})^2 + \frac{1}{2m^*} \left| \left(\frac{\nabla}{i} - e^* \mathbf{A} \right) \psi \right|^2 + \alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4, \quad (18.86)$$

$$\alpha = \alpha_0(T - T_c), \quad \beta > 0 \quad (18.87)$$

where F_n is the free energy of the normal conductor and ψ is the Cooper pair's wave function, commonly referred to as the order parameter. The equation of motion for ψ and $\mathbf{A}$ can be obtained by minimizing F , which gives

$$\frac{1}{2m^*} \left(\frac{\nabla}{i} - e^* \mathbf{A} \right)^2 \psi + \alpha \psi + \beta |\psi|^2 \psi = 0 \quad (18.88a)$$

$$\nabla \times (\nabla \times \mathbf{A}) = \mathbf{j} = \frac{e^*}{2m^*} [\psi^* (-i \nabla \psi) + (i \nabla \psi^*) \psi] - \frac{e^{*2}}{m^*} \psi^* \psi \mathbf{A} \quad (18.88b)$$

At $T > T_c$ the free energy has a minimum at $|\psi| = 0$. This can be seen by looking at the potential part of the free energy (i.e. the last two terms), but below the critical temperature, $T < T_c$, the minimum moves to $|\psi| = \psi_0 = \sqrt{-\alpha/\beta} \neq 0$ (see Fig. 18.23 for the potential shape). This is a Bose–Einstein condensate state, because as $T \rightarrow 0$ all the Cooper pairs lie in the ground state with density given by $|\psi_0|^2 = n_s$. Writing $\psi(x) = \psi_0 + \phi(x)$, we consider only the ground state, i.e. $\phi(x) = 0$. Then the kinetic energy term in Eq. (18.88b)

can be neglected. Taking a time derivative or rotation of Eq. (18.88b) and using $\mathbf{B} = \nabla \times \mathbf{A}$, we obtain

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} = \frac{\partial}{\partial t} \left(\frac{\mathbf{j}}{\lambda^2} \right), \quad \Delta \mathbf{B} - \frac{1}{\lambda^2} \mathbf{B} = 0, \quad \lambda = \sqrt{\frac{m^*}{e^*{}^2 n_s}} \quad (18.89)$$

λ is referred to as the (London) penetration depth.

The first equation gives the supercurrent with perfect conductivity because any electromagnetic field produces accelerating current and the second the Meissner effect. Equations (18.89) mean that the photon has acquired a mass $\mu = \hbar/\lambda c$.

Consider a medium with boundary at $x = 0$ (see Fig. 18.19). The region ($x < 0$) is the vacuum in the magnetic field $\mathbf{B} = (0, B, 0)$ and that ($x > 0$) is superconducting. Equations (18.89) have a solution $\mathbf{B}(x > 0) = (0, B e^{-x/\lambda}, 0)$, which connects with $\mathbf{B}(x < 0)$. One sees that the magnetic field penetrates only to the penetration depth $\sim \lambda$, which is typically hundreds of atomic distances, a microscopic size. At the boundary, the density of the Cooper pair does not rise abruptly, but reaches its bulk value over the coherence length defined by $\xi = \hbar/\sqrt{m^*|\alpha|}$. This can be seen from an approximate solution ($\psi \sim e^{-x/\xi}$) to Eq. (18.88a), which can be obtained by setting $\mathbf{A} = 0$ and $\beta|\psi|^2 \ll \alpha$. The ratio $\kappa = \lambda/\xi$, referred to as the Ginzburg–Landau parameter, discriminates between type I and type II superconductors:

$$\kappa \equiv \lambda/\xi, \quad \text{type I } \kappa < 1/\sqrt{2}, \quad \text{type II } \kappa > 1/\sqrt{2} \quad (18.90)$$

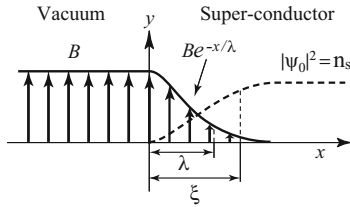


Figure 18.19 Meissner effect.

18.6

Unified Theory of the Electroweak Interaction

18.6.1

$SU(2) \times U(1)$ Gauge Theory

In Chap. 15, we learned that the weak interaction is of vector type, that it acts only on left-handed particles and that $W^\pm$ couple to isospin current, suggesting $W^\pm$

are members of an isospin triplet and the existence of a neutral boson W^0 . All this points to a gauge theory of $SU(2)$ symmetry. But there are two objections to be addressed. One is that the weak interaction shares part of the isospin current with the electromagnetic interaction and may not obey strict $SU(2)$ symmetry. The other is the short-range nature of the interaction, or equivalently the finite mass of $W^\pm$, which obviously contradicts the notion of $W^\pm$ being gauge particles. Deferring the mass problem until later, we will first address the fusion of the weak force with the electromagnetic force.

Since the electromagnetic current consists of isospin and remaining current [see Eqs. (15.98)], we consider the photon as a mixture of two gauge fields, one generated by the isospin and the other by a yet to be determined symmetry. We assume it to be $U(1)$ for simplicity. Then the electric charge consists of isospin and the $U(1)$ hypercharge, which are related to the electric charge by

$$Q = I_3 + \frac{Y}{2} \quad (18.91)$$

which defines the hypercharge of particles. The factor $1/2$ is introduced to make the formula the same as the Nishijima–Gell-Mann formula.¹³⁾ A new ingredient that did not exist in the isospin of the strong interaction is that we need to discriminate left- and right-handed particles as having different isospin and hypercharge. It is necessary because the isospin operator, the source of the weak force, acts only on left-handed particles and the hypercharge has to make a necessary adjustment to keep the charge Q independent of handedness. This means that we regard the left- and right-handed particles as different particles.

Let us consider a doublet consisting of (ν_e, e^-) for concreteness, though the following arguments are valid for any isospin doublet Ψ , including the quarks. A charged current made of a left-handed doublet

$$j_+^\mu = \bar{\nu}_L \gamma^\mu e_L = \bar{\Psi}_L \gamma^\mu \tau_+ \Psi_L \quad (18.92a)$$

$$\Psi = \begin{bmatrix} \nu \\ e \end{bmatrix}, \quad \Psi_L = \begin{bmatrix} \nu_L \\ e_L \end{bmatrix} \quad (18.92b)$$

couples with $W^\pm$, i.e. the weak interaction Lagrangian is assumed to be given by

$$\begin{aligned} -\mathcal{L}_{W1} &= \frac{g_w}{2} \bar{\Psi}_L \gamma^\mu (\boldsymbol{\tau} \cdot \mathbf{W}_\mu) \Psi_L \\ &= g_w \bar{\Psi}_L \gamma^\mu \left\{ \frac{1}{\sqrt{2}} (\tau_+ W_\mu^+ + \tau_- W_\mu^-) + \frac{1}{2} \tau_0 W_\mu^0 \right\} \Psi_L \end{aligned} \quad (18.93a)$$

$$W^\mp = \frac{1}{\sqrt{2}} (W_1 \pm W_2), \quad W^0 = W_3, \quad \tau_\pm = \frac{1}{2} (\tau_1 \pm i\tau_2) \quad (18.93b)$$

13) Note, the hypercharge thus defined is different from that of the strong interaction. In strong interactions, hypercharge is $Y = (B + S + C + B \cdots)/2$, where B is baryon number and $S, C, \dots$ are flavor numbers corresponding to strangeness, charmness, etc. and not sources of force fields.

where we have included the hypothetical W^0 to make the interaction consistent with $SU(2)$ gauge interaction [see Eq. (18.70)]. With the assumption that ν_R does not exist,¹⁴ e_R has no partner and constitutes a singlet. The hypercharge of ν_L , e_L , e_R is determined from Eq. (18.91)

$$Y(\nu_L) = Y(e_L) = -1, \quad Y(e_R) = -2 \quad^{15} \quad (18.94)$$

Denoting the gauge field associated with the hypercharge as B^μ , the Lagrangian satisfying independently $SU(2)$ and $U(1)$ symmetries [denoted as $SU(2) \times U(1)$] can be written as

$$\mathcal{L}_{W1} = \bar{\Psi}_L i \gamma^\mu D_\mu \Psi_L + \bar{e}_R i \gamma^\mu D_\mu e_R - \frac{1}{4} F_{\mu\nu} \cdot F^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \quad (18.95a)$$

$$D_\mu = 1\partial_\mu + i g_w \mathbf{W}_\mu \cdot \mathbf{t} + i(g_B/2) B_\mu Y \quad (18.95b)$$

The factor $1/2$ for g_B is for later convenience and has no special meaning. Since Ψ_L and e_R have different isospin and hypercharge, the isospin operator $\mathbf{t}$ and the hypercharge operator Y have different eigenvalues depending on the operand. For instance, $\mathbf{t}e_R = 0$, $Y\Psi_L = -\Psi_L$, $Ye_R = -2e_R$. Notice that the Lagrangian has no mass term. If it does, the left- and right-handed components mix and no unique quantum numbers can be assigned independently (see arguments in Sect. 4.3.5). In other words, in a world where both left- and right-handed particles respect independently gauge invariance (chiral invariance), the fermion cannot have mass. Let us close our eyes to massless fermions for a moment and proceed.

Extraction of the $\mathbf{W}_\mu^\pm$ part of the Lagrangian Eq. (18.95a) reproduces Eq. (18.93a). Extraction of terms including W_μ^0 and B_μ gives

$$\begin{aligned} -\mathcal{L}_{NC} &= \frac{g_w}{2} (\bar{\nu}_L \gamma^\mu \nu_L - \bar{e}_L \gamma^\mu e_L) W_\mu^0 \\ &\quad + \frac{g_B}{2} [-(\bar{\nu}_L \gamma^\mu \nu_L + \bar{e}_L \gamma^\mu e_L) - 2\bar{e}_R \gamma^\mu e_R] B_\mu \\ &= \frac{1}{2} (g_w W_\mu^0 - g_B B_\mu) (\bar{\nu}_L \gamma^\mu \nu_L) \\ &\quad - \frac{1}{2} (g_w W_\mu^0 + g_B B_\mu) (\bar{e}_L \gamma^\mu e_L) - g_B B_\mu \bar{e}_R \gamma^\mu e_R \end{aligned} \quad (18.96)$$

where (-1) and (-2) in front of $\bar{\nu}_L \gamma^\mu \nu_L$ and $\bar{e}_R \gamma^\mu e_R$ in the first line reflect their hypercharge. Since both gauge fields B and W^0 couple to the same current, $(\bar{\nu}_L \gamma^\mu \nu_L$ and $\bar{e}_{L(R)} \gamma^\mu e_{L(R)})$, they mix through fermion loops (Fig. 18.20). As the neutrino

14) Since the discovery of neutrino oscillation, this assumption is no longer true. However, the following discussion remains valid. The only necessary modification is to include ν_R in the Lagrangian, which is rather academic under the present circumstances. The assignment $I(e_R) = 0$ holds regardless of the neutrino mass.

15) For quarks, (u_L, d_L) couple with W , constituting an isodoublet, but u_R, d_R are isosinglets because they do not. Then $Y(u_L) = Y(d_L) = 1/3$, $Y(u_R) = 4/3$, $Y(d_R) = -2/3$. The same rule applies to (c, s) and (t, b) quarks.

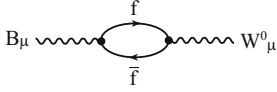


Figure 18.20 Mixing between B and W^0 occurs when both fields couple to the same fermion pair.

should not couple to the electromagnetic field, we define the combination that couples to the neutrino as Z_μ and its orthogonal combination as the electromagnetic field A_μ :

$$Z_\mu = \frac{1}{\sqrt{g_w^2 + g_B^2}} (g_w W_\mu^0 - g_B B_\mu) \equiv \cos \theta_W W_\mu^0 - \sin \theta_W B_\mu \quad (18.97a)$$

$$A_\mu = \frac{1}{\sqrt{g_w^2 + g_B^2}} (g_B W_\mu^0 + g_w B_\mu) \equiv \sin \theta_W W_\mu^0 + \cos \theta_W B_\mu \quad (18.97b)$$

which defines the Weinberg angle θ_W . Inserting Z_μ , A_μ defined in Eq. (18.97) back in Eq. (18.96), and rewriting their coefficients using $Q = I_3 + Y/2$, we obtain

$$\begin{aligned} -\mathcal{L}_{\text{NC}} &= g_w \sin \theta_W (\bar{e}_L \gamma^\mu Q e_L + \bar{e}_R \gamma^\mu Q e_R) A_\mu + \sqrt{g_w^2 + g_B^2} \\ &\quad \times [I_3 (\bar{\nu}_L \gamma^\mu \nu_L + \bar{e}_L \gamma^\mu e_L) - Q \sin^2 \theta_W (\bar{e}_L \gamma^\mu e_L + \bar{e}_R \gamma^\mu e_R)] Z_\mu \\ &= Q e \bar{\Psi} \gamma^\mu A_\mu \Psi + g_z \bar{\Psi} \gamma^\mu Z_\mu (I_{3L} - Q \sin^2 \theta_W) \Psi \end{aligned} \quad (18.98a)$$

$$e = g_w \sin \theta_W, \quad g_z = \sqrt{g_w^2 + g_B^2} = \frac{e}{\sin \theta_W \cos \theta_W} \quad (18.98b)$$

where I_{3L} means that it acts only on the left-handed part of the field. The first term is the electromagnetic interaction and the second a new weak neutral current interaction that couples to Z_μ . Because of mixing between $SU(2)$ and $U(1)$, the current that couples to the weak neutral boson Z_μ is a mixture of the left-handed neutral current and the electromagnetic current. Hereafter, we restrict the use of the word neutral current to that which couples exclusively to Z_μ unless otherwise stated.

We have, starting from the chiral world, where left- and right-handed particles are independent, succeeded in extracting a left-right symmetric current, which can be identified as the electromagnetic current. It is a gauge theory based on $SU(2) \times U(1)$ and contains the neutral weak gauge fields Z^0 as well as the electromagnetic field. Note, however, that the above formalism is valid only for massless gauge fields as well as for massless fermions. Adding mass terms by hand makes the formalism phenomenologically usable in many applications today. Using the above Lagrangian and equating the $W^\pm$ exchange transition with the four-Fermi interaction of the weak interaction, we obtain

$$\frac{G_F}{\sqrt{2}} = \frac{g_w^2}{8m_W^2} = \frac{e^2}{8m_W^2 \sin^2 \theta_W^2} \quad (18.99)$$

$\sin^2 \theta_W$ is referred to as the Weinberg angle, and if it can be determined by experiments, the W mass would also be determined. A shortcoming of the theory is that

the mass is put in by hand, thereby destroying the gauge invariance. No reliable higher order calculations are possible either.

18.6.2

Spontaneous Symmetry Breaking

We have learned in the previous section that we could construct a satisfactory theory, at least phenomenologically, by putting the mass terms in by hand. But by doing so, the gauge invariance is broken, making it unrenormalizable, with resultant loss of predictive power. Here we demonstrate that despite apparent breakdown of the symmetry, there are cases in which the symmetry is preserved deep in the roots of the mathematical formalism. We might well say the symmetry is not broken but hidden.

An equation with some symmetry generally has solutions that respect the same symmetry, but those solutions are not necessarily stable. Consider, for instance, a stick standing upright on a plane. Apply pressure straight down from above (Fig. 18.21). Within the elastic limit it is compressed, but beyond a certain pressure it is bent and stays that way. The solution is supposed to have rotational symmetry around an axis, but the realized stable solution obviously breaks it. This is a case in which the equation respects the symmetry but its solution does not. A remnant of the symmetry is there in that any direction can be chosen and the realized state has the same energy regardless of the chosen direction. Namely, the ground states are infinitely degenerate.

Another example where field theory can be applied is to a ferromagnetic substance. Ferromagnetism is produced as a result of electron spins all being aligned. At high temperature, the direction of the spins is random due to thermal motion, chaos in a sense, but since there is no special direction, the rotational symmetry is respected. When an external magnetic field is applied, all the spins are aligned and the rotational symmetry is broken. This is the case that the symmetry is broken by introducing a symmetry-breaking perturbation by hand in the system. The external field is not rotationally symmetric, which is the origin of the symmetry breaking of the realized state. However, even without a symmetry-breaking perturbation, the spins can be aligned. This happens below a critical point, the Curie temperature. Below it, spin–spin interaction, which by itself is rotationally symmetric, wins over thermal turbulence and the spins align themselves (Fig. 18.22a). The chosen direction is quite arbitrary, because the states are infinitely degenerate, but once a direction is chosen, all the neighboring spins follow the initiator's trigger and align

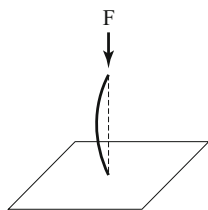


Figure 18.21 A stick under pressure. The bent position exhibits a ground state that breaks the rotational symmetry and infinitely degenerate.

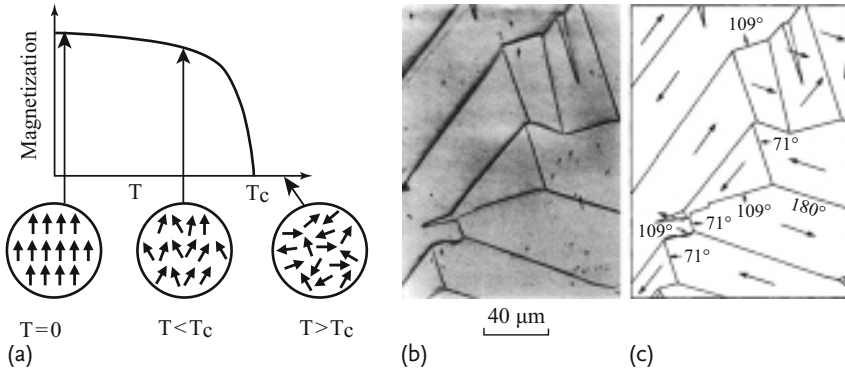


Figure 18.22 (a) Magnetization pattern as a function of temperature. Below the Curie temperature (T_c), spins are ordered and magnetization appears. (b) Domain patterns observed

on a $\{110\}$ surface away from a crystal edge. (c) Inferred domain magnetization directions. Figures taken from [309].

along the same direction; the chosen direction is frozen. No local perturbations can change it because it takes a vast amount of energy (binding energy per spin times the number of spins). Such processes are observed commonly in condensed matter physics. Common features of them are

1. There is some critical temperature T_c .
2. A configuration with good symmetry becomes unstable below the critical temperature and changes to another stable state where the original symmetry is broken.
3. Realized stable states (ground states) are continuously (infinitely) degenerate.

When the process happens, it is called spontaneous symmetry breaking. The chosen direction need not be the same in one place as in another if they are far apart. Because of this, domains of differing spin direction are formed (Fig. 18.22b). A little man living in such a world would never imagine that the equation that governs the direction of spins respects rotational symmetry. Such a macroscopic metamorphosis at a critical temperature has been known quite a while as a phase transition and its formalism is well developed in condensed matter physics. It was Nambu who introduced the idea of spontaneous symmetry breaking into particle physics. He investigated superconductivity from the viewpoint of gauge symmetry breakdown and applied it to particle physics [287, 289, 290].

The mathematical formalism of spontaneous symmetry breaking goes like this. Consider a self-interacting complex scalar field whose expectation value is taken to be the magnetization we talked about. Its Lagrangian has the form¹⁶⁾

16) The Hamiltonian that can be derived from Eq. (18.100) is referred to as the Gibbs free energy and ϕ is an order parameter in the terminology of condensed matter physics.

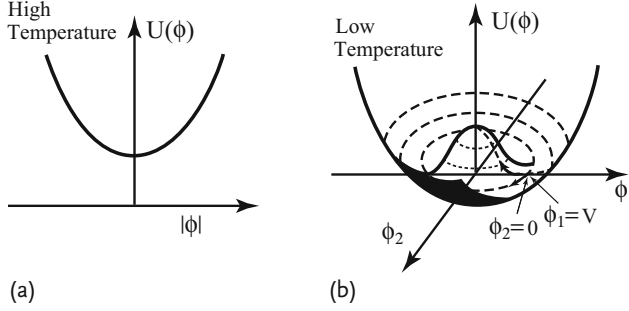


Figure 18.23 (a) At high temperature $T > T_c$, vacuum is where $\phi = 0$. (b) At low temperature $T < T_c$, $|\phi| = 0$ state is unstable, but the state with lowest energy is infinitely degenerate.

$$\mathcal{L}_G = \partial_\mu \phi \partial^\mu \phi - V(\phi) \quad (18.100a)$$

$$V(\phi) = \lambda \left(\phi^\dagger \phi + \frac{\mu^2}{2\lambda} \right)^2 = \mu^2 (\phi^\dagger \phi) + \lambda (\phi^\dagger \phi)^2 + V(0), \quad (18.100b)$$

$$\lambda > 0 \quad \mu^2 = a(T - T_c) \quad (18.100c)$$

The constant term $V(0)$ can be chosen arbitrarily, because we are only dealing with the energy difference or excitations from the ground state.¹⁷⁾ The Lagrangian is invariant under gauge transformation

$$\phi \rightarrow \phi' = e^{-i\alpha} \phi, \quad \phi^\dagger \rightarrow \phi'^\dagger = \phi^\dagger e^{i\alpha} \quad (18.101)$$

Note, when $|\phi|$ is to be interpreted as the magnetization, the rotational symmetry is in real space. Here we use rotational symmetry in the complex internal space because we want to discuss gauge symmetry. From now on we talk about the complex scalar field in particle physics and the field of magnetization interchangeably. When we talk about magnetization, the lowest energy state is referred to as the ground state.

At high temperature ($T > T_c$), the potential is of the form depicted in Fig. 18.23a and the lowest energy (the ground state) is realized with $|\phi| = 0$ or zero magnetization. At low temperature ($T < T_c$), however, $\mu^2 < 0$ and the shape of the potential is like a Mexican hat or the bottom of a wine bottle (Fig. 18.23b). $\phi = 0$ is an unstable solution, and the ground state is at $|\phi| = v/\sqrt{2} = \sqrt{-\mu^2/2\lambda}$; the magnetization is finite. The ground state can be anywhere that satisfies $|\phi| = v/\sqrt{2}$ along a circle on the bottom of the wine bottle. Putting $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$ we choose the ground state at

$$\phi_1 = v, \quad \phi_2 = 0, \quad v = \sqrt{-\frac{\mu^2}{\lambda}} \quad (18.102)$$

¹⁷⁾ $V(0)$ is considered as the vacuum energy. Its absolute value becomes important if gravity matters.

As the ground state is the vacuum in field theory, the field ϕ has a finite vacuum expectation value:

$$\langle \phi \rangle = \langle 0 | \phi | 0 \rangle = v / \sqrt{2} \quad (18.103)$$

Since we always measure excited energy (particle creation) from the ground state (vacuum), and the new vacuum is at $\phi_1 = v$, we define a new field

$$\phi'_1 = \phi_1 - v \quad (18.104)$$

The new field has the vacuum expectation value

$$\langle \phi'_1 \rangle = \langle \phi_2 \rangle = 0 \quad (18.105)$$

The field described in terms of ϕ'_1 , ϕ_2 does not respect the original gauge (rotational) symmetry. This is a case of spontaneous symmetry breaking. In field theory, once the vacuum has been chosen, it is fixed, because there are infinitely many degrees of freedom and no matter how small the microscopic energy needed to change the configuration, if multiplied by infinity, it is infinite. In the case of condensed matter physics, the number of degrees of freedom is large but finite (of the order of the Avogadro number) and the ground state can be changed by heating or cooling the bulk of the matter. The spontaneously broken symmetry in field theory is due to the vacuum that surrounds us, and we are the little men in the ferromagnet we talked about. We cannot change our vacuum unless we heat the whole universe! We are in one of the domains. Our whole universe is most likely in one domain.¹⁸⁾

Note that spontaneous symmetry breaking does not break the symmetry mathematically. Observed phenomena were simply reexpressed in terms of ϕ'_1 and ϕ_2 . Mathematically, no symmetry-breaking terms are introduced. All the phenomena can equally be described in terms of original variables that respect the symmetry, though in this case physical interpretation is much harder.

Nambu–Goldstone Boson What does the new field represent? In terms of ϕ'_1 , ϕ_2 , which we rewrite as ϕ_1 , ϕ_2 , the Lagrangian is expressed as

$$\begin{aligned} \mathcal{L}_G = & \frac{1}{2} [\partial_\mu \phi_1 \partial^\mu \phi_1 - (-2\mu^2) \phi_1^2] + \frac{1}{2} [\partial_\mu \phi_2 \partial^\mu \phi_2] \\ & + \left[\lambda v \phi_1 |\phi|^2 + \frac{\lambda}{4} |\phi|^2 \right] \end{aligned} \quad (18.106)$$

The first term represents a massive scalar field with mass $\sqrt{-2\mu^2}$, or the magnetization we talked about, the second is a massless scalar field and the third is the

18) The boundaries of domains are known as “defects” and they are commonly observed in condensed matter physics. There are conjectures that defects may have been formed at the big bang and their remnants may be observable as monopoles, cosmic strings, etc. If they are observed, the standard cosmic model has to be modified.

interaction between the two. A massless scalar boson appears if a continuous symmetry is spontaneously broken with infinitely degenerate ground (vacuum) states, a generalization of Nambu's proposal, which is referred to as the Goldstone theorem [180]. The massless boson is referred to as the Nambu–Goldstone boson. Does such a particle exist in a ferromagnet? The answer is yes, and it is called a magnon.

Although changing the chosen state to another is difficult, it is possible to give a small perturbation locally. Then the spin–spin interaction, which was the cause of the spontaneously broken state, transmits the perturbation to neighboring spins and the perturbation propagates, making a wave (Fig. 18.24). The magnon is a spin

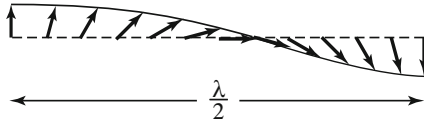


Figure 18.24 When the magnetization is aligned, a spin wave, a deviation from the ordered state, appears. In the long-wavelength limit, the ordered ground state is recovered.

This means that its quantum, called the magnon, is a massless particle commonly referred to as the Nambu–Goldstone boson.

wave. That its long-wavelength limit recovers the static ground state, i.e. $\omega = 0$ as $\lambda \rightarrow \infty$ or $k \rightarrow 0$, means the mass of the quantum is zero. Why does such an extra boson appear? Originally there were two fields to represent a continuous symmetry. After condensation, all particles (the field's quanta) are ordered and occupy the same state. There is no need of two fields to describe an ordered state; one is enough. Its quantum has to be massive to provide some stability to the state. The residual field is not constrained and its quantum is massless. One sees in Fig. 18.23b that in the ϕ_1 direction it is at the minimum of a potential and takes some effort to move it (excite the field, i.e. create a particle), but ϕ_2 lies on a flat circle, and needs no force to move it.

The question is: do we have Goldstone bosons as an indication of symmetry breaking?

18.6.3

Higgs Mechanism

In nature, no massless Goldstone bosons are known and there was difficulty in accepting the notion of spontaneously broken symmetry. However, there is a loophole in Goldstone's theorem and the Goldstone boson does not appear when the symmetry is associated with a long-range force like that created by the gauge field. In the absence of a long-range force, the wave in the Goldstone mode can propagate but the force's action does not reach beyond a few neighboring particles. If the force is of long range, particles far away are affected simultaneously, and they act coherently, constituting a longitudinal wave. In a plasma, as a result of the Coulomb force, a longitudinal wave called plasma oscillation is formed. Its action

screens the Coulomb force, converting it to a short-range force. In the language of field theory, the quantum of the plasma oscillations has mass, namely the photon has acquired mass and has become a plasmon. Higgs and others [131, 191, 205] showed that when the symmetry is broken in association with the gauge field, the Goldstone boson does not appear but instead the gauge boson acquires mass. We will see how.

Mass of the Gauge Boson When a gauge field is introduced, the Lagrangian containing the self-interacting scalar field (hereafter referred to as the Higgs field) and the gauge field is expressed as

$$\mathcal{L}_H = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (18.107a)$$

$$D_\mu = \partial_\mu + i q A_\mu, \quad V(\phi) = \mu^2 (\phi^\dagger \phi) + \lambda (\phi^\dagger \phi)^2 \quad (18.107b)$$

Let us see what happens after the gauge symmetry is spontaneously broken. It is convenient to decompose ϕ not linearly like $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$ but exponentially

$$\phi = \exp[i(\phi_2/\nu)] \frac{\phi_1}{\sqrt{2}} \rightarrow \exp[i(\phi_2/\nu)] \frac{\nu + \phi'_1}{\sqrt{2}} \quad (18.108)$$

Exponential Decomposition

Mathematically it is always possible to transform into this form as long as the number of degrees of freedom does not change. Then the difference of ϕ_2 in the two expressions is a matter of physical interpretation. ϕ_2 in this form is called a phase field. If expanded in powers of ϕ_2/ν , it is the same as the expansion $\phi = (\nu + \phi_1 + i\phi_2)/\sqrt{2}$ if higher order terms are neglected. Namely, as long as the excitation of the field is small, where the particle concept of the field is valid anyway, the difference is small. Remember, the field quantization is applied to harmonic oscillators or using up to quadratic terms in the Taylor expansion of the potential around the minimum and the rest is treated as interaction terms. The difference appears as a global feature of the field, which reflects the collective effects of the interaction. For instance, a local plane wave could globally be a part of spherical wave at far distance or circulating wave whose radius is large. In the spontaneous symmetry breaking, we are discussing what will happen if we extend our approximation up to the quartic power of the Taylor expansion. In Fig. 18.23b ϕ_2 is constrained to lie on a circle and has periodicity, hence its name. But locally, as long as it remains in the neighborhood of $\phi_2 = 0$, the linear treatment is equally valid.

Problem 18.5

Show that for $\phi'_1, \phi_2 \ll v$, Eq. (18.108) is equivalent to

$$\phi = \frac{1}{\sqrt{2}}(v + \phi'_1 + i\phi_2) \quad (18.109)$$

If we rename ϕ'_1 as ϕ_1 and replace the gauge field by

$$B_\mu = A_\mu + \frac{1}{qv} \partial_\mu \phi_2 \quad (18.110)$$

the Lagrangian takes the form

$$\begin{aligned} \mathcal{L}_H = & \frac{1}{2} [D_\mu \phi_1 D^\mu \phi_1 - (-2\mu^2)\phi_1^2] - \lambda v \phi_1^3 - \frac{\lambda}{4} \phi_1^4 \\ & - \frac{1}{4} F_{B\mu\nu} F_B^{\mu\nu} + \frac{(qv)^2}{2} B_\mu B^\mu \\ F_{B\mu\nu} = & \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \quad (18.111)$$

Problem 18.6

Derive Eq. (18.111).

In this expression, there is no Goldstone boson. Notice, the second line represents a massive vector field [see Eq. (5.22)]. ϕ_2 is absorbed as the third (longitudinal) component of the gauge field, which has mass $m_B = qv$. Note that the number of degrees of freedom is the same as before. Originally, there were one complex scalar field or two real scalar fields and a massless gauge vector boson, which has two degrees of freedom. After symmetry breaking, there are one real scalar field and a massive vector boson, which has three degrees of freedom. It is often said that the gauge boson became fat by eating a ghost. This is the process that a component of the Higgs field is converted to the longitudinal component of the gauge field to make it massive; it is called the Higgs mechanism.

In order to choose a vacuum from an infinite number of degenerate states, we have used the freedom of gauge transformation Eq. (18.101) as expressed in Eq. (18.108)

$$\phi \rightarrow e^{-i\alpha} \phi \iff \phi \rightarrow \frac{\phi_1}{\sqrt{2}} e^{i(\phi_2/v)} \quad (18.112)$$

In other words, the vacuum has been fixed by a special kind of gauge transformation. The gauge thus chosen is commonly referred to as the unitary gauge.

A Hamiltonian that can be extracted from a nonrelativistic version of Eq. (18.107) is a well-known formula in condensed matter physics and is referred to as the Ginzburg–Landau free energy for superconductivity (see the boxed paragraph on

p. 769 Confinement equation. 18.85). ϕ denotes the Cooper pair wave function. Below the critical temperature, the Cooper pair is mass produced (condensation) and matter is transformed to the superconducting phase. A superconductive medium repels the magnetic field by producing eddy currents to compensate the intruding magnetic perfectly (perfect diamagnetism referred to as the Meissner effect) except in a thin skin depth. This is due to condensed Cooper pairs acting coherently as a macroscopic quantum fluid. In the language of field theory, the photon (gauge field) has acquired mass in the superconductor and the electromagnetic force has changed to a short-range ($\sim$ skin depth) force with the range given by the Compton wavelength of its mass. In the superconductor, the gauge invariance is broken spontaneously. Adopting the Higgs mechanism is to consider that the vacuum that surrounds us is in the superconductive phase. The temperature of the phase transition is of the same order as the vacuum expectation value v .

What is the meaning of the vacuum expectation value? In magnetism, it means a finite magnetization in the ground state where all the spins are aligned, a situation in which a great number of particles occupy the same state, which is referred to as (Bose–Einstein) condensation in condensed matter physics. When the field has a finite vacuum expectation value, it means condensation to the ground state. All the particles act coherently, or as a macroscopic quantum state. In this state the characteristic field behavior is more that of a wave than that of particles. In the Higgs mechanism, the condensate is the Higgs field, an equivalent of a sea of Cooper pairs acting as a superfluid, which screens the force, converting it to a short-range force.

Fermion Mass We have explained the finite mass of the gauge boson as a result of spontaneous breakdown of the gauge symmetry. What about the fermion mass? The vanishing of the fermion mass was the result of chiral invariance. We shall show that the fermion mass can also be generated as a result of spontaneous symmetry breaking. Let us assume that the Higgs field has an interaction with the fermion field. Adopting a Yukawa-type interaction and denoting the fermion field as ψ , we can write the Lagrangian in Lorentz-invariant form

$$-\mathcal{L}_{\text{Yukawa}} = f \bar{\psi}_L \psi_R \phi + h.c. \quad (18.113)$$

The chiral invariance can be maintained by assigning suitable hypercharge to ψ_L , ψ_R and ϕ , namely to satisfy $-Y_{\psi_L} + Y_{\psi_R} + Y_\phi = 0$. When the symmetry is broken spontaneously, the Higgs field acquires vacuum expectation value, i.e. $\phi \rightarrow (v + \phi_1)/\sqrt{2}$,

$$\begin{aligned} -\mathcal{L}_{\text{Yukawa}} &= \left(\frac{fv}{\sqrt{2}} + \frac{f}{\sqrt{2}} \phi_1 \right) (\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R) \\ &= m_f \bar{\psi} \psi + \frac{m_f}{v} \bar{\psi} \psi \phi_1, \quad m_f = \frac{fv}{\sqrt{2}} \end{aligned} \quad (18.114)$$

Thus we have created the fermion mass term, but as a by-product we have also induced a fermion–Higgs interaction. There are no principles or guides to constrain fermion–Higgs couplings, hence the fermion mass is an additional arbitrary parameter. This applies also to the mass of the Higgs particle itself ($m_{\text{H}}^2 = -2\mu^2 = 2\lambda v^2$), which is proportional to the strength of the self-coupling constant λ and has to be determined experimentally.

What we have learned so far is summarized as follows. It is possible to generate mass terms in the gauge-invariant Lagrangian consisting of massless fermions and gauge bosons (referred to as the gauge sector). The mass term was generated by the interaction of the Higgs field with the gauge field as a result of spontaneous symmetry breaking. Its Lagrangian (Higgs sector) is also gauge invariant, as is seen from Eq. (18.107). Extracting the mass term only from the Higgs sector and adding it to the gauge sector violates the gauge invariance. But if we take into account the whole Higgs sector as well as the mass term, the gauge invariance is preserved. If the Higgs mechanism is right, we have to include the Higgs field in the calculation of higher order terms of gauge particle interactions. Above all, the Higgs particle itself has to exist.

18.6.4

Glashow–Weinberg–Salam Electroweak Theory

We are now ready to formulate a gauge-invariant electroweak theory. We discuss only (ν_e, e^-) doublets as before. Other doublets can be treated similarly.

Higgs Doublet We have succeeded in constructing a gauge-invariant theory with finite mass of the gauge boson conceptually. What remains to be done is to incorporate the Higgs mechanism into the weak interaction framework based on $SU(2) \times U(1)$. In order to do this, we need at least four Higgs fields. Three would disappear, giving mass to $W^\pm$ and Z , and the remaining one would become a massive Higgs field. There are many ways to incorporate four Higgs fields, but the simplest one, which Weinberg and Salam (WS) adopted, is to assume the existence of an isospin doublet of a complex scalar field and its charge conjugate. We express them as

$$\Phi = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix}, \quad \Phi^c = -i\tau_2 \Phi^\dagger = \begin{bmatrix} -\phi^{0+} \\ \phi^- \end{bmatrix} \quad (18.115)$$

Then using the Nishijima–Gell-Mann law, we have $Y(\Phi), Y(\Phi^c) = 1, -1$. When they acquire vacuum expectation values, we write them as

$$\Phi \rightarrow \begin{bmatrix} 0 \\ \frac{(v+\varphi)}{\sqrt{2}} \end{bmatrix}, \quad \Phi^c \rightarrow \begin{bmatrix} -\frac{(v+\varphi)}{\sqrt{2}} \\ 0 \end{bmatrix} \quad (18.116)$$

The Lorentz-invariant Yukawa interaction respecting $SU(2) \times U(1)$ symmetry is given by

$$-\mathcal{L}_{\text{mass}} = f_e \left[\bar{e}_R (\Phi^\dagger \Psi_L) + (\bar{\Psi}_L \Phi) e_R \right] \quad (18.117)$$

As the neutrino mass is assumed to be zero, we do not include its mass term here, but if one wants to give mass to the $I_3 = +1/2$ member of the doublets, one uses Φ^c instead of Φ .

$SU(2) \times U(1)$ The total Lagrangian for (ν_e, e^-) and $W^\pm, Z, \gamma$ and the Higgs respecting $SU(2) \times U(1)$ symmetry can be expressed as

$$\begin{aligned} \mathcal{L}_{\text{GWS}} = & \bar{\Psi}_L i \gamma^\mu D_\mu \Psi_L + \bar{e}_R i \gamma^\mu D_\mu e_R - \frac{1}{4} \mathbf{F}_{\mu\nu} \cdot \mathbf{F}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\ & + (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\phi) - f_e [\bar{e}_R (\Phi^\dagger \Psi_L) + (\bar{\Psi}_L \Phi) e_R] \end{aligned} \quad (18.118a)$$

$$D_\mu = \partial_\mu + i g_w \mathbf{W}_\mu \cdot \mathbf{t} + i (g_B/2) B_\mu \cdot Y \quad (18.118b)$$

$$\mathbf{F}_{\mu\nu} = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu - g_w \mathbf{W}_\mu \times \mathbf{W}_\nu \quad (18.118c)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (18.118d)$$

$$V(\phi) = \lambda \left(|\Phi|^2 + \frac{\mu^2}{2\lambda} \right)^2 \quad (18.118e)$$

The gauge sector is the first line of the Lagrangian, and the Higgs sector the second line. We need to rewrite $\mathbf{W}_\mu^0, B_\mu$ in terms of Z_μ, A_μ . When the Higgs field acquires a vacuum expectation value, one can obtain the Lagrangian of GWS theory. It is equivalent to Eq. (18.95) derived previously plus the mass term and the Higgs field.

Mass of W and Z The mass of W is already given in Eq. (18.99), but to derive the mass of Z , we need to know the Weinberg mixing angle and the Higgs field vacuum expectation value. They can be obtained by substituting

$$\Phi = \begin{bmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{bmatrix} \quad (18.119)$$

into the kinematic energy term (the first term in the second line) of the Higgs sector in Eq. (18.118) to give

$$m_W = \frac{g_w v}{2}, \quad m_Z = \frac{g_Z v}{2} = \frac{g_Z}{g_w} m_W = \frac{m_W}{\cos \theta_W} \quad (18.120)$$

The mass of W is also related to the Fermi coupling constant by Eq. (18.99). From this, if one can obtain the value of $\sin \theta_W$ experimentally, m_W, m_Z, v can be cal-

culated. Using experimentally observed value $\sin \theta_W = 0.2231$ [311]

$$m_W^2 = \frac{\sqrt{2}e^2}{8G_F \sin^2 \theta_W} = \left(\frac{37.3 \text{ GeV}}{\sin \theta_W} \right)^2 = (80.4 \text{ GeV})^2 \quad (18.121a)$$

$$m_Z = \frac{m_W}{\cos \theta_W} = 91.19 \text{ GeV} \quad (18.121b)$$

$$\nu = 1/\sqrt{\sqrt{2}G_F} = 250 \text{ GeV} \quad (18.121c)$$

Problem 18.7

Derive m_W , m_Z from Eq. (18.118).

18.6.5

Summary of GWS Theory

We have reached an amazing conclusion: that the mass is not an inherent property of a particle but one that has been acquired as a result of environmental change. The GWS theory is constructed on the gauge principle based on $SU(2) \times U(1)$ symmetry that is spontaneously broken. Requiring the existence of the Higgs field, which is responsible for the symmetry breaking, allowed the finite mass of the gauge bosons as well as that of fermions to be incorporated in the theory without breaking the gauge invariance. It is a unified theory of the electromagnetic and weak interactions¹⁹⁾ and hence is referred to as the electroweak theory. Renormalizability of the theory was proved by 't Hooft in 1971 [359, 360]. Therefore, it is mathematically closed, self-consistent and has strong predictive power. All dynamical phenomena caused by the electroweak force are calculable to arbitrary accuracy, at least in principle. The theory succeeded in reproducing all phenomena known at that time and so its proof resided in the prediction of new phenomena.

When Was the GWS Model Accepted as the Standard Model?

It is interesting to note that such a revolutionary theory was barely noticed upon its publication. According to citation records, the number of citations was 0 in 1967, 0 in 1968, 0 in 1969, 1 in 1970, 4 in 1971, 64 in 1972, 162 in 1973 and 226 in 1974 [102]. This includes Weinberg's own citations. Why was such an outstanding paper not noticed? Because it was uncertain if the theory was renormalizable. Weinberg and Salam guessed that it would be if the symmetry was broken spontaneously but left it unsolved. It was clear that the sudden increase of the citation is due to 't Hooft's paper [359, 360], which

¹⁹⁾ It is not unified in the strict sense because $SU(2)$ and $U(1)$ are not subgroups of a bigger group. The two coupling constants of $SU(2)$ and $U(1)$ are connected through Weinberg mixing, but there are still two coupling constants. It is unified in the sense that we can formulate two forces in a consistent and unified way.

provided a proof of renormalizability. Since then, Weinberg's seminal paper on the subject [383] boasts top citing to this day (6815 as of July 2009), surpassed only by the *Review of Particle Physics*, which is a compilation of all the world data in high-energy physics (see [348]). The first experimental evidence supporting the GWS model was the discovery of neutral current reactions in 1973 [199, 200], but what really established its validity beyond doubt was the discovery of W and Z in 1983 with exactly the mass predicted [372–375]. But why did it take as long as 16 years to prove the model? I would argue that the year the model was accepted by the community was 1978. By then, the majority already believed it, but there was one piece of counter evidence against it. Some atomic parity-violation experiments did not confirm the prediction that is now history. Convincing evidence to show that the GWS model, not the atomic experiment, was right was presented at ICHEP 78 (International Conference for High Energy Physics) by the SLAC group. It was the detection of parity violation in polarized electron scattering off a deuteron target [326, 327]. In addition, a compilation of measured Weinberg angles from a variety of experiments were shown to agree with each other [41] and singled the Weinberg–Salam model out from others. Ample evidence for QCD, including precision experiments on deep inelastic scattering, was also presented at the conference. The next year, Glashow, Weinberg and Salam were awarded the Nobel prize.

19 Epilogue

The first experimental evidence for GWS theory were the discovery of neutral current reactions [199, 200]. The fact that all the neutral current reactions can be reproduced with one universal parameter, the Weinberg angle, convinced people the rightness of the theory. But, the definitive evidence was provided by the discovery of W and Z [372–375] having exactly the right mass as predicted. Since then many experiments, for example, SPEAR at SLAC (7.4 GeV, 1972–1990), PETRA at DESY (40 GeV, 1978–1990) and TORISTAN at KEK (60 GeV, 1987–1996) confirmed the predictions of GWS theory. Above all, LEP at CERN (100 ~ 200 GeV, 1989–2000), in which the center of mass (CM) energy was set at $\sqrt{s} = m_Z = 92$ GeV, made a number of high-precision measurements confirming the higher order predictions of the electroweak theory. They could be compared with the Lamb shift experiment in QED in establishing the validity of the Standard Model to high precision (see for instance [184, 365]).

Among hadron colliders, the $S\bar{p}pS$ (Super Proton Synchrotron, $\sqrt{s} = 540 \rightarrow 630$ GeV, 1979–1990) discovered $W^\pm$ and Z in 1983. The Tevatron, which operated 1983–2009, discovered the top, the last quark in the Standard Model. Both of them provided a variety of hadron jet data confirming many predictions of QCD. B factories (PEPII at SLAC, $E = 3.1 + 9$ GeV, 1998–2008, and KEKB at KEK, $E = 3.5 + 8$ GeV, 1999 onwards) confirmed that CP-violation data fit well with the Kobayashi–Maskawa framework. The only major remaining issue in the Standard Model is the discovery of the Higgs particle, which is expected to be found soon when the LHC starts operation.¹⁾

All the known data so far, mainly obtained at electron–positron and hadron colliders, have been consistent with the Standard Model. The only evidence pointing beyond the Standard Model is neutrino oscillation, but whether this phenomenon is outside of GWS theory is debatable and depends on the exact definition of the Standard Model. This only shows how perfect the Standard Model is, but at the same time it is a disappointment in the sense that progress towards new physics has been very slow. Therefore we will close this book by making a few comments on what lies beyond the Standard Model.

1) The LHC is a p – p colliding accelerator at $\sqrt{s} = 14$ TeV and luminosity L of 10^{33} – 10^{34} $\text{cm}^{-2}\text{s}^{-1}$. It began operation in 2009. More discoveries mentioned later in this chapter are also expected.

19.1

Completing the Picture

Generation Puzzle and Mass Hierarchy Despite its great success, one caveat of the Standard Model is that it is powerless in explaining the generation (or family) structure. The three generations of family (u, d, ν_e, e^-), (c, s, ν_μ, μ^-), (t, b, ν_τ, τ^-) seem to share the same properties except their mass. The Standard Model cannot tell why there are three and only three generations. Their mass is also a mystery. Their mass hierarchical structure shown in Fig. 19.1 suggests a certain pattern, but in the Standard Model all the mass values are parameters to be adjusted. The mass is generated by the Higgs field and is expressed as $m_i = f_i v / \sqrt{2}$ GeV where $v = 1/\sqrt{\sqrt{2}G_F} = 250$ GeV is the vacuum expectation value of the Higgs field and f_i is its coupling strength [see Eqs. (18.114)(18.121c)]. There are no criteria to determine or constrain the strength of the Higgs coupling hence they are arbitrary.

Higgs Mechanism Another caveat is that the Standard Model does not have much predictive power for phenomena involving mass scales above ~ 1 TeV. This is because it is essentially a low energy phenomenological theory valid below the electroweak scale characterized by $v = 250$ GeV. The Higgs particle with its associated mass generating mechanism is at the core of the Standard Model, but we

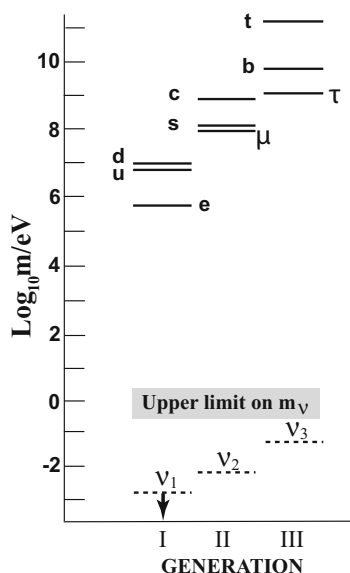


Figure 19.1 Mass hierarchy of elementary particles. The figure suggests a certain pattern which is an unsolved problem. Also notice, the mass of neutrinos are extremely small compared to others. For the neutrinos, only their mass differences and upper

limits of the absolute values are known [$\rightarrow$ Eq. (19.6)]. The dashed lines were drawn assuming $m_3 \gg m_2 \gg m_1$. The origin and hierarchical structure of the mass are outstanding problems of the Standard Model.

have no idea about the dynamical properties of the Higgs particle. Its Lagrangian was chosen for its simplicity and for theoretical convenience with no observational foundations. Unless we discover the Higgs and investigate its properties, the detailed dynamical structure of the Higgs condensation cannot be clarified. Without the knowledge we cannot predict interactions above the critical temperature where the phase transition and electroweak splitting take place. Comparing to the superconductivity with which the Higgs mechanism can be compared, we have the Ginzburg–Landau theory, but not the BCS (Bardeen–Cooper–Schrieffer) theory which clarified the dynamical mechanism of the superconductivity.

To produce the Higgs particle and understand its dynamics is the most urgent theme in the high energy particle physics. Hopefully, LHC (Large Hadron Collider at CERN) will give us hints on them.

19.2

Beyond the Standard Model

19.2.1

Neutrino Oscillation

So far the only firm experimental evidence beyond the Standard Model is neutrino oscillation, discovered in 1998. It is a phenomenon analogous to strangeness oscillation, where a neutrino of one flavor is converted to another flavor ($\nu_e \rightarrow \nu_\mu$, $\nu_\mu \rightarrow \nu_\tau$, etc.) (for a recent review see [42, 87]). It happens if the neutrinos have nonzero mass and mix with each other. It also means there is a right-handed neutrino.

Two-Body Oscillation The neutrino is distinguished by its flavor ($\nu_\alpha = \nu_e, \nu_\mu, \nu_\tau$). It is identified by its production in association with the corresponding charged lepton. Namely, the flavor states are the weak force eigenstates identified at the time of production. In general, they are not necessarily the mass eigenstates (ν_i , $i = 1 - 3$). If they are different and mix with each other, the flavor eigenstates are superpositions of mass eigenstates:

$$|\nu_\alpha\rangle = \sum_j U_{\alpha j} |\nu_j\rangle \quad (19.1)$$

Then as each eigenstate has different time dependence, the mixing rate changes as time passes inducing appearance of different flavor eigenstates. Assuming that the neutrino is stable, and considering the smallness of the neutrino mass, we can express the time development as

$$|\nu_\alpha(t)\rangle = \sum_j U_{\alpha j} |\nu_j\rangle e^{-iE_j t}, \quad E_j = \sqrt{p^2 + m_j^2} \simeq p + \frac{m_j^2}{2E} \quad (19.2)$$

For simplicity, we treat the oscillation between two flavors. Then there is only one independent mixing matrix element. In terms of a mixing angle θ , the flavor eigen-

state can be written as

$$\begin{aligned} |\nu_e\rangle &= \cos \theta |\nu_1\rangle + \sin \theta |\nu_2\rangle \\ |\nu_\mu\rangle &= -\sin \theta |\nu_1\rangle + \cos \theta |\nu_2\rangle \end{aligned} \quad (19.3)$$

Therefore, the probability that what was ν_e at $t = 0$ changes to ν_μ at time t is given by

$$P(\nu_e \rightarrow \nu_\mu; t) = |\langle \nu_\mu | \nu_e(t) \rangle|^2 = |\sin \theta \cos \theta (1 - e^{-i(E_1 - E_2)t})|^2 \quad (19.4a)$$

$$\simeq \sin^2 2\theta \sin^2 \frac{\Delta m^2}{4E} L = \sin^2 2\theta \sin^2 1.27 \frac{\Delta m^2 [(\text{eV})^2]}{E [\text{GeV}]} L [\text{km}] \quad (19.4b)$$

$$\Delta m^2 = |m_1^2 - m_2^2|, \quad L = ct \quad (19.4c)$$

These equations show that oscillation occurs only when there is mixing ($\theta \neq 0$) and also a mass difference exists ($\Delta m^2 \neq 0$). The probability that ν_e survives is given by

$$P(\nu_e \rightarrow \nu_e; t) = 1 - P(\nu_e \rightarrow \nu_\mu; t) \quad (19.5)$$

An optimum condition for the oscillation to be observed is given by $\Delta m^2 L/E \sim 1$. By careful choice of the value of L/E , a wide range of the mass region can be explored. The three-flavor-mixing formulae are different, but so far observed values do not change appreciably if analyzed in either way.

Two kinds of oscillation have been observed. One is disappearance of the atmospheric neutrino [$\nu_\mu (\bar{\nu}_\mu) \rightarrow \nu_\tau (\bar{\nu}_\tau)$: ν_τ not observed], which was detected by comparing downward-going neutrinos with upward-going neutrinos [357]. The energy of the neutrino E is 1–10 GeV, and the oscillation length $L \sim 10\,000$ km (the earth's diameter). The other is that of the solar neutrino, ($\nu_e \rightarrow \nu_\mu, \nu_\tau$) where E is 1–10 MeV and $L \sim 10^8$ km [349]. The observed mass-squared difference and the mixing angle are given by

$$\Delta m_{\text{atm}}^2 = 2.5 \pm 0.5 \times 10^{-3} \text{ eV}^2 \quad \sin^2 2\theta_{\text{atm}} \sim 1.0 \quad (19.6a)$$

$$\Delta m_{\odot}^2 = 8.0 \pm 0.3 \times 10^{-5} \text{ eV}^2 \quad \sin^2 2\theta_{\odot} \sim 0.86 \pm 0.04 \quad (19.6b)$$

Note that if the three masses are hierarchical ($m_3 \gg m_2 \gg m_1$), the mass difference is a measure of the neutrino mass itself and $m_\nu \lesssim 0.05$ eV is expected.

Implication As the mass of a particle in the Standard Model is generated by its coupling with the Higgs, the difference of the mass is ascribed to the difference in the Higgs coupling constant. The observed mass ranges from $\simeq 171$ GeV (m_{top}) down to 0.01 eV (m_ν). It is not likely that a single constant/mechanism can generate masses differing by an order of 10^{13} on the energy scale. Many attempts are under way to solve the mass hierarchy problem. The hierarchy is vertical (i.e. within a generation) as well as horizontal (i.e. across the generations). The vertical symmetry is treated in the grand unified theories (GUTs), to be described soon. Some

seek horizontal symmetry. Yet the neutrino mass problem is special because of its tiny mass as is demonstrated in Fig. 19.1. A promising scenario to explain it is the see-saw mechanism [163, 395], in which the existence of a very heavy right-handed Majorana neutrino makes the mass of the left-handed neutrino light. The model requires the neutrino to be a Majorana particle, whose antiparticle is indistinguishable from itself, and that the underlying physics is of high-energy origin. This naturally leads to GUTs.

After the discovery of neutrino oscillation, the importance of flavor mixing in the lepton sector and CP violation in the PMNS (Pontecorvo–Maki–Nakagawa–Sakata) matrix, which has the same meaning as the CKM (Cabibbo–Kobayashi–Maskawa) matrix in the quark sector, was recognized and at present various neutrino experiments are being planned in many laboratories across the world [42]. They are expected to provide complementary knowledge to what can be obtained at the high-energy frontier.

19.2.2

GUTs: Grand Unified Theories

GUTs were discussed long before the discovery of the neutrino oscillation. The success of the Standard Model immediately prompted the idea that the same recipe that had worked in unifying electroweak forces might also work if the strong force is included. There was another motivation to work on unified theories. The Standard Model has over twenty arbitrary parameters, including the three coupling constants of the electroweak and strong forces, fourteen masses of the weak bosons, quarks and leptons, three mixing parameters and one CP violation phase of the Kobayashi–Maskawa matrix (to be extended to the MNS matrix, which includes six parameters, including two extra Majorana phases), the number of generations, etc. This is to be contrasted with general relativity, which includes only one parameter, Newton’s gravitational constant. A unified theory should not have so many undecided parameters.

Another motivation is theoretical, i.e. the absence of the quantum anomaly. We have learned that there are two ways to break the symmetry that exists in the Lagrangian: explicit and spontaneous. In the former the symmetry-breaking term is put in by hand, for example by applying a magnetic field in an otherwise rotationally symmetric environment. The latter is realized when the ground state of the Lagrangian breaks the symmetry. A third kind appears only in quantum mechanics. When symmetry is conserved at the classical Lagrangian level, but broken when quantized, we talk about a quantum anomaly. When the anomaly exists in gauge theory, which happens in the Standard Model, the gauge symmetry is broken and the theory becomes nonrenormalizable. This is serious, but when the anomalies are added over all the members of a family in the Standard Model, they miraculously disappear. The vanishing of the anomaly is guaranteed by the fact that the

sum of the electric charge in a family adds up to zero:

$$\begin{aligned}\sum_i q_i &= (q_u + q_d) \times 3(\text{color}) + q_e + q_{\nu_e} \\ &= \left\{ \left(\frac{2}{3} \right) + \left(-\frac{1}{3} \right) \right\} \times 3 + (-1) + (0) = 0\end{aligned}\tag{19.7}$$

This fact suggests a strong connection between quarks and leptons. GUT models can naturally incorporate the mechanism of anomaly cancellation.

Since the electroweak theory is a gauge theory based on $SU(2) \times U(1)$ and QCD on $SU(3)$, it is natural to think of a larger group that contains $SU(3) \times SU(2) \times U(1)$. The first GUT theory was based on $SU(5)$ [166], a minimal extension of the group that contains the Standard Model. Extrapolating the running coupling constant of QCD and others, it is generally agreed that GUT unification is only possible at extremely high energies $\sim 10^{14}$ – 10^{16} GeV, commonly referred to as the GUT (energy) scale. Major predictions of GUT were the instability of the proton and the existence of the very heavy magnetic monopole at the GUT scale. The failure to discover the monopole prompted the inflation model, which later became part of standard cosmological theories. Naive GUT models like the one based on simple $SU(5)$ are probably excluded by proton decay experiments (see for instance [252]), which did not observe the decay mode $p \rightarrow e^+ \pi^0$ and set the upper limit of the lifetime as $\gtrsim 10^{33-34} / \text{BR}(p \rightarrow e \pi^0)$ years. But GUTs that take into account supersymmetry predict that the dominant decay mode of the proton is $p \rightarrow K^+ \nu$ and are still viable models.

19.2.3

Supersymmetry

Supersymmetry (SUSY for short) is a symmetry that connects fermions with bosons and considers them as one and the same particle but belonging to different states. If the symmetry exists, all bosons (or fermions) should have fermionic (bosonic) partners with the same mass, constituting super-multiplets. We know there are no such particles. Obviously supersymmetry is broken if it exists. Although no hints of supersymmetry are around experimentally, there are at least two strong theoretical motivations for its existence.

One is the naturalness problem in GUTs. GUT models consider the strong and electroweak forces as united with a common coupling strength. At the GUT energy ($\sim 10^{16}$ GeV), the first spontaneous symmetry breaking occurs and the strong and electroweak forces are split. The second symmetry breaking occurs at the electroweak scale ($\sim 1/\sqrt{G_F} \sim$ a few hundred GeV), where the weak gauge bosons acquire mass and are separated from the electromagnetic force. GUT models contain a variety of extra gauge and Higgs bosons (hereafter referred to as GUT particles) associated with the overarching symmetry, some of which have GUT scale mass induced by spontaneous symmetry breaking. When one wants to calculate quantum corrections of some observables, one has to deal with the two energy scales at every order of the perturbation series, which is referred to as the hierarchy problem.

For the electroweak observables, quantum corrections due to GUT particles with energy scale $\sim 10^{16}$ GeV have to be adjusted to the electroweak scale ~ 100 GeV. Since quantum corrections to bosonic mass like that of Higgs appear in quadratic form, one has to make a fine tuning of order $(10^{16}/10^2)^2 \sim 10^{28}$ at every order of the perturbation calculations. This is considered unnatural and referred to as the “naturalness problem”.

Three types of recipes for its remedy have been proposed. One is to consider the electroweak gauge bosons $W^\pm$, Z as composites of more fundamental fermions; among these the most notable was the technicolor model. The binding energy plays the role of a built-in cutoff. Form factors describing the penetration effect of gauge bosons at GUT scale can be safely assumed to vanish. However, despite many efforts, no viable models have been built and the so-called technicolor models have been largely abandoned.

The second is the addition of extra dimensions, which will be discussed later. The third is supersymmetry. SUSY requires that for every bosonic loop (i.e. higher order) correction, there exists a corresponding fermionic loop and they exactly cancel each other. If the symmetry is broken, it is still under control provided the breaking scale is of the electroweak scale. All the corrections due to GUT particles are canceled away. There is one piece of circumstantial evidence in support of supersymmetry. The three coupling constants α (QED), $\alpha_W[SU(2)]$ and α_s (QCD) are running if quantum corrections are taken into account. The quantum corrections are calculated by including known particles in the loops. Their mass determines at which energy scale they have to be included. It was shown that the observed coupling constants, if extrapolated to GUT scale, do not meet together, in other words, the three forces do not unify. On the other hand, if supersymmetry is included somewhere above the electroweak scale, the number of particles contributing to the quantum corrections doubles and it was shown that all three coupling constants meet together [20, 253]. SusyGUTs (see Fig. 19.2), i.e. GUTs with built-in supersymmetry, are favored over non-SusyGUTs.

The second reason for invoking supersymmetry is the desire to include gravity in unified theories. The symmetry operator changes fermions to bosons or vice versa and hence is a spinor. Amazingly, its commutator is identified as the energy-momentum operator. If it is gauged, i.e. if the operators are made local and include the corresponding gauge fields, they can accommodate general coordinate transformations, which means gravity is included. The quantized gauge theory of supersymmetry is referred to as supergravity. However, supergravity as a mathematical framework for unification has not been very successful. There is an inherent divergence problem associated with gravity being a tensor force. The problem may be avoided if one considers the particle not as a point but a string. This hope triggered the idea of superstring theory.

If supersymmetry exists, every known particle should have its SUSY partner with the same mass but with spin differing by $1/2$ (see Table 19.1). Since no such particles are observed, supersymmetry is broken. But if the symmetry is broken at ~ 1 TeV level, we expect the symmetry to recover at higher energy and that SUSY particles will be discovered with mass in the few hundred GeV region. Since SUSY

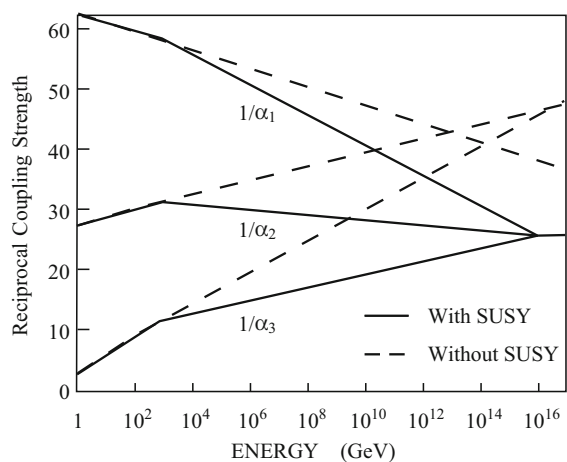


Figure 19.2 Unification of electroweak and strong forces. Plotted are the inverse of α_1 , α_2 and α_3 , which are coupling strengths of $U(1)$, $SU(2)$ and $SU(3)$ gauge symmetries. The couplings do not meet at a point without

supersymmetry but they do with supersymmetry. The supersymmetry is assumed to be broken below ~ 1 TeV, which is the cause of the kinks in the figure.

requires the superpartner to belong to the same multiplet as the known particles, the coupling strength is the same as their partners. Therefore, reaction topologies and their reaction rates are calculable if their mass is known. The design of the detector at LHC takes this into account.

Table 19.1 Superparticles.

Particle	spin	Superpartner	spin	Name
ν_e, ν_μ, ν_τ	1/2	$\hat{\nu}_e, \hat{\nu}_\mu, \hat{\nu}_\tau$	0	sneutrino
e, μ, τ	1/2	$\hat{e}, \hat{\mu}, \hat{\tau}$	0	slepton
u, c, t	1/2	$\hat{u}, \hat{c}, \hat{t}$	0	squark
d, s, b	1/2	$\hat{d}, \hat{s}, \hat{b}$	0	squark
γ	1	$\hat{\gamma}$	1/2	photino ^a
$W^\pm$	1	$\hat{W}^\pm$	1/2	wino
Z	1	$\hat{Z}$	1/2	zino ^a
g	1	$\hat{g}$	1/2	gluino
h, H, A^b	0	$\hat{h}, \hat{H}, \hat{A}$	1/2	higgsino ^a
$H^\pm{}^b$	0	$\hat{H}^\pm$	1/2	chargino
G^c	2	$\hat{G}$	3/2	gravitino

^a Photino, zino and higgsinos are mixed to make 5 neutralinos $\chi_{01}-\chi_{05}$.

^b Additional Higgs are required in the supersymmetric model.

^c Graviton.

In the simplest supersymmetric model, each ordinary particle makes a doublet with its superpartner with spin differing by $1/2$. In many models, SUSY particles are considered to carry R-parity to distinguish them from the ordinary particles. The left-handed and right-handed quarks and leptons separately make doublets with their superpartners having spin 0 because conversion of one to the other is mutual and one cannot convert one SUSY particle with $s = 0$ to q_L and q_R at the same time. Because of R-parity the lightest SUSY particle is stable, which is usually assigned to the neutralino or gravitino. It is a prime candidate for dark matter, which is discussed later.

19.2.4

Superstring Model

See [186]. The string as an elementary particle was first proposed as a model for the hadron. We have learned that the intra-quark forces share many characteristics with the string, Regge trajectories among them (recall discussions in Sects. 13.7.2 and 14.5.3). Above all, it was pointed out that the duality of the Veneziano model was well explained by the string model. However, the hadronic string model had too many unwanted particles and it could not explain asymptotic freedom, so it was abandoned eventually.

When the string model was reconsidered at the Planck energy scale ($\sim 10^{19}$ GeV), the unwanted particles at the GeV–TeV scale turned out to include gravitons, and the model was revived as a candidate for the unified theory of all forces, including gravity. The string model considers the fundamental particle not as a point particle but a one-dimensional string (or even higher dimensional brane) and each particle species is one of its vibrational modes. It is a quantized gauge theory with supersymmetry incorporated in it, and also an extended general relativity in ten-dimensional space-time à la Kaluza–Klein.

The reason for thinking that the string resides in ten-dimensional space-time is the absence of the quantum anomaly. As we live in four-dimensional space-time, the extra six-dimensional space is either curled up to a small size or our four-dimensional brane world floats in a ten-dimensional bulk. The string model even has the possibility of determining the space-time structure of the universe uniquely. Theorists became really taken with it and it is often dubbed the theory of everything (TOE). Many fascinating ideas have been proposed but are unfortunately untestable, because the model works essentially at the Planck energy scale. There are very few subjects if any that can be related to low-energy phenomena for experimental tests. If the only guiding principle is the beauty of the theory, there are too many possibilities and the research field runs the potential danger of becoming more like mathematics or philosophy than physics. We certainly need some observational guidance.

19.2.5

Extra Dimensions

The idea of extra dimensions has already been mentioned in Sect. 18.2.8 in connection with Kaluza–Klein theory. All the superstring models have built-in extra dimensions, but here a different perspective on the extra dimensions is presented.

Gravity is different from other forces in many respects. The graviton has spin 2, its strength is extremely weak, and it is closely related to the space-time structure. If we want to treat gravity on an equal footing with other forces, we have to go to the Planck energy, which is some 10^{17} orders of magnitude apart in scale. The idea of extra dimensions [29] offers the possibility of solving this hierarchy problem. If space-time is $4 + D$ ($D = 1, 2, \dots$) dimensional, gravity spreads in $(3 + D)$ -dimensional space. Then Gauss's law tells us that the gravitational force flux (hence its strength) decreases $\sim r^{-(2+D)}$. Let us denote the gravitational constant in $(4 + D)$ -dimensional space-time as $G_D \equiv M_D^{-2}$. G_D is the fundamental gravitational constant in $(3 + D)$ -dimensional space just like $G_N = 1/M_{\text{Pl}}^2$ is in three-dimensional space. We want to have M_D not too different from the electroweak scale, which we conveniently set at 1 TeV. The gravity under consideration should not contradict Newton's inverse square law. Therefore, we assume the extra D -dimensional space is curled up at the scale of $R = (M_D)^{-1}$. Then the distance in the extra dimensional direction saturates at $r \sim R$ and the power of the force propagation decreases to $\sim r^{-2}$, recovering the usual Newton's law. The dimensional argument constrains the force as

$$\sim \frac{1}{M_D^2} \frac{m_1 m_2}{r^2} \frac{1}{(M_D R)^D} = G_N \frac{m_1 m_2}{r^2} = \frac{1}{M_{\text{Pl}}^2} \frac{m_1 m_2}{r^2} \quad (19.8a)$$

$$\begin{aligned} \therefore R &= \left(\frac{M_{\text{Pl}}^2}{M_D^{2+D}} \right)^{1/D} \simeq \left(\frac{M_D}{1 \text{ TeV}} \right)^{-(D+2)/D} \times (10)^{(32/D)-17} \text{ cm} \\ &= \begin{cases} \sim 0.4 \text{ light years} & D = 1 \\ \sim 1 \text{ mm} & D = 2 \\ \sim 5 \times 10^{-7} \text{ cm} & D = 3 \\ \sim 10^{-9} \text{ cm} & D = 4 \end{cases} \quad (19.8b) \end{aligned}$$

where m_1, m_2 are masses of the particles introduced for convenience and which disappear from the final expression. Since for $r \ll R$, $f \sim r^{-(2+D)}$, and for $r \gg R$, $f \sim r^{-2}$, $D = 1$ is ruled out. However, Newton's law has never been tested at or below the scale of 1 mm, hence $D \geq 2$ is allowed. In other words, the fundamental gravity energy scale being ~ 1 TeV is not excluded if one allows extra dimensions. Then the hierarchy problem may not exist.

The idea of extra dimensions triggered many models and tests of them have been proposed. The values listed above are somewhat more limited now, but the basic idea remains valid. One may refer to [311], p. 1272 for recent developments. In some models like one we described above, only the graviton lives in the extra dimensions and ordinary particles are confined to the four-dimensional Minkowski

space. In other models, gauge particles or all the Standard Model particles reside in the extra dimensions, too. Gravitons (denoted as G) are produced in $e^- + e^+ \rightarrow \gamma (W^\pm, Z) + G$ or hadronic colliders ($q + \bar{q} \rightarrow \text{gluon} + G$). Since they are emitted in the extra-dimensional space as well as in real space, the signal in the hadron collider is a monojet + missing energy.

The momentum of the emitted graviton in the extra dimensions appears in the guise of a mass in our space:

$$E = \sqrt{\mathbf{p}^2 + \sum_{i=1}^D p_i^2 + m^2} \equiv \sqrt{\mathbf{p}^2 + m_{\text{KK}}^2} \quad (19.9)$$

which is referred to as the KK (Kaluza–Klein) tower. Its spectrum is discrete because of the finite space size of the extra dimensions, but the spacing is small ($\Delta E \sim \hbar c/R \sim 10^{-4}$ eV for $D = 2$) and is almost continuous. Therefore, although the production rate is small for one KK graviton, the phase space volume for emitting the KK gravitons is large and the total production rate is at a detectable level. The detection of virtual graviton effects and mini black holes has also been discussed.

19.2.6

Dark Matter

The existence of dark matter has been clarified by astrophysical observations. It is so named because it does not radiate any electromagnetic waves and its existence is only detected by its gravitational force.

The existence of dark matter was first suggested by Zwicky in 1933 [401] to explain the violent random motion of galaxies in the Coma Cluster and revived by Rubin and others in their observations of rotational curves of spiral galaxies [337] in the 1970s. The motion of neutral hydrogens detected as 21.1 cm radio waves (see Fig. 4.1) showed that the rotational velocity of spiral galaxies, including the Milky Way, stays constant far beyond the boundary of shining stars and unambiguously proved the existence of dark matter (Fig. 19.3). Since then dark matter has also been observed in many other places, including gravitational lensing. On the theoretical side, the properties of dark matter have been investigated in conjunction with cosmic expansion to produce the observed large-scale structure of the galaxy distribution and cosmic microwave background radiation.

Problem 19.1

Show that the rotational velocity goes down $\sim \sqrt{M/r}$ if the mass is concentrated at the center and that to make the rotational velocity constant the mass of a galaxy has to grow $\sim r$.

What are the properties of dark matter? It is electrically as well as color-wise neutral, which means that the interaction with ordinary matter is weak. Its properties do not fit with known particles in the Standard Model. The best candidate for dark

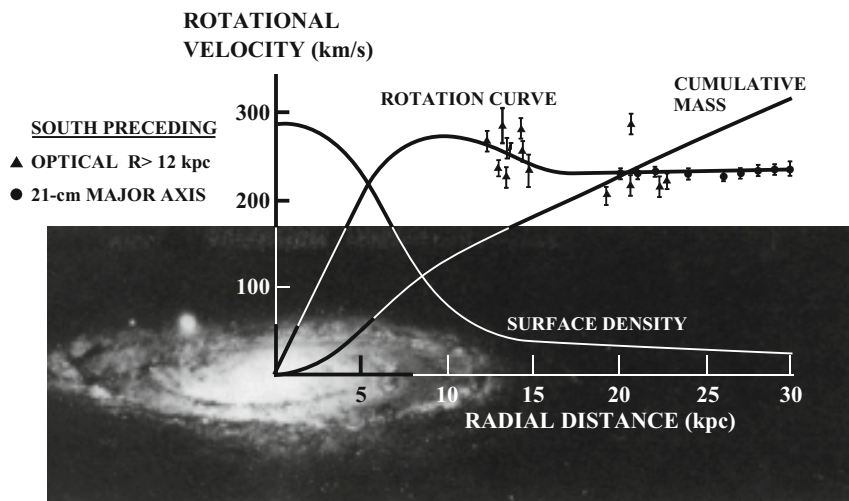


Figure 19.3 An image of the Andromeda galaxy with its rotation curve (velocity of gas clouds and stars orbiting the center of the galaxy vs. distance from the center). The con-

stancy (flatness) of the rotation curve beyond the point where the light ends indicates the presence of dark matter that holds the galaxy together. Adapted from [307].

matter is the stable lightest neutral SUSY particle, known as the neutralino, which is a mixture of the superpartners of the photon, Z and Higgs particle. Another candidate is the axion, a hypothetical very light (pseudo)-scalar particle introduced to explain the so-called strong CP problem in QCD [314, 315].²⁾ From cosmological argument it is known that the dark matter is cold, i.e. its temperature is low or it exists as a mass cluster but not moving appreciably.

Many experiments to search for dark matter are under way. The idea is to catch it in the halo of the Milky Way galaxy. The solar system, including the earth, is orbiting around the center of the galaxy, which arouses winds of dark matter. The signal is a tiny ionization loss liberated by collision of slow dark matter with the detector material. There is also a hope that it may be created directly in the LHC.

19.2.7

Dark Energy

The existence of the cosmological constant in the cosmic evolution equation was suspected for a while by analysis of large scale structure formations. The first direct evidence was given by observation of far away type I supernovae, which exhibited the accelerating expansion of the universe [320, 331] (Fig. 19.4).

Just as the Standard Model of particle physics was established in the 1970s, the cosmological Standard Model (referred to as the concordance model) was estab-

2) A term in the QCD Lagrangian, which is the color extension of $\mathbf{E} \cdot \mathbf{B}$ in QED, violates both P and T invariance. Theoretically it is allowed but phenomenologically it does not exist. The reason is not understood, which is referred to as the "strong CP problem".

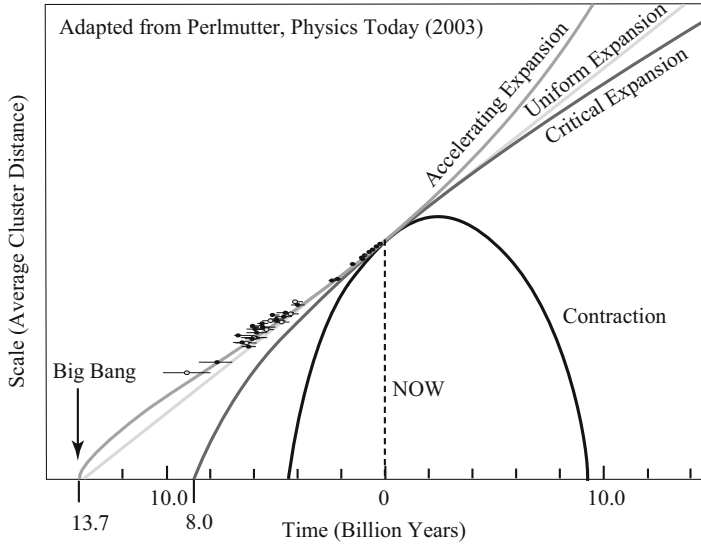


Figure 19.4 History of cosmic expansion, as measured by the high-redshift supernovae (black data points). Given energy compositions, time evolution curves can be calculated from the Friedman equation, assuming flat

cosmic geometry. Without dark energy, the acceleration is always decelerating and the curve is convex upwards. With dark energy, the expansion is decelerating at the beginning but turns to acceleration later [319].

lished in 2003 when the precision data of WMAP³⁾ was analyzed and a consistent view of the dynamical history of the universe was established [242, 351]. According to the concordance model, the energy of our universe is occupied mostly by dark energy and dark matter (see Fig. 19.5). The age of the universe was determined to be 13.69 ± 0.13 billion years [125]. Another outcome was that the space-time curvature of the universe was found to be consistent with zero.

What is the dark energy? How do we know that it is there? It can only be proved indirectly by detecting the acceleration of the expanding universe and analysis of cosmic evolution. We will elaborate a bit on why we believe that dark energy exists by considering the cosmic expansion. It is governed by the Friedman equation

$$H^2 = \frac{8\pi}{3c^2} G_N \rho - \frac{k c^2}{a^2} + \frac{\Lambda c^2}{3} \quad (19.10)$$

where $H = \dot{a}/a$ is the expansion rate of the universe, $\rho = \rho_{\text{matter}} + \rho_{\text{radiation}} + \rho_{\text{vacuum}}$ is the energy density, k is the curvature of space-time and Λ is the cosmological constant. The value of H at present time (denoted as H_0) is referred as the Hubble constant. a is a scale factor and can be any spatial distance, but conveniently one might think of it as the size of the universe or average intergalaxy cluster distances.

3) The Wilkinson Microwave Anisotropy Probe, which was launched in 2001. It has mapped the cosmic microwave background (CMB) radiation (the oldest light in the universe) and produced the first fine-resolution (0.2°) full-sky map of the microwave sky. Details can be obtained from [292].

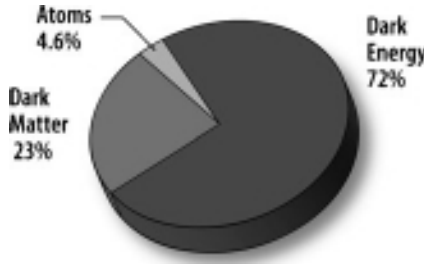


Figure 19.5 WMAP data reveal that the cosmic energy content includes only 4.6% atoms, the building blocks of stars and planets. Dark matter comprises 23% of the universe. This matter, different from atoms, does not emit or absorb light. It has only been detected in-

directly by its gravity. 72% of the universe is composed of “dark energy”, which produces antigravity. This energy, distinct from dark matter, is responsible for the present-day acceleration of the universal expansion. Credit: NASA / WMAP Science Team [292].

The equation is derived from Einstein’s equation of general relativity. However, it can be made to appeal to intuitive understanding (albeit vaguely) if one derives it using only a knowledge of undergraduate mechanics and thermodynamics. Apply Newton’s familiar equation of motion in a gravitational field

$$\frac{1}{2}mv^2 - G \frac{mM}{r} = E \quad (19.11)$$

to a test particle (mass m) at the periphery (r) of a sphere speeding away from the center of the sphere anywhere in the universe. The test particle which can be imagined as a galaxy feels the gravity from the mass inside the sphere but not from outside. Convert $v \rightarrow \dot{r}$, $M \rightarrow (4\pi/3)(\rho/c^2)r^3$, enlarge the notion of mass density to include radiation as well as vacuum energy density, $2E/mc^2 \rightarrow -k$, and one gets the Friedman equation.

Problem 19.2

Show the critical density of the universe is given by

$$\rho_c = \frac{3H^2}{8\pi G} = 1.88 \times 10^{-29} h^2 \text{ g/cm}^3, \quad h = 0.73 \quad (19.12)$$

and confirm the value using the observed value of present expansion rate $H_0 = 73.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $1 \text{ Mpc} = 10^6 \text{ pc}$, $1 \text{ pc} = 3.262 \text{ lightyears}$. The critical energy density is defined as that to give $k = 0$ in the absence of the cosmological constant. It is also the density to separate eternal expansion or eventual collapse of the universe without Λ .

By taking time derivatives of Eq. (19.10) and applying the first law of thermodynamics

$$d(\rho V) + P dV = 0^4 \quad (19.13)$$

we can obtain the cosmic acceleration equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3c^2}(\rho + 3P) + \frac{\Lambda c^2}{3} \quad (19.14)$$

The first term in the parentheses is familiar: Newton's universal attractive force. The pressure (P) in the second term gives the difference acquired by going to general relativity. The cosmological constant Λ , if positive, gives the repulsive force.

Einstein's original motivation for introducing the cosmological constant was to make the universe stationary because he did not know the universe was expanding. The original equation without the cosmological constant was unstable, the universe perpetually expanding or shrinking. To make the universe static, he needed a counter term to cancel the attractive force. Later, when he heard of Hubble's discovery of the expanding universe, he abandoned it, saying it was his biggest blunder in his life. He found no reason to keep it in the equation. It is interesting to note that today's physicists take a different attitude toward the unknown. People are beginning to believe that what is mathematically possible is likely to happen, as is exemplified by Gell-Mann's words: "Everything which is not forbidden is compulsory".

The pressure term could give a repulsive force if it is negative and dominates the first term in Eq. (19.14). Indeed, by inspecting Eqs. (19.10) and (19.14) and recalling that the vacuum has negative pressure $P = -\rho$ (see arguments in boxed paragraph of Sect. 5.5.1), one can easily see that the cosmological constant can be identified as the constant vacuum energy density given by $\rho_{\text{vac}} = \Lambda c^2 / (8\pi G_N)$. Now we have understood that the accelerating expansion ($\ddot{a} > 0$) means that the negative pressure permeates all over the universe. Note that, unlike a cluster of mass, which can exert force $\sim M/r^2$, the vacuum energy is not clustered, and as can be seen from Eq. (19.14) the strength of the force increases linearly with distance. This is why repulsive gravity is not observed. Only at cosmic distances is sufficient vacuum energy accumulated to be perceptible. The cosmic expansion is the only phenomenon (so far) where repulsive gravity can be seen in operation.

Problem 19.3

Derive Eq. (19.14) using Eqs. (19.10) and (19.13).

As is easily derived from Eq. (19.10), the constant and dominant energy density forces the universe to expand exponentially. It is believed that the universe went through such an epoch, known as inflation, before the thermal hot universe developed. Since we really do not know if the acceleration is due to the cosmological constant, which stays constant as its name suggests, or some dynamical object, which

4) Assumption of the uniform universe restricts net thermal flow to vanish in any sampled volume.

can generate time varying negative pressure, the name “dark energy” is adopted to include all possible candidates for the cause of acceleration.

Note, in spontaneous symmetry breaking we learned in Sect. 18.6.2 that the vacuum moves from $\phi = 0$ to $|\phi| = v/\sqrt{2}$. Then the difference in the potential energy $V(0) - V(v/\sqrt{2})$ is liberated as vacuum energy. Namely, a scalar field that can break symmetry spontaneously has the capability of creating vacuum energy. Cosmic inflation is believed to have been produced by such a scalar field, referred to as the inflaton. As the concordance model says that at present the dark energy dominates over all other types of energy, we are, sort of, in the second inflationary stage. The dark energy could be static vacuum energy (equivalent to the cosmological constant) or of dynamical origin. If there is a self-interacting scalar field whose potential energy overwhelms its kinetic energy, it delivers vacuum energy, which varies with time. This is referred to as quintessence, which means the fifth element after the ancient Greek four elements of earth, wind, fire and water, or perhaps modern u , d , v and e ? Whether the dark energy will eventually die away, as in some of the quintessence models, or stay constant, as is the case for the cosmological constant, determines the fate of the universe.

From a particle physicist’s point of view, dark matter and dark energy are new challenges. What have we been doing? After all the long journey of the quest for the fundamental constituents of matter, what we have found occupies no more than 5% of the total cosmic energy budget.

Appendix A

Spinor Representation

A.1

Definition of a Group

A group G is a set of elements with a definite operation that satisfy the following conditions.

1. Closure: For any two elements A, B , a product AB exists and also belongs to the group.
2. Associativity: For any three elements, $(AB)C = A(BC)$.
3. Identity element: A unit element 1 which does not cause anything to happen exists. For all elements A , $1A = A1 = A$.
4. Inverse element: For any element A , an inverse A^{-1} exists and $A^{-1}A = AA^{-1} = 1$.

If $AB = BA$, the set is called commutative or abelian group.

Example A.1

Integers $\mathbb{Z}$ with addition “+” as the operation and zero as an identity operation. Inverse is subtraction.

Example A.2

Rational numbers $\mathbb{Q}$ without zero with operation multiplication “ $\times$ ” and the number 1 as identity element. Zero has to be excluded, because it does not have an inverse element. Other examples are parallel movement or rotation in two-dimensional space (a plane).

If $AB \neq BA$, the set is called noncommutative or nonabelian group.

Example A.3

Space rotation $O(3)$ which is a member of more general group $O(N)$ which keeps the length of the N -dimensional real variables space $r = \sqrt{\sum_{i=1}^N x_i^2}$ invariant.

Group Representation The operations of a group are abstract. By using the group representation they are transformed to a format that is calculable. A familiar representation is a collection of matrices $U(g)$ that act on a vector space $|\Psi\rangle \in V$, such that for all group elements $g \in G$, there exists a corresponding matrix $U(g)$ that satisfies

$$\begin{aligned} U(A)U(B) &= U(AB) \\ U(A^{-1}) &= U^{-1}(A) \\ U(1) &= \mathbf{1} \quad (\text{Unit matrix}) \end{aligned} \quad (\text{A.1})$$

The collection of operand vectors is called the representation space, and the dimensionality n of the vectors is called the dimension of the space.

Note, hermitian matrices do not form a group unless it is commutative.

Example A.4

$U(N)$ is a transformation that keeps the metric of the N -dimensional complex variables space $\ell^2 = \sum_{i=1}^N |u_i|^2$ invariant. If we express u_i 's as a vector $\psi^\top = (u_1, u_2, \dots, u_N)$, then $N \times N$ unitary matrices are the representation of the $U(N)$ group elements. When $\det U = 1$, it is referred to as the special unitary group $SU(N)$.

Subgroup When a subset of a group makes a group by itself under the same multiplication (transformation) law, it is called subgroup.

Subgroup of $SU(N)$: Any $SU(N)$ group has an abelian discrete subgroup H for which

$$\text{An element } h \in H, \quad h = e^{i2\pi k/N} \quad k = 0, 1, 2, \dots, N-1 \quad (\text{A.2})$$

ordinarily designated as Z_N . It is the “center” of the group whose elements commute with all other elements of the group.

A.1.1

Lie Group

When the group element is an analytical function of continuous parameters, it is called a **Lie group**. Any element of the Lie group can be expressed as

$$A(\theta_1, \theta_2, \dots, \theta_n) = \exp \left(i \sum_{i=1}^n \theta_i F_i \right) \quad (\text{A.3})$$

where $n = N^2 - 1$ for the $SU(N)$ group. F_i 's satisfy commutation relations called Lie algebra.

$$[F_i, F_j] = i f_{ijk} F_k \quad (\text{A.4})$$

F_i 's are called generators of the group, f_{ijk} structure constants. Since the whole representation matrices are analytic functions smoothly connected to the unit matrix in the limit of $\theta_i \rightarrow 0$, the Lie algebra Eq. (A.4) completely determines the local structure of the group.¹⁾ f_{ijk} can be made totally antisymmetric in the indices ijk . Determination of the generator is not unique, because different F'_i can be made from any independent linear combination of F_i .²⁾ We are concerned with a compact Lie group in which the range of parameters is finite and closed within a finite volume in the space. For example, the rotation in a two-dimensional plane, $0 \leq \theta \leq 2\pi$ and is compact. In the Lorentz group, $-1 < v < 1$ and is finite but not closed. If another parameter rapidity is used $-\infty < \eta < \infty$ and hence, the group is not compact.

A.2 $SU(2)$

We define a spinor as a base vector of $SU(2)$ group representation in two-dimensional complex variable space.

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \quad (\text{A.6})$$

Then representation matrices are expressed by 2×2 unitary matrices with unit determinant. ξ transforms under $SU(2)$

$$\xi \rightarrow \xi' = U\xi \quad \text{or} \quad \begin{bmatrix} \xi'_1 \\ \xi'_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \quad (\text{A.7})$$

The condition for U is

$$U^{-1} = U^\dagger, \quad \text{and} \quad \det U = 1 \quad (\text{A.8})$$

which constrains the U matrix to be of the form

$$U = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}, \quad |a|^2 + |b|^2 = 1 \quad (\text{A.9})$$

Therefore, the number of independent parameters in $SU(2)$ is three. Equation (A.9) means

$$\begin{aligned} \xi'_1 &= a\xi_1 + b\xi_2 \\ \xi'_2 &= -b^*\xi_1 + a^*\xi_2 \end{aligned} \quad (\text{A.10})$$

1) An example of global structure is "connectedness". For example $SU(2)$ and $O(3)$ have the same structure constants, but $SU(2)$ is simply connected and $O(3)$ is doubly connected and there is a one-to-two correspondence between the two.

2) For instance, instead of angular momentum J_i , we can use $J_\pm = J_x \pm iJ_y$ which satisfies

$$[J_+, J_-] = 2J_z, \quad [J_z, J_\pm] = \pm J_\pm \quad (\text{A.5})$$

Taking the complex conjugate of the equations and rearranging, we have

$$\begin{aligned} (-\xi_2^*)' &= a(-\xi_2^*) + b(\xi_1^*) \\ (\xi_1^*)' &= -b^*(-\xi_2^*) + a^*(\xi_1^*) \end{aligned} \quad (\text{A.11})$$

Therefore

$$\xi^c \equiv \begin{bmatrix} -\xi_2^* \\ \xi_1^* \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \xi_1^* \\ \xi_2^* \end{bmatrix} = -i\sigma_2 \xi^* \quad (\text{A.12})$$

transforms in the same way as ξ .

As there are three traceless hermitian matrices in $SU(2)$, which are taken as Pauli matrices, a general form of the unitary matrix with unit determinant takes a form of

$$U = \exp \left[-i \frac{\boldsymbol{\sigma} \cdot \boldsymbol{\theta}}{2} \right] = \cos \frac{\theta}{2} - \sin \frac{\theta}{2} i \boldsymbol{\sigma} \cdot \mathbf{n} \quad (\text{A.13})$$

$$U^\dagger = \exp \left[i \frac{\boldsymbol{\sigma} \cdot \boldsymbol{\theta}}{2} \right] = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} i \boldsymbol{\sigma} \cdot \mathbf{n} \quad (\text{A.14})$$

$s_i = \sigma_i/2$ are generators of the $SU(2)$ group and satisfy

$$[s_i, s_j] = i \varepsilon_{ijk} s_k \quad (\text{A.15})$$

which is the same commutation relation as that of angular momentum. To see connections with the $O(3)$ (three-dimensional rotation) group, let us take a look at the transformation properties of a product of ξ 's, $H_{ij} = \xi_i \xi_j^*$. Taking H_{ij} as components of a hermitian matrix, it can be expanded in terms of Pauli matrices.

$$H = a\mathbf{1} + \mathbf{b} \cdot \boldsymbol{\sigma} = a\mathbf{1} + \begin{bmatrix} b_z & b_x - ib_y \\ b_x + ib_y & -b_z \end{bmatrix} = \mathbf{A} + \mathbf{B} \quad (\text{A.16})$$

The matrix H transforms under $SU(2)$

$$H \rightarrow H' = U H U^\dagger = \mathbf{A}' + \mathbf{B}' = a'\mathbf{1} + \mathbf{b}' \cdot \boldsymbol{\sigma} \quad (\text{A.17})$$

$$\mathbf{A}' = a'\mathbf{1} = U \mathbf{A} U^\dagger = U a \mathbf{1} U^\dagger = a \mathbf{1} \quad \therefore \quad a' = a \quad (\text{A.18})$$

It shows the coefficient of the unit matrix is invariant. To extract a , we take trace of both sides of Eq. (A.18) and obtain

$$2a = \text{Tr}[H] = \xi_1 \xi_1^\dagger + \xi_2 \xi_2^\dagger = |\xi_1|^2 + |\xi_2|^2 = \xi^\dagger \xi \quad (\text{A.19})$$

We see that $SU(2)$ keeps the norm of the spinor invariant. Considering that the unitary transformation keeps the value of determinant invariant,

$$\det \mathbf{B}' = \det \mathbf{U} \mathbf{B} \mathbf{U}^\dagger = \det \mathbf{B} \quad \therefore \quad b_x'^2 + b_y'^2 + b_z'^2 = b_x^2 + b_y^2 + b_z^2 \quad (\text{A.20})$$

This means the set of three parameters $\mathbf{b} = (b_x, b_y, b_z)$ defined in Eq. (A.17) behaves as a vector in three-dimensional real space, i.e. a vector in $O(3)$. We shall show explicitly that components of $\mathbf{b}$ exactly follow the relation Eq. (3.3).

To extract b_x from Eq. (A.17), we multiply both sides of the equation by σ_x and take trace. Using $\text{Tr}[\sigma_x^2] = 2$, $\text{Tr}[\sigma_x] = \text{Tr}[\sigma_x \sigma_y] = \text{Tr}[\sigma_x \sigma_z] = 0$, we obtain

$$b_x = \frac{1}{2} \text{Tr}[\sigma_x \mathbf{H}] = \frac{1}{2} (\sigma_x)_{ij} \xi_j \xi_i^\dagger = \xi_i^\dagger (\sigma_x)_{ij} \xi_j = \xi^\dagger \sigma_x \xi \quad (\text{A.21a})$$

similarly

$$b_y = \xi^\dagger \sigma_y \xi, \quad b_z = \xi^\dagger \sigma_z \xi \quad (\text{A.21b})$$

Therefore we can write

$$\mathbf{b} = \xi^\dagger \boldsymbol{\sigma} \xi \quad (\text{A.22})$$

Let us show that $\mathbf{b}$ indeed behaves as a vector in three-dimensional real space. Operation of $SU(2)$ transforms $\xi \rightarrow \xi' = \mathbf{U} \xi$. Rotation on axis z by θ transforms

$$b_x \rightarrow b'_x = \xi'^\dagger \sigma_x \xi' = \xi^\dagger \mathbf{U}^\dagger \sigma_x \mathbf{U} \xi \quad (\text{A.23})$$

$$= \xi^\dagger \left(\cos \frac{\theta}{2} + \sin \frac{\theta}{2} i \sigma_z \right) \sigma_x \left(\cos \frac{\theta}{2} - \sin \frac{\theta}{2} i \sigma_z \right) \xi \quad (\text{A.24})$$

Inserting $\sigma_z \sigma_x \sigma_z = -\sigma_x$, $\sigma_z \sigma_x = i \sigma_y$, we obtain

$$b'_x = \cos \theta b_x - \sin \theta b_y \quad (\text{A.25a})$$

Similarly

$$b'_y = \sin \theta b_x + \cos \theta b_y \quad (\text{A.25b})$$

$$b'_z = b_z \quad (\text{A.25c})$$

Thus, $\mathbf{b}$ behaves exactly as a vector in $O(3)$. We may define a spinor whose bi-linear form (A.22) behaves like a vector. Loosely speaking, it is like a square root of the vector.

Composites of Spinors Let us consider products of two independent spinors ξ and η , which may be thought as representing two particles of spin $1/2$,

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}, \quad \eta = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \quad (\text{A.26a})$$

$$s = \frac{1}{\sqrt{2}} (\xi_1 \eta_2 - \xi_2 \eta_1), \quad \mathbf{v} = \begin{cases} v_x = -\frac{1}{\sqrt{2}} (\xi_1 \eta_1 - \xi_2 \eta_2) \\ v_y = \frac{1}{\sqrt{2}i} (\xi_1 \eta_1 + \xi_2 \eta_2) \\ v_z = \frac{1}{\sqrt{2}} (\xi_1 \eta_2 + \xi_2 \eta_1) \end{cases} \quad (\text{A.26b})$$

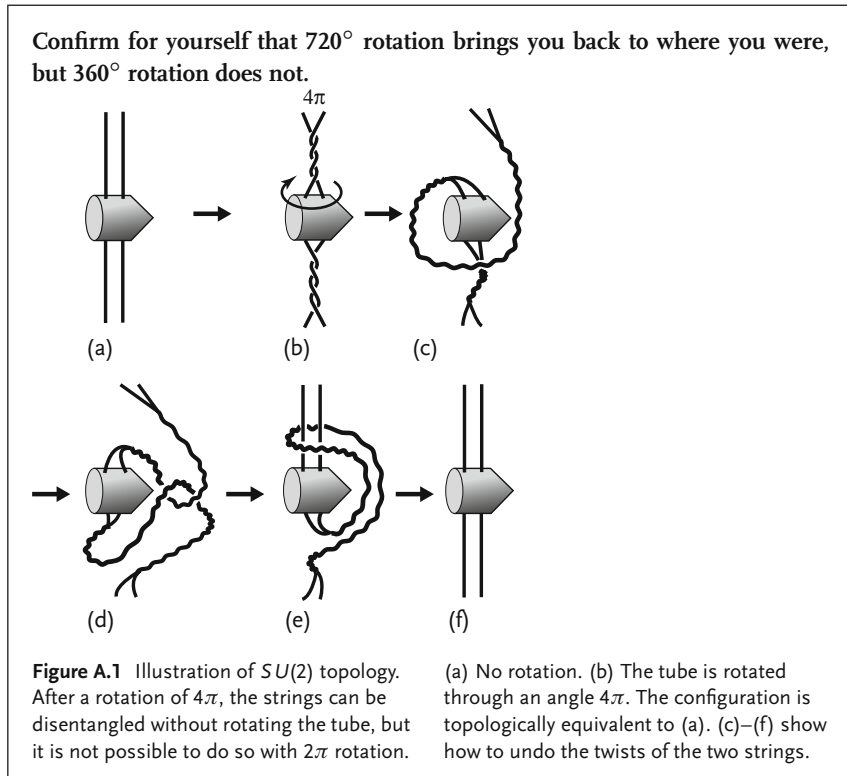
It is easy to show that s and $\mathbf{v}$ behave like a scalar and a vector, namely as spin 0 and 1 particles.

From the above arguments, we realize that $SU(2)$ operation on ξ and that of $O(3)$ on vector (x, y, z) are closely related. Writing down explicitly operations on spin part

$$\mathbf{U} = e^{-i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2} = \cos \frac{\theta}{2} - i \sin \frac{\theta}{2} \boldsymbol{\sigma} \cdot \mathbf{n} \Leftrightarrow \mathbf{R} = e^{-i\mathbf{j}\cdot\boldsymbol{\theta}} = \cos \theta - i \sin \theta \mathbf{S} \cdot \mathbf{n} \quad (\text{A.27})$$

we see that while in $SU(2)$, 4π rotation takes the spinor back to where it started, it takes 2π rotation in $O(3)$. Between the range $2\pi \sim 4\pi$, $\mathbf{U} \rightarrow -\mathbf{U}$ corresponds to the same operation in $O(3)$. The correspondence is 1 : 2. This is the reason that $O(3)$ realizes rotations only on integer spin particles, while $SU(2)$ is capable of handling those of half integer spin.

There are many reasons to believe that $SU(2)$ is active in our real world. For instance, we know that quarks and leptons, the most fundamental particles of matter have spin 1/2. A direct evidence has been obtained by a neutron interference experiment using a monolith crystal which has explicitly shown that it takes 4π to go back to the starting point [86, 103]. To confirm it personally, the reader may try a simple experiment shown in Figure A.1 in the boxed paragraph.



A.3

Lorentz Operator for Spin 1/2 Particle

Referring to Eq. (3.77) and Eq. (3.74) in the text, combinations of spin operator $\mathbf{S}$ and boost operator $\mathbf{K}$

$$\mathbf{A} = \frac{1}{2}(\mathbf{S} + i\mathbf{K}), \quad \mathbf{B} = \frac{1}{2}(\mathbf{S} - i\mathbf{K}) \quad (\text{A.28})$$

satisfy the following equalities

$$[A_i, A_j] = i\varepsilon_{ijk}A_k \quad (\text{A.29a})$$

$$[B_i, B_j] = i\varepsilon_{ijk}B_k \quad (\text{A.29b})$$

$$[A_i, B_j] = 0 \quad (\text{A.29c})$$

Both $\mathbf{A}$ and $\mathbf{B}$ satisfy commutation relations of $SU(2)$, and they commute. In other words, the Lorentz group can be expressed as a direct product of two groups $SU(2) \times SU(2)$, and hence its representations can be specified by a couple of spins (s_A, s_B). Let us consider the most interesting case where one has spin 1/2 and the other 0.

$$(0, 1/2) \quad \mathbf{A} = 0 \rightarrow \mathbf{K} = +i\mathbf{S} = +i\frac{\boldsymbol{\sigma}}{2} \quad (\text{A.30a})$$

$$(1/2, 0) \quad \mathbf{B} = 0 \rightarrow \mathbf{K} = -i\mathbf{S} = -i\frac{\boldsymbol{\sigma}}{2} \quad (\text{A.30b})$$

Correspondingly there are two kinds of spinors which we denote ϕ_L and ϕ_R respectively. They transform under Lorentz operation as

$$\phi_L \xrightarrow{L} \phi'_L = \exp\left[-i\frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\theta} - \frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\eta}\right] \phi_L \equiv \mathbf{M} \phi_L \quad (\text{A.31a})$$

$$\phi_R \xrightarrow{L} \phi'_R = \exp\left[-i\frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\theta} + \frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\eta}\right] \phi_R \equiv \mathbf{N} \phi_R \quad (\text{A.31b})$$

which means that the two spinors behave in the same way under rotation, but transform differently under the boost operation.

Note there is no matrix T that connects each other by transformation $N = TMT^{-1}$, i.e. in mathematical language they are inequivalent. Actually they are connected by

$$N = \zeta M^* \zeta^{-1}, \quad \zeta = -i\sigma_2 \quad (\text{A.32})$$

A.3.1

$SL(2, \mathbb{C})$ Group

Lorentz transformation on spinors is represented by 2×2 complex matrix with unit determinant and is called $SL(2, \mathbb{C})$ group.

$$\xi' = \mathbf{S}\xi; \quad \mathbf{S} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad ad - bc = 1 \quad (\text{A.33})$$

The matrix $\mathbf{S}$ comprises 6 parameters which are related to the three angles and the three rapidities in the Lorentz transformation. To see it, we look again at the transformation properties of a product of spinor components $H_{ij} = \xi_i \xi_j^*$ ($i, j = 1 \sim 2$). In a similar manner as Eq. (A.16), we decompose it to two parts, one with unit matrix and the other with Pauli matrices

$$\mathbf{H} = b^0 \mathbf{1} + \mathbf{b} \cdot \boldsymbol{\sigma} = \begin{bmatrix} b^0 + b_z & b_x - i b_y \\ b_x + i b_y & b^0 - b_z \end{bmatrix} \xrightarrow{\mathbf{S}} \mathbf{H}' = \mathbf{S} \mathbf{H} \mathbf{S}^\dagger \quad (\text{A.34})$$

Here, as the matrix $\mathbf{S}$ is not unitary, $\mathbf{S}^\dagger \mathbf{S} \neq \mathbf{1}$, hence the first and the second term mix. However, the value of the determinant is invariant which leads

$$\det \mathbf{H}' = \det[\mathbf{S} \mathbf{H} \mathbf{S}^\dagger] = |\det \mathbf{S}|^2 \det \mathbf{H} = \det \mathbf{H} \quad (\text{A.35})$$

$$\therefore b^{0'2} - b_x'^2 - b_y'^2 - b_z'^2 = b^{02} - b_x^2 - b_y^2 - b_z^2 \quad (\text{A.36})$$

therefore $SL(2, \mathbb{C})$ operation has been proved to be equivalent to the Lorentz transformation.

To see the difference between ϕ_L and ϕ_R , let us see transformation properties of two vector-like quantities made of constant matrices sandwiched between the two spinors.

$$a^\mu \equiv \phi_R^\dagger \sigma^\mu \phi_R, \quad a_\mu \equiv \phi_R^\dagger \sigma_\mu \phi_R \quad (\text{A.37a})$$

$$b^\mu \equiv \phi_L^\dagger \bar{\sigma}^\mu \phi_L, \quad b_\mu \equiv \phi_L^\dagger \bar{\sigma}_\mu \phi_L \quad (\text{A.37b})$$

$$\sigma^\mu \equiv (\mathbf{1}, \boldsymbol{\sigma}) = \bar{\sigma}_\mu, \quad \bar{\sigma}^\mu \equiv (\mathbf{1}, -\boldsymbol{\sigma}) = \sigma_\mu \quad (\text{A.37c})$$

Put

$$\mathbf{S}_\pm = \exp(\pm \eta \sigma_x / 2), \quad \mathbf{S}_\pm^\dagger = \mathbf{S}_\pm \quad (\text{A.38})$$

and define dashed quantities by

$$\begin{cases} \phi_R'^\dagger \sigma^\mu \phi_R' & \equiv \phi_R^\dagger \sigma^\mu \phi_R \\ \phi_L'^\dagger \bar{\sigma}_\mu \phi_L' & \equiv \phi_L^\dagger \bar{\sigma}_\mu \phi_L \end{cases} \quad (\text{A.39})$$

then by the Lorentz transformation, they change as

$$\begin{cases} \sigma^\mu & \rightarrow \sigma'^\mu = S_+^\dagger \sigma^\mu S_+ \\ \bar{\sigma}_\mu & \rightarrow \bar{\sigma}'_\mu = S_-^\dagger \sigma_\mu S_- \end{cases} \quad (\text{A.40})$$

Calculating “0,1” components under a boost in the x -direction

$$\sigma^{0'}(\bar{\sigma}_0') = \mathbf{S}_\pm^\dagger \mathbf{1} \mathbf{S}_\pm = e^{\pm \sigma_x \eta} = \cosh \eta \pm \sinh \eta \sigma_x \quad (\text{A.41a})$$

$$\begin{aligned} \sigma^{1'}(\bar{\sigma}_1') &= \mathbf{S}_\pm^\dagger \sigma_x \mathbf{S}_\pm = \left(\cosh \frac{\eta}{2} \pm \sinh \frac{\eta}{2} \sigma_x \right) \sigma_x \left(\cosh \frac{\eta}{2} \pm \sinh \frac{\eta}{2} \sigma_x \right) \\ &= (\sinh \eta \pm \cosh \eta \sigma_x) \end{aligned} \quad (\text{A.41b})$$

The transformation is the same as Eq. (3.66). Therefore, we have proved that a^μ , b^μ are contravariant vectors and a_μ , b_μ covariant vectors. We often refer σ^μ , $\bar{\sigma}^\mu$ as Lorentz vectors, but note that it is meaningful only if sandwiched between the two spinors. It is also clear that $\phi_R^\dagger \phi_L$, $\phi_L^\dagger \phi_R$ are Lorentz scalars.

A.3.2

Dirac Equation: Another Derivation

Here, we try to make some correspondence between the spinors and particles considering a special transformation with $\boldsymbol{\theta} = 0$. We can consider a particle with momentum $\mathbf{p}$ as boosted from its rest state. Then there are two distinct states of the particle expressed as

$$\phi_L(\mathbf{p}) = \exp\left[-\frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\eta}\right] \phi_L(0) \quad (\text{A.42a})$$

$$\phi_R(\mathbf{p}) = \exp\left[+\frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\eta}\right] \phi_R(0) \quad (\text{A.42b})$$

Let ϕ_L, ϕ_R be any two-component spinors behaving like ϕ_L, ϕ_R , then $\phi_L^\dagger(1, -\boldsymbol{\sigma})\phi_L$, $\phi_R^\dagger(1, \boldsymbol{\sigma})\phi_R$ are contravariant vectors and their dot products with the energy-momentum vector $\phi_L^\dagger(E + \boldsymbol{\sigma} \cdot \mathbf{p})\phi_L$, $\phi_R^\dagger(E - \boldsymbol{\sigma} \cdot \mathbf{p})\phi_R$ must be scalars. Considering $\phi_L^\dagger\phi_R$, $\phi_R^\dagger\phi_L$ are also scalars, it follows that $(E + \boldsymbol{\sigma} \cdot \mathbf{p})\phi_L$, $(E - \boldsymbol{\sigma} \cdot \mathbf{p})\phi_R$ behave like ϕ_R, ϕ_L and hence must be proportional to them. Using a relation

$$E + \boldsymbol{\sigma} \cdot \mathbf{p} = m (\cosh \eta + \sinh \eta \boldsymbol{\sigma} \cdot \mathbf{n}) = m \exp(\boldsymbol{\sigma} \cdot \boldsymbol{\eta}) \quad (\text{A.43})$$

and putting the proportionality constant C ,

$$\begin{aligned} (E + \boldsymbol{\sigma} \cdot \mathbf{p})\phi_L(\mathbf{p}) &= m \exp(\boldsymbol{\sigma} \cdot \boldsymbol{\eta}) \exp\left[-\frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\eta}\right] \phi_L(0) \\ &= m \exp\left[+\frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\eta}\right] \phi_L(0) \\ &= C \phi_R(\mathbf{p}) = C \exp\left[+\frac{\boldsymbol{\sigma}}{2} \cdot \boldsymbol{\eta}\right] \phi_R(0) \end{aligned} \quad (\text{A.44})$$

It follows that $C = m$ and $\phi_L(0) = \phi_R(0)$. The latter equality is physically a simple statement that at rest there is no means to distinguish ϕ_L from ϕ_R . Applying the same argument to $(E - \boldsymbol{\sigma} \cdot \mathbf{p})\phi_R$ and combining both results give

$$(E + \boldsymbol{\sigma} \cdot \mathbf{p})\phi_L(\mathbf{p}) = m\phi_R(\mathbf{p}) \quad (\text{A.45a})$$

$$(E - \boldsymbol{\sigma} \cdot \mathbf{p})\phi_R(\mathbf{p}) = m\phi_L(\mathbf{p}) \quad (\text{A.45b})$$

This is the celebrated Dirac equation written in terms of two-component spinors³⁾. Historically, the equation was created first and its solution followed. Here, we followed the inverse logic starting from the solution and found the equation. We see

3) Equation (A.45) is written in Weyl representation. In some text books, Dirac representation is used where they are connected by

$$\psi_D = T \psi_W, \quad \{\boldsymbol{\alpha}, \boldsymbol{\beta}\}_D = T \{\boldsymbol{\alpha}, \boldsymbol{\beta}\}_W T^{-1}, \quad T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (\text{A.46a})$$

that it is a relation that 2-component spinors must satisfy and is the simplest one that a spin 1/2 particle respecting the Lorentz invariance has to obey.

Multiplying $(E - \boldsymbol{\sigma} \cdot \mathbf{p})$ on both sides of (A.45a), and using relation $(E - \boldsymbol{\sigma} \cdot \mathbf{p})(E + \boldsymbol{\sigma} \cdot \mathbf{p}) = E^2 - \mathbf{p}^2$ and (A.45b), we obtain

$$(E^2 - \mathbf{p}^2 - m^2)\phi_R = 0, \quad \text{similarly} \quad (E^2 - \mathbf{p}^2 - m^2)\phi_L = 0 \quad (\text{A.47})$$

They satisfy the Einstein relation. If $\phi_L(\mathbf{p})$ and $\phi_R(\mathbf{p})$ are expressed as a spacetime function (i.e. take their Fourier transform), they satisfy the Klein–Gordon equation.

Problem A.1

Define

$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu] \quad (\text{A.48a})$$

$$(S^{23}, S^{31}, S^{12}) = (S_1, S_2, S_3) = \frac{1}{2}\boldsymbol{\Sigma} = \frac{1}{2} \begin{bmatrix} \boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma} \end{bmatrix} \quad (\text{A.48b})$$

$$(S^{01}, S^{02}, S^{03}) = (K_1, K_2, K_3) = \frac{i}{2}\boldsymbol{\alpha} = \frac{1}{2} \begin{bmatrix} -i\boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & i\boldsymbol{\sigma} \end{bmatrix} \quad (\text{A.48c})$$

Show that $\mathbf{S}$ and $\mathbf{K}$ satisfy the same commutation relations as Eq. (3.74), thus qualify as Lorentz operators.

Looking at Eq. (A.42), we see the plane wave solution of the Dirac equation is indeed nothing but a Lorentz boosted spin wave function at rest as it should be from our argument.

Appendix B

Coulomb Gauge

The Lagrangian for an interacting electromagnetic field is given by

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - qj_{\mu}A^{\mu} = -\frac{1}{2}(\mathbf{B}^2 - \mathbf{E}^2) - qj_{\mu}A^{\mu} \quad (\text{B.1a})$$

$$F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}, \quad A^{\mu} = (\phi, \mathbf{A}), \quad j^{\mu} = (\rho, \mathbf{j}) \quad (\text{B.1b})$$

The Euler-Lagrange equation produces the Maxwell equation

$$\partial_{\mu}F^{\mu\nu} = \partial_{\mu}\partial^{\mu}A^{\nu} - \partial^{\nu}(\partial_{\mu}A^{\mu}) = qj^{\nu} \quad (\text{B.2})$$

In the Coulomb gauge, we impose the gauge condition

$$\nabla \cdot \mathbf{A} = 0 \quad (\text{B.3})$$

Using the condition, the 0th component of the Maxwell equation becomes

$$\nabla^2\phi(x) = -q\rho \quad (\text{B.4})$$

which has solution with boundary condition $\phi \xrightarrow{x \rightarrow \infty} 0$

$$\phi(x) = \frac{q}{4\pi} \int d^3x' \frac{\rho(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} \quad (\text{B.5})$$

Notice, the action of the charge at $\mathbf{r}'$ on a point $\mathbf{x}$ is not retarded, but instantaneous. Defining the conjugate momentum fields by

$$\pi^0 = \frac{\delta \mathcal{L}}{\delta \dot{\phi}} = 0 \quad (\text{B.6a})$$

$$\pi^k = \frac{\delta \mathcal{L}}{\delta \dot{A}_k} = -\dot{A}^k - \partial_k \phi = E^k \quad (\text{B.6b})$$

The Hamiltonian density is given by

$$\begin{aligned} \mathcal{H} &= \frac{\delta \mathcal{L}}{\delta \dot{A}_{\mu}} \dot{A}_{\mu} - \mathcal{L} = \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) + \mathbf{E} \cdot \nabla \phi + qj_{\mu}A^{\mu} \\ &= \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) - q\rho\phi + qj_{\mu}A^{\mu} = \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) - \mathbf{j} \cdot \mathbf{A} \end{aligned} \quad (\text{B.7a})$$

where in going to the second line partial integration was taken. This Hamiltonian seems to lack the Coulomb potential term, but actually it is hidden in the E^2 . To see it, we set

$$E_{\parallel} = -\nabla\phi, \quad E_{\perp} = -\dot{\mathbf{A}} \quad (\text{B.8})$$

then

$$\mathbf{E} = \mathbf{E}_{\perp} + \mathbf{E}_{\parallel}, \quad \nabla \times \mathbf{E}_{\parallel} = 0, \quad \nabla \cdot \mathbf{E}_{\perp} = 0 \quad (\text{B.9})$$

by partial integration, we can show that

$$\begin{aligned} |\mathbf{E}|^2 &= |\mathbf{E}_{\perp} + \mathbf{E}_{\parallel}|^2 = |\mathbf{E}_{\perp} - \nabla\phi|^2 = |\mathbf{E}_{\perp}|^2 + |\mathbf{E}_{\parallel}|^2 \\ |\mathbf{E}_{\parallel}|^2 &= \nabla\phi \cdot \nabla\phi = -(\nabla^2\phi)\phi = q\rho\phi \end{aligned} \quad (\text{B.10})$$

Using the above relations, the Hamiltonian is now expressed as

$$\mathcal{H} = \frac{1}{2}(\mathbf{B}^2 + \mathbf{E}_{\perp}^2) + \frac{q}{2}\rho\phi - \mathbf{j} \cdot \mathbf{A}_{\perp} \quad (\text{B.11})$$

The Coulomb term has reappeared. Notice the factor 1/2. The Hamiltonian can now be written as

$$\begin{aligned} H &= \int d^3\mathbf{x} \left[\frac{1}{2} \{ (\nabla \times \mathbf{A}_{\perp})^2 + (\dot{\mathbf{A}}_{\perp})^2 - \mathbf{j} \cdot \mathbf{A}_{\perp} \} \right. \\ &\quad \left. + \frac{q^2}{2} \iint d^3\mathbf{x} d^3\mathbf{x}' \frac{\rho(\mathbf{x})\rho(\mathbf{x}')}{4\pi|\mathbf{x} - \mathbf{x}'|} \right] \end{aligned} \quad (\text{B.12})$$

Thus the degree of freedom corresponding to the scalar potential is replaced with the instantaneous Coulomb potential and the rest are written in terms of transversely polarized potential. They are not manifestly Lorentz invariant, but a detailed investigation can show that Lorentz invariance is respected. Therefore, when intuitive and physical interpretation is desired, the Coulomb gauge is convenient. But for easier calculation, the covariant gauge is more convenient.

B.1

Quantization of the Electromagnetic Field in the Coulomb Gauge

Since $\pi^0 = 0$, the scalar component does not propagate meaning it is not a dynamical object hence we cannot quantize A^0 . We leave it as a classical potential and quantize only the three fields $\mathbf{A}$ according to the standard procedure with the conditions

$$[A^i(\mathbf{x}), \pi^j(\mathbf{y})]_{t_x=t_y} = i\delta_{\perp ij}\delta^3(\mathbf{x} - \mathbf{y}) \quad (\text{B.13a})$$

$$[A^i(\mathbf{x}), A^j(\mathbf{y})]_{t_x=t_y} = [\pi^i(\mathbf{x})\pi^j(\mathbf{y})]_{t_x=t_y} = 0 \quad (\text{B.13b})$$

$$\delta_{\perp ij}\delta^3(\mathbf{x} - \mathbf{y}) = \int \frac{d^3k}{(2\pi)^3 2\omega} \left(\delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2} \right) e^{i\mathbf{k} \cdot (\mathbf{x} - \mathbf{y})} \quad (\text{B.13c})$$

The reason for using $\delta_{\perp ij}$ instead of δ_{ij} is to make the commutators consistent with the Coulomb gauge condition (B.3), so that both sides of the commutators are divergenceless. The field quantization in terms of the annihilation or creation operators is

$$A(x) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}, \lambda = \pm} \frac{1}{\sqrt{2\omega V}} \left[a_{\mathbf{k}, \lambda} \boldsymbol{\epsilon}(\mathbf{k}, \lambda) e^{-ik \cdot x} + a_{\mathbf{k}, \lambda}^\dagger \boldsymbol{\epsilon}(\mathbf{k}, \lambda)^* e^{ik \cdot x} \right] \quad (\text{B.14a})$$

$$\sum_{\lambda = \pm} \varepsilon^i(\lambda) \varepsilon^j(\lambda) = \delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2} \quad (\text{B.14b})$$

$$[a_{\mathbf{k}, \lambda}, a_{\mathbf{k}', \lambda'}^\dagger] = \delta_{\mathbf{k}\mathbf{k}'} \delta_{\lambda\lambda'} \quad (\text{B.14c})$$

$$[a_{\mathbf{k}, \lambda}, a_{\mathbf{k}', \lambda'}] = [a_{\mathbf{k}, \lambda}^\dagger, a_{\mathbf{k}', \lambda'}^\dagger] = 0 \quad (\text{B.14d})$$

Only the transverse polarization is included. The gauge condition requires

$$\nabla \cdot \mathbf{A} = 0 \quad \rightarrow \quad \begin{cases} \mathbf{k} \cdot \boldsymbol{\epsilon}(\mathbf{k}, \lambda = \pm) = 0 \\ \sum_{\lambda = \pm} \varepsilon^i(\mathbf{k}, \lambda) \varepsilon^j(\mathbf{k}, \lambda)^* = \delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2} \end{cases} \quad (\text{B.15})$$

The quantization is carried out in the same manner as two independent Klein-Gordon fields. The only subtlety is in the definition of transverse delta function to make the commutator relations consistent with the gauge requirement.

The advantage of the Coulomb gauge is clearness of its physical image, but the disadvantage is that the form of the propagator and hence the calculation of the transition amplitude becomes rather complicated. To derive the Feynman propagator in the Coulomb gauge, we introduce a timelike vector $\eta_\mu = (1, 0, 0, 0)$, which is orthogonal to the polarization vectors i.e. $\eta \cdot \varepsilon^\lambda = 0$. We form four orthonormal vectors $\eta_\mu, \varepsilon_\mu^1, \varepsilon_\mu^2, \bar{k}_\mu$, where $\bar{k}_\mu$ is a vector defined by

$$\bar{k}_\mu = \frac{k_\mu - (k \cdot \eta) \eta_\mu}{[(k \cdot \eta) - k^2]^{1/2}} \quad (\text{B.16})$$

As $k \cdot \varepsilon = \eta \cdot \varepsilon = 0$, it is easy to prove that $\bar{k}$ is a spacelike unit vector, i.e. $\bar{k}^2 = -1$, and orthogonal to the polarization vectors ($\bar{k} \cdot \varepsilon = 0$). Then we have

$$g_{\mu\nu} = \eta_\mu \eta_\nu - \bar{k}_\mu \bar{k}_\nu - \sum_{\lambda=1}^2 \varepsilon_\mu^{\lambda*}(k) \varepsilon_\nu^\lambda(k) \quad (\text{B.17})$$

In the covariant Feynman gauge, the propagator is expressed by simply

$$\frac{\sum_{\lambda=0}^3 \varepsilon_\mu^{\lambda*}(k) \varepsilon_\nu^\lambda(k)}{k^2 + i\varepsilon} = \frac{-g_{\mu\nu}}{k^2 + i\varepsilon} \quad (\text{B.18})$$

whereas in the Coulomb gauge the numerator is replaced by

$$\begin{aligned} \sum_{\lambda=1}^2 \varepsilon_\mu^{\lambda*}(k) \varepsilon_\nu^\lambda(k) &= -g_{\mu\nu} + \eta_\mu \eta_\nu - \bar{k}_\mu \bar{k}_\nu \\ &= -g_{\mu\nu} - \frac{k_\mu k_\nu}{(k \cdot \eta)^2 - k^2} + \frac{(k \cdot \eta)(k_\mu \eta_\nu + \eta_\mu k_\nu)}{(k \cdot \eta)^2 - k^2} - \frac{k^2 \eta_\mu \eta_\nu}{(k \cdot \eta)^2 - k^2} \end{aligned} \quad (\text{B.19})$$

Appendix C

Dirac Matrix and Gamma Matrix Traces

C.1

Dirac Plane Wave Solutions

$$\begin{aligned}
 \not{p} &\equiv p_\mu \gamma^\mu \\
 (\not{p} - m)u(p) &= 0, \quad (\not{p} + m)v(p) = 0 \\
 \bar{u}(p)(\not{p} - m) &= 0, \quad \bar{v}(p)(\not{p} + m) = 0 \\
 u_s^\dagger(p)u_r(p) &= v_s^\dagger(p)v_r(p) = 2E\delta_{rs} \\
 \bar{u}_s(p)u_r(p) &= -\bar{v}_s(p)v_r(p) = 2m\delta_{rs} \\
 \bar{u}_s(p)v_r(p) &= \bar{v}_s(p)u_r(p) = 0 \\
 \sum_{r=\pm 1/2} u_r(p)\bar{u}_r(p) &= \not{p} + m, \quad \sum_{r=\pm 1/2} v_r(p)\bar{v}_r(p) = \not{p} - m
 \end{aligned} \tag{C.1a}$$

Gordon Equality

$$\begin{aligned}
 \bar{u}_s(p')\gamma^\mu u_r(p) &= \frac{1}{2m}\bar{u}_s(p')[(\not{p}' + \not{p})^\mu + i\sigma^{\mu\nu}(p' - p)_\nu]u_r(p) \\
 \bar{u}_s(p')\gamma^\mu\gamma^5 u_r(p) &= \frac{1}{2m}\bar{u}_s(p')[(\not{p}' - \not{p})^\mu + i\sigma^{\mu\nu}(p' + p)_\nu]\gamma^5 u_r(p)
 \end{aligned} \tag{C.1b}$$

C.2

Dirac γ Matrices

$$\begin{aligned}
 \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu &= 2g^{\mu\nu} \\
 \gamma^5 &= i\gamma^0\gamma^1\gamma^2\gamma^3 = -\frac{i}{4!}\sum \varepsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma \\
 (\gamma^5)^2 &= 1, \quad \gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0 \\
 \sigma^{\mu\nu} &= \frac{i}{2}[\gamma^\mu, \gamma^\nu], \quad \sigma^{0i} = i\alpha^i, \quad \sigma^{ij} = \varepsilon_{ijk}\sigma^k \\
 \gamma^0(\gamma^\mu)^\dagger\gamma^0 &= \gamma^\mu, \quad \gamma^0(\gamma^5)^\dagger\gamma^0 = -\gamma^5 \\
 \gamma^0(\gamma^5\gamma^\mu)^\dagger\gamma^0 &= \gamma^5\gamma^\mu, \quad \gamma^0(\sigma^{\mu\nu})^\dagger\gamma^0 = \sigma^{\mu\nu}
 \end{aligned} \tag{C.2a}$$

$$\begin{aligned}
\gamma^5 \gamma_\sigma &= -\frac{i}{3!} \varepsilon_{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \\
\gamma^\mu \gamma^\nu \gamma^\rho &= g^{\mu\nu} \gamma^\rho - g^{\mu\rho} \gamma^\nu + g^{\nu\rho} \gamma^\mu - i \gamma^5 \varepsilon^{\mu\nu\rho\sigma} \gamma_\sigma
\end{aligned} \tag{C.2b}$$

Charge Conjugation Matrices

$$\begin{aligned}
C &= i \gamma^2 \gamma^0, \quad C^T = C^\dagger = -C, \quad C C^\dagger = 1, \quad C^2 = -1 \\
C \gamma^{\mu T} C^{-1} &= -\gamma^\mu, \quad C \gamma^{5T} C^{-1} = \gamma^5 \\
C(\gamma^5 \gamma^\mu)^T C^{-1} &= \gamma^5 \gamma^\mu, \quad C \sigma^{\mu\nu T} C^{-1} = -\sigma^{\mu\nu}
\end{aligned} \tag{C.3a}$$

Chiral (Weyl) Representation

$$\alpha = \begin{bmatrix} -\sigma & 0 \\ 0 & \sigma \end{bmatrix}, \quad \beta = \gamma^0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \gamma = \beta \alpha = \begin{bmatrix} 0 & \sigma \\ -\sigma & 0 \end{bmatrix} \tag{C.4a}$$

$$\gamma^5 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} i\sigma^2 & 0 \\ 0 & -i\sigma^2 \end{bmatrix} \tag{C.4b}$$

Pauli–Dirac Representation

$$\alpha = \begin{bmatrix} 0 & \sigma \\ \sigma & 0 \end{bmatrix}, \quad \beta = \gamma^0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \gamma = \beta \alpha = \begin{bmatrix} 0 & \sigma \\ -\sigma & 0 \end{bmatrix} \tag{C.5a}$$

$$\gamma^5 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad C = i \gamma^2 \gamma^0 = \begin{bmatrix} 0 & -i\sigma^2 \\ -i\sigma^2 & 0 \end{bmatrix} \tag{C.5b}$$

Writing wave functions in the Weyl and Pauli–Dirac representations as Φ_W , Φ_D , they are mutually connected by

$$\Phi_D = S \Phi_W, \quad \gamma_D = S \gamma_W S^{-1}, \quad S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \tag{C.6}$$

C.2.1

Traces of the γ Matrices

$$\begin{aligned}
\gamma_\mu \gamma^\mu &= 4 \\
\gamma_\mu \not{A} \gamma^\mu &= -2 \not{A} \\
\gamma_\mu \not{A} \not{B} \gamma^\mu &= 4 (A \cdot B) \\
\gamma_\mu \not{A} \not{B} \not{C} \gamma^\mu &= -2 \not{C} \not{B} \not{A} \\
\gamma_\mu \not{A} \not{B} \not{C} \not{D} \gamma^\mu &= 2 [\not{D} \not{A} \not{B} \not{C} + \not{C} \not{B} \not{A} \not{D}]
\end{aligned} \tag{C.7a}$$

$$\begin{aligned}
\text{Tr}[1] &= 4, \quad \text{Tr}[\gamma^\mu] = \text{Tr}[\gamma^5] = 0 \\
\text{Tr}[\gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_{2n+1}}] &= 0 \\
\text{Tr}[\gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_n}] &= (-1)^n \text{Tr}[\gamma^{\mu_n} \dots \gamma^{\mu_2} \gamma^{\mu_1}] \\
\text{Tr}[\gamma^\mu \gamma^\nu] &= 4g^{\mu\nu} \\
\text{Tr}[\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma] &= 4[g^{\mu\nu} g^{\rho\sigma} + g^{\mu\sigma} g^{\nu\rho} - g^{\mu\rho} g^{\nu\sigma}] \\
\text{Tr}[\gamma^5 \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma] &= -4i \varepsilon^{\mu\nu\rho\sigma} = 4i \varepsilon_{\mu\nu\rho\sigma} \\
\text{Tr}[\gamma^5] &= \text{Tr}[\gamma^5 \gamma^\mu] = \text{Tr}[\gamma^5 \gamma^\mu \gamma^\nu] = \text{Tr}[\gamma^5 \gamma^\mu \gamma^\nu \gamma^\rho] = 0
\end{aligned} \tag{C.7b}$$

C.2.2

Levi-Civita Antisymmetric Tensor

$$\varepsilon^{ijk} = \varepsilon_{ijk} = \begin{cases} +1 & ; ijk = \text{even permutation of } 123 \\ -1 & ; ijk = \text{odd permutation of } 123 \\ 0 & ; \text{if any two indices are the same} \end{cases} \tag{C.8a}$$

$$\varepsilon^{\mu\nu\rho\sigma} = -\varepsilon_{\mu\nu\rho\sigma} = \begin{cases} +1 & ; \mu\nu\rho\sigma = \text{even permutation of } 0123 \\ -1 & ; \mu\nu\rho\sigma = \text{odd permutation of } 0123 \\ 0 & ; \text{if any two indices are the same} \end{cases} \tag{C.8b}$$

$$\begin{aligned}
\varepsilon^{\mu\nu\rho\sigma} \varepsilon^{\alpha\beta\gamma\delta} &= -\det[g^{\lambda\lambda'}], \quad \lambda = \mu\nu\rho\sigma, \quad \lambda' = \alpha\beta\gamma\delta \\
\varepsilon^{\alpha\nu\rho\sigma} \varepsilon_{\alpha}{}^{\beta\gamma\delta} &= -\det[g^{\lambda\lambda'}], \quad \lambda = \nu\rho\sigma, \quad \lambda' = \beta\gamma\delta \\
\varepsilon^{\alpha\beta\rho\sigma} \varepsilon_{\alpha\beta}{}^{\gamma\delta} &= -2(g^{\rho\gamma} g^{\sigma\delta} - g^{\rho\delta} g^{\sigma\gamma}) \\
\varepsilon^{\alpha\beta\gamma\sigma} \varepsilon_{\alpha\beta\gamma}{}^{\delta} &= -6g^{\sigma\delta} \\
\varepsilon^{\alpha\beta\gamma\delta} \varepsilon_{\alpha\beta\gamma\delta} &= -24
\end{aligned} \tag{C.8c}$$

C.3

Spin Sum of $|\mathcal{M}_{fi}|^2$

$$\begin{aligned}
K^{\mu\nu}(p', p) &= \sum_{r,s=1,2} [\bar{u}_s(p') \gamma^\mu (a - b\gamma^5) u_r(p)] \\
&\quad \times [\bar{u}_s(p') \gamma^\nu (a - b\gamma^5) u_r(p)]^* \\
&= \text{Tr}[\gamma^\mu (a - b\gamma^5)(\not{p}' + m)(a^* + b^* \gamma^5) \gamma^\nu (\not{p} + m')] \\
&= 4[(|a|^2 + |b|^2)\{p'^\mu p^\nu + p'^\nu p^\mu - g^{\mu\nu}(p' \cdot p)\} \\
&\quad + (|a|^2 - |b|^2)g^{\mu\nu} m' m - 2i \text{Re}[(ab^*) \varepsilon^{\mu\nu\rho\sigma} p'_\rho p_\sigma]]
\end{aligned} \tag{C.9}$$

For antiparticles, replace m or m' or both $\rightarrow -m, -m'$.

C.3.1

A Frequently Used Example

$e(p_1) + \mu(p_2) \rightarrow e(p_3) + \mu(p_4)$ (1) VV scattering; $a = 1, b = 0, m_1 = m_3 = m, m_2 = m_4 = M$

$$\begin{aligned}
 K_V^{\mu\nu}(p_3, p_1) &= 4 \left[\{p_3^\mu p_1^\nu + p_3^\nu p_1^\mu - g^{\mu\nu}(p_3 \cdot p_1)\} + m^2 g^{\mu\nu} \right] \\
 &= 4 \left[p_3^\mu p_1^\nu + p_3^\nu p_1^\mu + g^{\mu\nu} \left(\frac{q^2}{2} \right) \right], \quad q = p_1 - p_3 \\
 K_V^{\mu\nu}(p_3, p_1) K_{V\mu\nu}(p_4, p_2) &= 16 \left[2(p_1 \cdot p_2)(p_3 \cdot p_4) + 2(p_1 \cdot p_4)(p_2 \cdot p_3) + q^2 \{ (p_1 \cdot p_3) \right. \\
 &\quad \left. + (p_2 \cdot p_4) \} + q^4 \right] \quad (C.10a) \\
 &= 8[(s - m^2 - M^2)^2 + (u - m^2 - M^2)^2 + 2t(m^2 + M^2)] \\
 &= 16 \left[(s - m^2 - M^2)(m^2 + M^2 - u) + t(m^2 + M^2) + \frac{t^2}{2} \right] \\
 &\simeq 8(s^2 + u^2) \quad \text{for } s \gg m^2, M^2
 \end{aligned}$$

The spin averaged matrix element becomes

$$|\overline{\mathcal{M}}_{fi}|^2 = \frac{1}{4} \sum_{\text{spin}} |\mathcal{M}_{fi}|^2 \xrightarrow{s \gg m^2, M^2} \frac{2e^4}{q^4} (s^2 + u^2) \quad (C.10b)$$

The cross section in the CM frame is given by

$$\begin{aligned}
 \left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} &= \frac{p_3}{p_1} \frac{1}{64\pi^2 s} |\overline{\mathcal{M}}_{fi}|^2 \\
 &= \frac{\alpha^2}{q^4} \frac{s^2 + u^2}{2s} = \frac{\alpha^2}{q^4} \frac{s}{2} \left[1 + \left(\frac{1 + \cos \theta}{2} \right)^2 \right] \quad (C.10c)
 \end{aligned}$$

(2) LL, LR, RL, RR scattering; (L: $a = b = 1/2$, R: $a = -b = 1/2$)

$$\begin{aligned}
 K_L^{\mu\nu}(p_3, p_1) &= 2 \left[\{p_3^\mu p_1^\nu + p_3^\nu p_1^\mu - g^{\mu\nu}(p_3 p_1)\} - i \varepsilon^{\mu\nu\rho\sigma} p_{3\rho} p_{1\sigma} \right] \\
 K_R^{\mu\nu}(p_3, p_1) &= 2 \left[\{p_3^\mu p_1^\nu + p_3^\nu p_1^\mu - g^{\mu\nu}(p_3 p_1)\} + i \varepsilon^{\mu\nu\rho\sigma} p_{3\rho} p_{1\sigma} \right] \\
 K_L^{\mu\nu}(p_3, p_1) K_{L\mu\nu}(p_4, p_2) &= K_R^{\mu\nu}(p_3, p_1) K_{R\mu\nu}(p_4, p_2) \\
 &= 16(p_1 \cdot p_2)(p_3 \cdot p_4) = 4(s - m^2 - M^2)^2 \\
 &\simeq 4s^2 \quad s \gg m^2, M^2 \\
 K_L^{\mu\nu}(p_3, p_1) K_{R\mu\nu}(p_4, p_2) &= K_R^{\mu\nu}(p_3, p_1) K_{L\mu\nu}(p_4, p_2) \\
 &= 16(p_1 \cdot p_4)(p_2 \cdot p_3) = 4(u - m^2 - M^2)^2 \\
 &\simeq 4u^2 \quad u \gg m^2, M^2 \quad (C.10d)
 \end{aligned}$$

$e(p_1) + \bar{e}(p_2) \rightarrow \mu(p_3) + \bar{\mu}(p_4)$ Here, we consider only VV scattering. Then, $K^{\mu\nu}(p', p)$ in Eq. (C.9) becomes

$$\begin{aligned} K_V^{\mu\nu}(p', p) &= 4[\{p'^{\mu} p^{\nu} + p'^{\nu} p^{\mu} - g^{\mu\nu}(p' \cdot p)\} - m^2 g^{\mu\nu}] \\ &= 4\left[p'^{\mu} p^{\nu} + p'^{\nu} p^{\mu} - g^{\mu\nu}\left(\frac{s}{2}\right)\right], \quad s = (p + p')^2 \end{aligned} \quad (C.11)$$

$$\begin{aligned} K_V^{\mu\nu}(p_2, p_1) K_{V\mu\nu}(p_4, p_3) &= 16[2(p_1 \cdot p_3)(p_2 \cdot p_4) + 2(p_1 \cdot p_4)(p_2 \cdot p_3) \\ &\quad - s\{(p_1 \cdot p_2) + (p_3 \cdot p_4)\} + s^2], \\ &= 8[(t - m^2 - M^2)^2 + (u - m^2 - M^2)^2 + 2s(m^2 + M^2)] \end{aligned} \quad (C.12)$$

In the CM frame,

$$\begin{aligned} E_1 &= E_2 = \sqrt{p_1^2 + m^2}, \quad E_3 = E_4 = \sqrt{p_3^2 + M^2}, \\ s &= 4E_1^2 = 4E_3^2 = 4E_1 E_3, \quad p_1^2 = \frac{s - 4m^2}{4}, \quad p_3^2 = \frac{s - 4M^2}{4} \\ (p_1 \cdot p_3)(p_2 \cdot p_4) &= (E_1 E_3 - p_1 p_3 \cos \theta)^2, \\ (p_1 \cdot p_4)(p_2 \cdot p_3) &= (E_1 E_3 + p_1 p_3 \cos \theta)^2 \end{aligned} \quad (C.13)$$

Inserting Eq. (C.13) in Eq. (C.12), we have

$$\begin{aligned} K_V^{\mu\nu}(p_2, p_1) K_{V\mu\nu}(p_4, p_3) &= 16\left[4(E_1^2 E_3^2 + p_1^2 p_3^2 \cos^2 \theta) + s(m^2 + M^2)\right] \end{aligned} \quad (C.14)$$

$$= 4s^2 \left[(1 + \cos^2 \theta) + 4 \frac{(m^2 + M^2)}{s} \sin^2 \theta + 16 \frac{m^2 M^2}{s^2} \cos^2 \theta \right] \quad (C.15)$$

$$\begin{aligned} \left. \frac{d\sigma}{d\Omega} \right|_{\text{CM}} &= \frac{p_3}{p_1} \frac{1}{64\pi^2 s} \frac{1}{4} \frac{(4\pi\alpha)^2}{s^2} K_V^{\mu\nu}(p_2, p_1) K_{V\mu\nu}(p_4, p_3) \\ &= \frac{\alpha^2}{4s} \frac{p_3}{p_1} \left[(1 + \cos^2 \theta) + 4 \frac{(m^2 + M^2)}{s} \sin^2 \theta \right. \\ &\quad \left. + 16 \frac{m^2 M^2}{s^2} \cos^2 \theta \right] \end{aligned} \quad (C.16)$$

The total cross section is

$$\sigma_{\text{TOT}} = \frac{4\pi\alpha^2}{3s} \frac{p_3}{p_1} \left[1 + 2 \frac{(m^2 + M^2)}{s} + \frac{4m^2 M^2}{s^2} \right] \quad (C.17a)$$

$$= \frac{4\pi\alpha^2}{3s} \frac{\beta_f}{\beta_i} \left[1 + \frac{2 - \beta_i^2 - \beta_f^2}{2} + \frac{(1 - \beta_i^2)(1 - \beta_f^2)}{4} \right] \quad (C.17b)$$

where $\beta_{i,f} = p_{i,f}/E_{i,f}$ is the particle's initial and final velocity in CM.

C.3.2

Polarization Sum of the Vector Particle

Polarization vectors ($\varepsilon^\mu(\lambda)$) are orthogonal to the momentum q^μ of the particle, the photon by the Lorentz condition and the massive boson by Proca's equation:

$$q_\mu \varepsilon^\mu(\lambda) = 0 \quad (\text{C.18})$$

Only the transverse polarizations are allowed for the real photon, and they satisfy

$$\sum_{\lambda=\pm} \varepsilon^i(\lambda) \varepsilon^j(\lambda) = \delta_{ij} - \frac{q^i q^j}{|q|^2} \quad (\text{C.19})$$

For the case of the massive vector boson,

$$\sum_{\lambda=\pm, l} \varepsilon^\mu(\lambda) \varepsilon^\nu(\lambda) = - \left(g^{\mu\nu} - \frac{q^\mu q^\nu}{m^2} \right) \quad (\text{C.20})$$

This can be understood by first considering the polarization vectors in the particle rest coordinate system, then boosting to the direction of the momentum. Define four polarization vectors $\varepsilon^\mu(\lambda)$, ($\lambda = 0 \sim 3$) in the rest frame of the particle by

$$\begin{aligned} \varepsilon^\mu(0) &= (1, 0, 0, 0) \\ \varepsilon^\mu(1) &= (0, 1, 0, 0) \\ \varepsilon^\mu(2) &= (0, 0, 1, 0) \\ \varepsilon^\mu(3) &= (0, 0, 0, 1) \end{aligned} \quad (\text{C.21a})$$

Alternatively we can adopt circular polarizations for the transverse components. In the coordinate system where the 4-momentum of the particle is given by

$$q^\mu = (\omega, 0, 0, |q|) \quad (\text{C.21b})$$

The four polarization vectors are boosted to

$$\begin{aligned} \varepsilon^\mu(s) &= \left(\frac{\omega}{m}, 0, 0, \frac{|q|}{m} \right) \\ \varepsilon^\mu(+) &= -\frac{1}{\sqrt{2}} (0, 1, +i, 0) \\ \varepsilon^\mu(-) &= \frac{1}{\sqrt{2}} (0, 1, -i, 0) \\ \varepsilon^\mu(l) &= \left(\frac{|q|}{m}, 0, 0, \frac{\omega}{m} \right) \end{aligned} \quad (\text{C.21c})$$

Here, we used names $\pm$ for circular, scalar for the 0th and longitudinal for the third component of the polarization. The direction of the longitudinal photon is parallel to the momentum. They satisfy the orthogonality condition

$$\varepsilon^\mu(\lambda) \varepsilon_\mu(\lambda')^* = g^{\lambda\lambda'} \quad (\text{C.21d})$$

and the completeness condition

$$\sum_{\lambda=\text{all}} g^{\lambda\lambda'} \varepsilon^\mu(\lambda) \varepsilon^\nu(\lambda')^* = g^{\mu\nu} \quad (\text{C.21e})$$

Requirement of the constraint Eq. (C.18) removes the scalar polarization and the completeness condition becomes Eq. (C.20).

For virtual photons with time-like momentum (i.e. $q^2 > 0$), the above discussion can be extended by simply replacing m by $\sqrt{q^2}$.

For space-like momentum ($q^2 < 0$), at least one of the three polarization vectors has to be time-like. We cannot use $\varepsilon^\mu(s)$ of Eq. (C.21c), because it does not satisfy the constraint Eq. (C.18). The condition Eq. (C.18) is satisfied by using $\varepsilon^\mu(l)$ but replacing the mass with $\sqrt{-q^2}$, because

$$\varepsilon(l)_\mu \varepsilon^\mu(l) = \frac{q^2 - \omega^2}{-q^2} = 1 > 0 \quad (\text{C.22a})$$

What about the completeness condition? For space-like q^μ , there is a coordinate system that makes $|q| \neq 0$, $q^0 = 0$, because the energy momentum of the space like virtual photon is created by a charged particle, namely $q = p_i - p_f$ where p_i , p_f are the time-like momenta of the particle before and after emission of the photon. In that coordinate system $\varepsilon^\mu(l) = (1, 0, 0, 0)$ which agrees with $\varepsilon^\mu(0)$ defined in Eq. (C.21). Accordingly the metric for the summation has to be changed and the completeness condition is kept unchanged.

$$\sum_{\lambda=\pm 1, l \rightarrow 0} (-1)^{\lambda+1} \varepsilon^\mu(\lambda) \varepsilon^\nu(\lambda)^* = - \left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2} \right) \quad (\text{C.23})$$

$\varepsilon^\mu(l)$ is time-like and is called scalar polarization.

C.4

Other Useful Formulae

$$\begin{aligned} A + B + C &\equiv (2\pi)^4 \delta^4(P - p_3 - p_4) \frac{d^3 p_3 d^3 p_4}{(2\pi)^6 2 E_3 2 E_4} \\ &\quad \times [a + b_\mu p_3^\mu + c_{\mu\nu} p_3^\mu p_4^\nu] \\ &= \frac{\lambda(1, x, y)}{8\pi} \left[a + b_\mu P^\mu \frac{1+x-y}{2} \right. \\ &\quad \left. + c_{\mu\nu} \frac{1}{2} \left\{ (1-x-y) P^\mu P^\nu - \frac{\lambda^2(1, x, y)}{6} (4 P^\mu P^\nu - P^2 g^{\mu\nu}) \right\} \right] \\ E_3 &= \sqrt{p_3^2 + m_3^2}, \quad E_4 = \sqrt{p_4^2 + m_4^2} \\ \lambda(1, x, y) &= (1 + x^2 + y^2 - 2x - 2y - 2xy)^{1/2}, \quad x = \frac{m_3^2}{p^2}, \quad y = \frac{m_4^2}{p^2} \end{aligned} \quad (\text{C.24})$$

Especially, when $m_3 = m_4 = 0$

$$\begin{aligned} A &= \frac{a}{8\pi}, \quad B = \frac{b_\mu P^\mu}{16\pi}, \\ C &= \frac{c_{\mu\nu}}{96\pi} (2P^\mu P^\nu + P^2 g^{\mu\nu}) \end{aligned} \quad (\text{C.25})$$

To prove Eq. (C.24), use Lorentz invariance to show that

$$A = X(P^2), \quad B = b_\mu P^\mu Y(P^2), \quad C = c_{\mu\nu} [P^\mu P^\nu Z(P^2) + g^{\mu\nu} W(P^2)] \quad (\text{C.26})$$

and perform the calculation in the P rest frame.

Appendix D

Dimensional Regularization

D.1

Photon Self-Energy

We demonstrate the method of dimensional regularization by calculating the photon self-energy loop diagram.

$$-i\Pi^{\mu\nu} = -\int \frac{d^4k}{(2\pi)^4} \text{Tr} \left[(-ie\gamma^\mu) \frac{i(\not{k} + m)}{k^2 - m^2 + i\varepsilon} (-ie\gamma^\nu) \right. \\ \left. \times \frac{i(\not{k} - \not{q} + m)}{(k - q)^2 - m^2 + i\varepsilon} \right] \quad (\text{D.1})$$

and prove that

$$\Pi^{\mu\nu}(q^2) = (g^{\mu\nu}q^2 - q^\mu q^\nu) \Pi_\gamma(q^2) \quad (\text{D.2a})$$

$$\Pi_\gamma(q^2) = \Pi_\gamma(0) + \hat{\Pi}_\gamma(q^2) \quad (\text{D.2b})$$

in which $\hat{\Pi}_\gamma(q^2)$ is a finite analytical function of q^2 .

The trace calculation yields

$$\text{Tr}[\dots] = \text{Tr}[(\not{k} + m)\gamma^\nu(\not{k} - \not{q} + m)\gamma^\mu] \\ = 4[2k^\mu k^\nu - (k^\mu q^\nu + k^\nu q^\mu) - g^{\mu\nu}(k^2 - k \cdot q - m^2)] \\ \therefore -i\Pi^{\mu\nu} = -4e^2 \int \frac{d^4k}{(2\pi)^4} \\ \times \frac{[2k^\mu k^\nu - (k^\mu q^\nu + k^\nu q^\mu) - g^{\mu\nu}(k^2 - k \cdot q - m^2)]}{[k^2 - m^2 + i\varepsilon][(k - q)^2 - m^2 + i\varepsilon]} \quad (\text{D.3})$$

which is a diverging integral for $k \rightarrow \infty$. The above integral Eq. (D.1) contains a $O(k^2)$ term in the numerator, which is quadratically divergent. In dimensional regularization, it appears as a $D = 2$ pole, as the quadratic divergence in $D = 4$ becomes logarithmic in $D = 2$ space. The gauge invariance, however, guarantees that the divergence is at most of logarithmic type, as is shown below. Since the integral is well defined for $D < 2$ and is an analytic function of D , it is allowed to

assume that D is an integer larger than 4, perform the integration formally, then take the limit $D \rightarrow 4$.

To perform the integration in D -dimensional space, we introduce

$$k^\mu = (k^0, k^1, k^2, k^3) \rightarrow (k^0, k^1, \dots, k^{D-1}) \equiv (k^0, \mathbf{k}) \quad (\text{D.4a})$$

$$\frac{d^4 k}{(2\pi)^4} \rightarrow \mu^{4-D} \frac{d^D k}{(2\pi)^D} \quad (\text{D.4b})$$

$$g^{\mu\nu} g_{\mu\nu} = 4 \rightarrow g^{\mu\nu} g_{\mu\nu} = D \quad (\text{D.4c})$$

$$k^2 = (k^0)^2 - \mathbf{k}^2 = (k^0)^2 - \sum_{i=1}^{D-1} (k^i)^2 \quad (\text{D.4d})$$

μ is an energy scale parameter introduced to keep the physical dimension. Although introduction of μ is redundant in the following argument, we keep it because it is indispensable when we extend our discussion to QCD. We introduce the following notation for later convenience:

$$B_0(q^2, m^2) = \int_k \frac{1}{(1)(2)} \quad (\text{D.5a})$$

where we have defined

$$\int_k \equiv \frac{16\pi^2}{i} \mu^{4-D} \int \frac{d^D k}{(2\pi)^D} \quad (\text{D.5b})$$

$$(1) \equiv k^2 - m^2 + i\varepsilon, \quad (2) \equiv (k - q)^2 - m^2 + i\varepsilon \quad (\text{D.5c})$$

Applying the Feynman parameter formula

$$\frac{1}{ab} = \int_0^1 dx \frac{1}{[(1-x)a + xb]^2} \quad (\text{D.6})$$

The integrand can be transformed to a more convenient form for integration

$$\begin{aligned} -i\Pi^{\mu\nu} &= -4e^2 \int_0^1 dx \int \frac{d^4 k}{(2\pi)^4} \\ &\quad \times \frac{[2k^\mu k^\nu - (k^\mu q^\nu + k^\nu q^\mu) - g^{\mu\nu}(k^2 - k \cdot q - m^2)]}{[(k - xq)^2 - (m^2 - x(1-x)q^2)]^2} \end{aligned} \quad (\text{D.7})$$

Let us calculate B_0 first. It is expressed as

$$B_0 = \int_k \frac{1}{(1)(2)} = \frac{16\pi^2}{i} \mu^{4-D} \int_0^1 dx \frac{d^D k}{(2\pi)^D} \frac{1}{[(k - xq)^2 - M^2]^2} \quad (\text{D.8a})$$

$$M^2 = -q^2 x(1-x) + m^2 - i\varepsilon \quad (\text{D.8b})$$

Changing the integration value from $k \rightarrow k' = k - xq$, and introducing $\omega^2 = \sum_{i=1}^{D-1} (k^i)^2 + M^2$, we find the denominator becomes

$$(k^0)^2 - \omega^2 + i\varepsilon = [k^0 - (\omega - i\varepsilon)][k^0 - (-\omega + i\varepsilon)] \quad (\text{D.9})$$

It has poles at $\omega - i\varepsilon$ and $-\omega + i\varepsilon$. Therefore, we can rotate the integral path from $k^0 = -\infty \sim \infty \rightarrow -i\infty \sim i\infty$, which is called the Wick rotation. Then writing

$$k^2 = (k^0)^2 - \sum_{i=1}^{D-1} \rightarrow -r^2 = \sum_{i=0}^{D-1} (k^i)^2 \quad (\text{D.10})$$

As r is the length in the Euclidean space, we can adopt the polar coordinate system and integrate all the angular variables, giving the D -dimensional solid angle Ω_D . It can be calculated as follows. Using the equality

$$\int_{-\infty}^{\infty} dx e^{-x^2} = \sqrt{\pi} \Rightarrow \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_D e^{-x_1^2 \cdots -x_D^2} = \pi^{D/2} \quad (\text{D.11})$$

recalculating it using the polar coordinate

$$\Omega_D \int_0^{\infty} e^{-r^2} r^{D-1} dr = \pi^{D/2} \quad (\text{D.12})$$

Using the definition of the Γ function

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt \quad (\text{D.13})$$

we obtain

$$\Omega_D = \frac{2\pi^{D/2}}{\Gamma(D/2)} \quad (\text{D.14})$$

Now we can perform the integration in Eq. (D.8). At this point, we extend the power of the denominator from two to more general n for later convenience. Then in the Wick rotated coordinate, we obtain

$$\begin{aligned} i \int \frac{d^D k}{(-r^2 - M^2)^2} &\Rightarrow (-1)^n i \int \frac{d^D k}{(r^2 + M^2)^n} = (-1)^n i \Omega_D \int_0^{\infty} \frac{r^{D-1} dr}{(r^2 + M^2)^n} \\ &= (-1)^n i \Omega_D \frac{M^{D-2n}}{2} B\left(\frac{D}{2}, n - \frac{D}{2}\right) \end{aligned} \quad (\text{D.15a})$$

$$B(p, q) = \int_0^{\infty} dt \frac{t^{q-1}}{(1+t)^{p+q}} = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \quad (\text{D.15b})$$

Then

$$B_0 = \frac{16\pi^2 \mu^{4-D}}{i} \frac{(-1)^n i}{(2\pi)^D} \int_0^1 dx \frac{2\pi^{D/2}}{\Gamma(D/2)} \frac{M^{D-2n}}{2} \frac{\Gamma(D/2)\Gamma(n-D/2)}{\Gamma(n)} \quad (\text{D.16a})$$

$$= (-1)^n \int_0^1 dx \left[\frac{(4\pi\mu^2)^{2-D/2}}{(M^2)^{n-D/2}} \right] \frac{\Gamma(n-\frac{D}{2})}{\Gamma(n)} \quad (\text{D.16b})$$

As one can see, $n = 2$, $D = 4$ is a pole of the Γ function. The residue at the pole can be calculated by using infinite product representation for Γ

$$\frac{1}{\Gamma(z)} = z e^{\gamma_E z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n} \quad (\text{D.17})$$

Setting $n = 2$ and defining an infinitesimal number $\varepsilon = 4 - D$, and using

$$\Gamma\left(\frac{\varepsilon}{2}\right) = \frac{2}{\varepsilon} - \gamma_E + O(\varepsilon) \quad (\text{D.18a})$$

$$x^{\varepsilon/2} = e^{(\varepsilon/2)\ln x} = 1 + \frac{\varepsilon}{2} \ln x + O(\varepsilon^2) \quad (\text{D.18b})$$

$$\begin{aligned} B_0(q^2, m^2) &= \Delta - \int_0^1 dx \ln M^2 \\ &= \Delta - \int_0^1 dx \ln[-q^2 x(1-x) + m^2 - i\varepsilon] \end{aligned} \quad (\text{D.19a})$$

$$\Delta = \frac{2}{\varepsilon} - \gamma_E + \ln(4\pi\mu^2) \quad (\text{D.19b})$$

where $\gamma_E \simeq 0.5772$ is the Euler constant. The second term in B_0 is a finite analytical function. The divergent term is included in the first term as the pole of the complex D -plane. γ_E and $\ln(4\pi\mu^2)$ in Δ always accompany $2/\varepsilon$ as a unit and cancel in the observable quantities. Therefore, it is to be subtracted together with $2/\varepsilon$ in the renormalization prescription. In QCD, where there is no special point to normalize the coupling constant with the experimental value, $\ln\mu^2$ is kept as an adjustable parameter.

To calculate other terms in Eq. (D.7), the following formulas are useful and easy to verify:

$$\int \frac{d^D k}{[k^2 - M^2 + i\varepsilon]^n} = i(\pi)^{D/2} \frac{(-1)^n}{(M^2)^{n-D/2}} \frac{\Gamma(n-\frac{D}{2})}{\Gamma(n)} \quad (\text{D.20a})$$

$$\int \frac{k^\mu d^D k}{[k^2 - M^2 + i\varepsilon]^n} = 0 \quad (\text{D.20b})$$

$$\int \frac{k^2 d^D k}{[k^2 - M^2 + i\varepsilon]^n} = i(\pi)^{D/2} (-1)^{n-1} \frac{D/2}{(M^2)^{n-1-D/2}} \frac{\Gamma(n-1-\frac{D}{2})}{\Gamma(n)} \quad (\text{D.20c})$$

$$\int \frac{k^\mu k^\nu d^D k}{[k^2 - M^2 + i\varepsilon]^n} = \frac{g^{\mu\nu}}{D} \int \frac{k^2 d^D k}{[k^2 - M^2 + i\varepsilon]^n} \quad (\text{D.20d})$$

To derive the third equation, $\Gamma(z+1) = z\Gamma(z)$ is used. The last equation can be proved by noting that the integral must be proportional to $g^{\mu\nu}$ by Lorentz invariance and that $g^{\mu\nu}g_{\mu\nu} = D$.

To calculate $\Pi^{\mu\nu}$, we rewrite k in the numerator of Eq. (D.7) with $k'^\mu = k^\mu - xq^\mu$ and drop terms linear in k' as they vanish:

$$\begin{aligned} -i\Pi^{\mu\nu} &= -4e^2 \int_0^1 dx \int \frac{d^4 k}{(2\pi)^4} \\ &\quad \times \frac{2k^\mu k^\nu - (k^\mu q^\nu + k^\nu q^\mu) - g^{\mu\nu}(k^2 - k \cdot q - m^2)}{[(k - xq)^2 - (m^2 - x(1-x)q^2)]^2} \\ &\Rightarrow -4e^2 \int_0^1 dx \int \frac{d^D k'}{(2\pi)^D} \\ &\quad \times \frac{2k'^\mu k'^\nu - 2x(1-x)q^\mu q^\nu - g^{\mu\nu}[k'^2 - q^2 x(1-x) - m^2]}{[k'^2 - M^2 + i\varepsilon]^2} \\ M^2 &= m^2 - x(1-x)q^2 \end{aligned} \quad (\text{D.21})$$

Inserting Eq. (D.20), we obtain

$$\begin{aligned} \Pi^{\mu\nu} &= \frac{4e^2}{16\pi^2} \int_0^1 dx \left[\frac{4\pi\mu^2}{M^2} \right]^{2-D/2} \left[g^{\mu\nu} M^2 \left(\frac{D}{2} - 1 \right) \Gamma \left(1 - \frac{D}{2} \right) \right. \\ &\quad \left. + \left\{ -2x(1-x)q^\mu q^\nu + g^{\mu\nu}[q^2 x(1-x) + m^2] \right\} \Gamma \left(2 - \frac{D}{2} \right) \right] \end{aligned} \quad (\text{D.22a})$$

As $(D/2 - 1)\Gamma(1 - D/2) = -\Gamma(2 - D/2)$, we see that the pole at $D = 2$ which represents the quadratic divergence vanishes. Reduction of the power in the divergent term by the gauge invariance has worked!

$$\begin{aligned} \Pi^{\mu\nu} &= \frac{2\alpha}{\pi} (g^{\mu\nu} q^2 - q^\mu q^\nu) \int_0^1 dx x(1-x) \left[\frac{4\pi\mu^2}{M^2} \right]^{2-D/2} \Gamma \left(2 - \frac{D}{2} \right) \\ &= \frac{2\alpha}{\pi} (g^{\mu\nu} q^2 - q^\mu q^\nu) \int_0^1 dx x(1-x) [\Delta - \ln(M^2 - i\varepsilon)] \\ &= (g^{\mu\nu} q^2 - q^\mu q^\nu) \\ &\quad \times \left[\frac{\alpha}{3\pi} \Delta - \frac{2\alpha}{\pi} \int_0^1 dx x(1-x) \ln[-q^2 x(1-x) + m^2 - i\varepsilon] \right] \end{aligned} \quad (\text{D.23a})$$

Using $\Pi^{\mu\nu} = (g^{\mu\nu} q^2 - q^\mu q^\nu) \Pi_\gamma$, we finally obtain

$$\Pi_\gamma = \Pi_\gamma(0) + \hat{\Pi}_\gamma(q^2) \quad (\text{D.24a})$$

$$\Pi_\gamma(0) = \frac{\alpha}{3\pi} (\Delta - \ln m^2) \quad (\text{D.24b})$$

$$\hat{\Pi}_\gamma(q^2) = -\frac{2\alpha}{\pi} \int_0^1 dx \, x(1-x) \ln \left[1 - \frac{q^2}{m^2} x(1-x) - i\varepsilon \right] \quad (\text{D.24c})$$

D.2

Electron Self-Energy

The electron self energy is given by

$$i\Sigma_e(p) = \int \frac{d^4 k}{(2\pi)^4} \frac{-ig_{\mu\nu}}{(p-k)^2 - \mu^2 + i\varepsilon} (ie\gamma^\mu) \frac{i(\not{k} + m)}{k^2 - m^2 + i\varepsilon} (ie\gamma^\nu) \quad (\text{D.25})$$

where μ is a small mass introduced to avoid the infrared divergence. Using the Feynman parametrization, it becomes

$$\begin{aligned} i\Sigma_e(p) &= -e^2 \int_0^1 dx \int \frac{d^4 k}{(2\pi)^4} \frac{\gamma_\mu \not{k} \gamma^\mu + m \gamma_\mu \gamma^\mu}{[(k-xp)^2 - M^2 + i\varepsilon]^2} \\ &= -2e^2 \int_0^1 dx \int \frac{d^4 k}{(2\pi)^4} \frac{2m - x \not{p}}{[k^2 - M^2 + i\varepsilon]^2} \end{aligned} \quad (\text{D.26a})$$

$$M^2 = -p^2 x(1-x) + (1-x)m^2 + x\mu^2 - i\varepsilon \quad (\text{D.26b})$$

where we have used the relation $\gamma_\mu \not{k} \gamma^\mu = -2 \not{k}$, $\gamma_\mu \gamma^\mu = 4$. This is the same integral as Eq. (D.8) except for the numerator and its result is already given by Eq. (D.19). Thus

$$\Sigma_e(p) = \frac{\alpha}{2\pi} \int_0^1 dx (2m - x \not{p}) [-\Delta + \ln M^2] \quad (\text{D.27})$$

We can easily calculate δm and δZ_2 defined in Eq. (8.68)

$$\Sigma_e(p) = \delta m + [\delta Z_2 + \hat{\Pi}_e(\not{p})](\not{p} - m_R) \quad (\text{D.28a})$$

$$\delta m = \Sigma_e(p)|_{\not{p}=m_R} = \frac{\alpha}{2\pi} m_R \left[-\frac{3}{2} \Delta + \int_0^1 dx (2-x) \ln M_R^2 \right] + O(\alpha^2) \quad (\text{D.28b})$$

In deriving the last equation we used $m = m_R + \delta m = m_R + O(\alpha)$.

$$\begin{aligned} \delta Z_2 &= \left. \frac{\partial \Sigma_\epsilon(p)}{\partial \not{p}} \right|_{\not{p}=m_R} \\ &= \frac{\alpha}{2\pi} \left[+\frac{\Delta}{2} - \int_0^1 dx \left\{ x \ln M_R^2 + \frac{2m_R^2 x(1-x)(2-x)}{M_R^2} \right\} \right] + O(\alpha^2) \end{aligned} \quad (\text{D.28c})$$

$$M_R^2 = (1-x)^2 m_R^2 + x\mu^2 - i\varepsilon \quad (\text{D.28d})$$

In calculating M_R^2 and its derivative, we have used $p^2 = \not{p}\not{p}$ and

$$\begin{aligned} M_R^2 &= M^2|_{\not{p}=m_R} = -m_R^2 x(1-x) + (1-x)(m_R + \delta m)^2 + x\mu^2 - i\varepsilon \\ &= (1-x)^2 m_R^2 + x\mu^2 - i\varepsilon + O(\alpha) \end{aligned} \quad (\text{D.29})$$

We see, the first derivative of $\Sigma_\epsilon(p)$ still includes the divergent term. But $\hat{\Pi}_\epsilon(\not{p})$ is finite and satisfies the relation

$$\hat{\Pi}_\epsilon(\not{p})|_{\not{p}=m_R} = 0 \quad (\text{D.30})$$

Appendix E

Rotation Matrix

E.1

Angular Momentum Operators

An operator to rotate a physical object by an angle α around the axis having a unit vector $\mathbf{n}(\theta, \phi)$ is given by

$$R = \exp[-\alpha(\mathbf{n} \cdot \mathbf{J})] \quad (\text{E.1})$$

The angular momentum operator $J_i = L_i + S_i$ satisfies the commutation relation

$$[J_i, J_j] = i\epsilon_{ijk}J_k \quad (\text{E.2})$$

$J^2 = J_1^2 + J_2^2 + J_3^2$ commutes with all J_i , and has an eigenvalue $j(j+1)$, $j = 0, 1/2, 1, \dots$. One of J_i (conventionally J_3) and J^2 can be diagonalized simultaneously, so the eigenvector can be specified by two values, j, j_3 . Writing $J_{\pm} = J_1 \pm iJ_2$

$$\begin{aligned} J^2|jm\rangle &= j(j+1)|jm\rangle \\ J_3|jm\rangle &= m|jm\rangle \\ J_{\pm}|jm\rangle &= \sqrt{j(j+1) - m(m \pm 1)}|jm \pm 1\rangle \end{aligned} \quad (\text{E.3})$$

For spinless particle, $\mathbf{J} = \mathbf{L}$ and

$$L_i = x_j p_k - x_k p_j = -i(x_j \partial_k - x_k \partial_j) \quad (\text{E.4})$$

Eigenfunctions of the orbital angular momentum are spherical harmonic functions. They are, for $L^2 = L(L+1)$, $L_3 = M$, $-L \leq M \leq L$, given as

$$Y_{LM}(\theta, \phi) = (-1)^{(M+|M|)/2} \left[\frac{2L+1}{4\pi} \frac{(L-|M|)!}{(L+|M|)!} \right]^{1/2} P_L^{|M|}(\cos \theta) e^{iM\phi} \quad (\text{E.5a})$$

$$P_L^{|M|}(z) = (1-z)^{|M|/2} \frac{1}{2^L L!} \frac{d^{L+|M|}}{dz^{L+|M|}} (z^2 - 1)^L \quad (\text{E.5b})$$

$$Y_{L,-M}(\theta, \phi) = (-1)^M [Y_{LM}(\theta, \phi)]^* \quad (\text{E.5c})$$

$$\begin{aligned}
Y_{00} &= \sqrt{\frac{1}{4\pi}}, \quad Y_{10} = \sqrt{\frac{3}{4\pi}} \cos \theta, \quad Y_{1\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \\
Y_{20} &= \sqrt{\frac{5}{4\pi}} \frac{3 \cos^2 \theta - 1}{2}, \quad Y_{2\pm 1} = \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi}, \\
Y_{2\pm 2} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi} \\
Y_{30} &= \sqrt{\frac{7}{4\pi}} \frac{5 \cos^3 \theta - 3 \cos \theta}{2}, \\
Y_{31} &= \mp \sqrt{\frac{21}{64\pi}} \sin \theta (5 \cos^2 \theta - 1) e^{\pm i\phi} \\
Y_{32} &= \sqrt{\frac{105}{32\pi}} \sin^2 \theta \cos \theta e^{\pm 2i\phi}, \quad Y_{33} = \mp \sqrt{\frac{35}{64\pi}} \sin^3 \theta e^{\pm 3i\phi}
\end{aligned} \tag{E.6}$$

For $L = 0$, $J = S$, and S_i are $(2S + 1) \times (2S + 1)$ matrices. For $S = 1/2$, $S = \sigma/2$, where σ is the Pauli spin matrix.

$$S_1 = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad S_2 = \frac{1}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad S_3 = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{E.7}$$

The eigenfunctions are

$$\chi_{+1/2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \chi_{-1/2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{E.8}$$

For $S = 1$, we can adopt eigenfunctions

$$e_{+1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_{-1} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{E.9}$$

For which we can easily construct $S_{\pm}$ using Eq. (E.3)

$$S_+ = \begin{bmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{bmatrix}, \quad S_- = \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix}, \tag{E.10}$$

Then

$$\begin{aligned}
S_1 &= \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad S_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \\
S_3 &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}
\end{aligned} \tag{E.11}$$

E.2

Addition of the Angular Momentum

Adding J_1 and J_2 vectorially

$$J = J_1 + J_2 \quad (\text{E.12})$$

The range of the eigenvalues J, M are

$$|j_1 - j_2| \leq J \leq j_1 + j_2, \quad M = m_1 + m_2 \quad (\text{E.13})$$

The eigenvectors of the J are given by

$$|JM\rangle = \sum_{m_1+m_2=M} C_{j_1 m_1, j_2 m_2}^{JM} |j_1 m_1\rangle |j_2 m_2\rangle \quad (\text{E.14})$$

$C_{j_1 m_1, j_2 m_2}^{JM}$ are called the Clebsch–Gordan coefficients and their general forms are given in [334]. For the case, $j_2 = 1/2, 1$ they are given by

Table E.1 $C_{j_1, M-m; 1/2, m}^{JM}$

$J \backslash m$	$m = 1/2$	$m = -1/2$
$j + 1/2$	$\left[\frac{j_1 + M + 1/2}{2j_1 + 1} \right]^{1/2}$	$\left[\frac{j_1 - M + 1/2}{2j_1 + 1} \right]^{1/2}$
$j - 1/2$	$-\left[\frac{j_1 - M + 1/2}{2j_1 + 1} \right]^{1/2}$	$\left[\frac{j_1 + M + 1/2}{2j_1 + 1} \right]^{1/2}$

Table E.2 $C_{j_1, M-m; 1, m}^{JM}$

$J \backslash m$	$m = 1$	$m = 0$	$m = -1$
$j_1 + 1$	$\left[\frac{(j_1 + M)(j_1 + M + 1)}{(2j_1 + 1)(2j_1 + 2)} \right]^{1/2}$	$\left[\frac{(j_1 - M + 1)(j_1 + M + 1)}{(2j_1 + 1)(j_1 + 1)} \right]^{1/2}$	$\left[\frac{(j_1 - M)(j_1 - M + 1)}{(2j_1 + 1)(2j_1 + 2)} \right]^{1/2}$
j_1	$-\left[\frac{(j_1 + M)(j_1 - M + 1)}{2j_1(j_1 + 1)} \right]^{1/2}$	$\frac{M}{[j_1(j_1 + 1)]^{1/2}}$	$\left[\frac{(j_1 - M)(j_1 + M + 1)}{2j_1(j_1 + 1)} \right]^{1/2}$
$j_1 - 1$	$\left[\frac{(j_1 - M)(j_1 - M + 1)}{2j_1(2j_1 + 1)} \right]^{1/2}$	$-\left[\frac{(j_1 - M)(j_1 + M)}{j_1(2j_1 + 1)} \right]^{1/2}$	$\left[\frac{(j_1 + M + 1)(j_1 + M)}{2j_1(2j_1 + 1)} \right]^{1/2}$

E.3

Rotational Matrix

The effect of a rotation on the z-axis simply gives a phase factor to the eigenfunction $|j m\rangle$, but that on y-axis generates states of different m . As, $|j m\rangle$ makes a complete

system if the value of j does not change, the new state can be expanded by their linear combination:

$$\therefore e^{-i\theta J_2} |j m\rangle = \sum_n d_{nm}^j(\theta) |j n\rangle \quad (\text{E.15})$$

$d_{nm}^j(\theta)$ is called the rotation matrix and is defined by

$$d_{nm}^j(\theta) = \langle j n | e^{-i\theta J_2} | j m \rangle \quad (\text{E.16})$$

The properties of d_{nm}^j function are

$$\begin{aligned} d_{nm}^j(\theta) &= d_{-m, -n}^j(\theta) = (-1)^{n-m} d_{mn}^j(\theta) \\ d_{nm}^j(\theta) &= (-1)^{j+n} d_{n, -m}^j(\pi - \theta) \\ d_{nm}^j(\pi) &= (-1)^{j+n} \delta_{n, -m} \\ d_{nm}^j(-\pi) &= (-1)^{n-j} \delta_{n, -m} \end{aligned} \quad (\text{E.17})$$

The normalization is given by

$$\int d \cos \theta d_{nm}^{j_1}(\theta) d_{nm}^{j_2}(\theta) = \frac{2}{2j_1 + 1} \delta_{j_1, j_2} \quad (\text{E.18})$$

A general formula of rotation matrices are given by Wigner [334], but for $j = 1/2, 1$, it is easily obtained by inserting Eqs. (E.7) and (E.11) into Eq. (E.16):

$$\begin{aligned} \begin{bmatrix} d_{1/2, 1/2}^{1/2} & d_{1/2, -1/2}^{1/2} \\ d_{-1/2, 1/2}^{1/2} & d_{-1/2, -1/2}^{1/2} \end{bmatrix} &= e^{-i(\frac{\theta}{2})\sigma_2} = \sum_n \frac{1}{n!} \left(-i \frac{\theta}{2} \sigma_2 \right)^n \\ &= \cos \frac{\theta}{2} - \sin \frac{\theta}{2} (i\sigma_2) = \begin{bmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix} \end{aligned} \quad (\text{E.19})$$

Similarly, for $S = 1$

$$\begin{bmatrix} d_{1,1}^1 & d_{1,0}^1 & d_{1,-1}^1 \\ d_{0,1}^1 & d_{0,0}^1 & d_{0,-1}^1 \\ d_{-1,1}^1 & d_{-1,0}^1 & d_{-1,-1}^1 \end{bmatrix} = \begin{bmatrix} \frac{1+\cos \theta}{2} & -\frac{\sin \theta}{\sqrt{2}} & \frac{1-\cos \theta}{2} \\ \frac{\sin \theta}{\sqrt{2}} & \cos \theta & -\frac{\sin \theta}{\sqrt{2}} \\ \frac{1-\cos \theta}{2} & \frac{\sin \theta}{\sqrt{2}} & \frac{1+\cos \theta}{2} \end{bmatrix} \quad (\text{E.20})$$

Slightly more general formulae are given in [217].

$j = L = \text{Integer}$

$$\begin{aligned} d_{00}^L &= P_L(\theta) \\ d_{10}^L(\theta) &= -[L(L+1)]^{-1/2} \sin \theta P_L'(\theta) \\ d_{20}^L(\theta) &= [(L-1)L(L+1)(L+2)]^{-1/2} [2P_{L-1}'(\theta) - L(L-1)P_L(\theta)] \\ P_L'(z) &= \frac{d}{dz} P_L(z) \end{aligned} \quad (\text{E.21})$$

$$j = L + 1/2$$

$$\begin{aligned}
 d_{1/2,1/2}^j(\theta) &= (j + 1/2)^{-1} \cos \frac{\theta}{2} (P'_{j+1/2} - P'_{j-1/2}) \\
 d_{-1/2,1/2}^j(\theta) &= (j + 1/2)^{-1} \sin \frac{\theta}{2} (P'_{j+1/2} + P'_{j-1/2}) \\
 d_{1/2,3/2}^j(\theta) &= (j + 1/2)^{-1} \sin \frac{\theta}{2} \left(\sqrt{\frac{j-1/2}{j+3/2}} P'_{j+1/2} + \sqrt{\frac{j+3/2}{j-1/2}} P'_{j-1/2} \right) \\
 d_{-1/2,3/2}^j(\theta) &= (j + 1/2)^{-1} \sin \frac{\theta}{2} \left(-\sqrt{\frac{j-1/2}{j+3/2}} P'_{j+1/2} + \sqrt{\frac{j+3/2}{j-1/2}} P'_{j-1/2} \right)
 \end{aligned}
 \tag{E.22}$$

Appendix F

C, P, T Transformation

We summarize here the transformation properties of fields. Only the Dirac fields and their bilinears are given. Corresponding scalars, vectors etc. that obey the Klein–Gordon equations are similar to the corresponding Dirac bilinears. $|q, p, s_z\rangle$ means the state of a particle (q) or antiparticle ($\bar{q}$) having momentum $\mathbf{p}$, and spin component $s = s_z$. $(t, \mathbf{x})$ stands for $(t, \mathbf{x})$. Combined operations usually give different phases depending on the order of the operation. But the Dirac bilinears do not because $\eta\eta^* = 1$.

$$\begin{aligned}
 P : \quad & (t, \mathbf{x}) \rightarrow (t, -\mathbf{x}) \\
 & P|q, p, s\rangle = \eta_P|q, -p, s\rangle \\
 & \psi(t, \mathbf{x}) \rightarrow \gamma^0\psi(t, -\mathbf{x}) \\
 & \bar{\psi}(t, \mathbf{x}) \rightarrow \bar{\psi}(t, -\mathbf{x})\gamma^0 \\
 S \quad & \bar{\psi}_1\psi_2 \rightarrow \bar{\psi}_1\psi_2 \\
 P \quad & i\bar{\psi}_1\gamma^5\psi_2 \rightarrow -i\bar{\psi}_1\gamma^5\psi_2^* \\
 V \quad & \bar{\psi}_1\gamma^\mu\psi_2 \rightarrow \bar{\psi}_1\gamma^\mu\psi_2 \\
 A \quad & \bar{\psi}_1\gamma^\mu\gamma^5\psi_2 \rightarrow -\bar{\psi}_1\gamma^\mu\gamma^5\psi_2 \\
 T \quad & \bar{\psi}_1\sigma^{\mu\nu}\psi_2 \rightarrow \bar{\psi}_1\sigma_{\mu\nu}\psi_2
 \end{aligned} \tag{F.1}$$

* i is attached to make $i\bar{\psi}\gamma^5\psi$ hermitian.

*

$$\begin{aligned}
 C : \quad & C|q, p, s\rangle = \eta_C|\bar{q}, p, s\rangle \\
 & \psi(t, \mathbf{x}) \rightarrow C\bar{\psi}^T(t, \mathbf{x}) = i\gamma^2\gamma^0\bar{\psi}^T(t, \mathbf{x}) \\
 & \bar{\psi}(t, \mathbf{x}) \rightarrow -\psi^T(t, \mathbf{x})C^{-1} \\
 & C^{-1}\gamma^\mu C = -\gamma^{\mu T} \\
 & C = -C^\dagger = -C^{-1} = -C^T = C^* \\
 S \quad & \bar{\psi}_1\psi_2 \rightarrow \bar{\psi}_2\psi_1 \\
 P \quad & i\bar{\psi}_1\gamma^5\psi_2 \rightarrow i\bar{\psi}_2\gamma^5\psi_1 \\
 V \quad & \bar{\psi}_1\gamma^\mu\psi_2 \rightarrow -\bar{\psi}_2\gamma^\mu\psi_1 \\
 A \quad & \bar{\psi}_1\gamma^\mu\gamma^5\psi_2 \rightarrow \bar{\psi}_2\gamma^\mu\gamma^5\psi_1 \\
 T \quad & \bar{\psi}_1\sigma^{\mu\nu}\psi_2 \rightarrow -\bar{\psi}_2\sigma^{\mu\nu}\psi_1
 \end{aligned} \tag{F.2}$$

$$\begin{aligned}
T : \quad & (t, x) \rightarrow (-t, x) \\
& (\text{c-number}) \rightarrow (\text{c-number})^* \\
& \langle f : \text{out} | H | i : \text{in} \rangle = \langle f : \text{out} | T^{-1} T H T^{-1} T | i : \text{in} \rangle \\
& = \langle T f : \text{in} | H | T i : \text{out} \rangle^* \\
& T | q, p, s \rangle = \eta_T | q, -p, -s \rangle \\
& \psi(t, x) \rightarrow B \psi(-t, x) \\
& = i \gamma^1 \gamma^3 \psi(-t, x) = -i \gamma^5 C \psi(-t, x) \\
& \bar{\psi}(t, x) \rightarrow \bar{\psi}(-t, x) B^{-1} = \bar{\psi}(-t, x) \gamma^1 \gamma^3 \\
& B^{-1} \gamma^{\mu*} B = \gamma_{\mu}, \quad B^{\dagger} = B^{-1} = -B^T = B
\end{aligned} \tag{F.3}$$

$$\begin{aligned}
S \quad & \bar{\psi}_1 \psi_2 \rightarrow \bar{\psi}_1 \psi_2 \\
P \quad & i \bar{\psi}_1 \gamma^5 \psi_2 \rightarrow -i \bar{\psi}_1 \gamma^5 \psi_2 \\
V \quad & \bar{\psi}_1 \gamma^{\mu} \psi_2 \rightarrow \bar{\psi}_1 \gamma_{\mu} \psi_2 \\
A \quad & \bar{\psi}_1 \gamma^{\mu} \gamma^5 \psi_2 \rightarrow \bar{\psi}_1 \gamma_{\mu} \gamma^5 \psi_2 \\
T \quad & \bar{\psi}_1 \sigma^{\mu\nu} \psi_2 \rightarrow -\bar{\psi}_1 \sigma_{\mu\nu} \psi_2
\end{aligned}$$

$$\begin{aligned}
CP : \quad & (t, x) \rightarrow (t, -x) \\
& CP | q, p, s \rangle = \eta_{CP} | \bar{q}, -p, s \rangle \\
& \psi(t, x) \rightarrow \pm i \gamma^2 \bar{\psi}^T(t, -x) \\
& \bar{\psi}(t, x) \rightarrow \mp \psi^T(t, -x) i \gamma^2 \\
S \quad & \bar{\psi}_1 \psi_2 \rightarrow \bar{\psi}_2 \psi_1 \\
P \quad & i \bar{\psi}_1 \gamma^5 \psi_2 \rightarrow -i \bar{\psi}_2 \gamma^5 \psi_1 \\
V \quad & \bar{\psi}_1 \gamma^{\mu} \psi_2 \rightarrow -\bar{\psi}_2 \gamma_{\mu} \psi_1 \\
A \quad & \bar{\psi}_1 \gamma^{\mu} \gamma^5 \psi_2 \rightarrow -\bar{\psi}_2 \gamma_{\mu} \gamma^5 \psi_1 \\
T \quad & \bar{\psi}_1 \sigma^{\mu\nu} \psi_2 \rightarrow -\bar{\psi}_2 \sigma_{\mu\nu} \psi_1
\end{aligned} \tag{F.4}$$

$$\begin{aligned}
CPT : \quad & (t, x) \rightarrow (-t, -x) \\
& (\text{c-number}) \rightarrow (\text{c-number})^* \\
& \langle f : \text{out} | H | i : \text{in} \rangle = \langle f : \text{out} | \Theta^{-1} \Theta H \Theta^{-1} \Theta | i : \text{in} \rangle \\
& = \langle \Theta f : \text{in} | H | \Theta i : \text{out} \rangle^* \\
& \Theta | q, p, s \rangle = \eta_{\Theta} | \bar{q}, p, -s \rangle \\
& \psi(t, x) \rightarrow \pm i \gamma_0 \gamma^5 \bar{\psi}^T(t, -x) \\
& \bar{\psi}(t, x) \rightarrow \pm \psi^T(t, -x) i \gamma_0 \gamma^5 \\
S \quad & \bar{\psi}_1 \psi_2 \rightarrow \bar{\psi}_2 \psi_1 \\
P \quad & i \bar{\psi}_1 \gamma^5 \psi_2 \rightarrow i \bar{\psi}_2 \gamma^5 \psi_1 \\
V \quad & \bar{\psi}_1 \gamma^{\mu} \psi_2 \rightarrow -\bar{\psi}_2 \gamma^{\mu} \psi_1 \\
A \quad & \bar{\psi}_1 \gamma^{\mu} \gamma^5 \psi_2 \rightarrow -\bar{\psi}_2 \gamma^{\mu} \gamma^5 \psi_1 \\
T \quad & \bar{\psi}_1 \sigma^{\mu\nu} \psi_2 \rightarrow \bar{\psi}_2 \sigma^{\mu\nu} \psi_1
\end{aligned} \tag{F.5}$$

Appendix G

$SU(3)$, $SU(n)$ and the Quark Model¹⁾

G.1

Generators of the Group

The set of all the linear transformations that preserve the length $L = \sqrt{\sum_{i=1}^N |u_i|^2}$ in the N -dimensional coordinate space of N complex variables $u_i (i = 1 \sim N)$ is called $U(N)$. With the further restriction of the determinant being 1, it is called $SU(N)$, a special unitary group. The number of independent variables is N^2 for $U(N)$ and $N^2 - 1$ for $SU(N)$. Any representation matrix of the $SU(3)$ group operating on n -dimensional vectors can be expressed as $n \times n$ unitary matrices, which is a function of $N^2 - 1$ independent continuous variables θ_i and traceless hermitian matrices F_i

$$U = \exp \left(i \sum_{i=1}^{N^2-1} \theta_i F_i \right), \quad F_i^\dagger = F_i, \quad \text{Tr}[F_i] = 0, \quad (\text{G.1})$$

The traceless condition makes the determinant of the matrix equal to one. It can be proved as follows. If we write $U = e^{iQ}$, Q is a hermitian matrix. Therefore, it can be diagonalized by using some unitary matrix $S^{-1}QS = \text{diag}\{Q\}$. Then

$$\begin{aligned} \det U &= \det[S^{-1} \exp(iQ) S] = \det \left[\sum S^{-1} \frac{(iQ)^n}{n!} S \right] \\ &= \det \left[\sum \frac{(S^{-1}iQS)^n}{n!} \right] = \det[\text{diag}\{e^{iQ}\}] = \prod_i e^{i\eta_i} \\ &= \exp \left(i \sum_i \eta_i \right) = e^{i\text{Tr}[Q]} \quad (= 1 \text{ if } \text{Tr}[Q] = 0) \end{aligned} \quad (\text{G.2})$$

where η_i 's are eigenvalues of the diagonalized matrix Q . $N^2 - 1$ F_i 's satisfy commutation relations called Lie algebra:

$$[F_i, F_j] = i f_{ijk} F_k \quad (\text{G.3})$$

1) References are [89, 101, 156, 165]

The F_i 's are called the generators of the group and f_{ijk} are structure constants. Since all the representation matrices are analytic functions smoothly connected to the unit matrix in the limit of $\theta_i \rightarrow 0$, the Lie algebra Eq. (G.3) completely determines the local structure of the group.²⁾ f_{ijk} can be made totally antisymmetric in the indices ijk . Determination of the generator is not unique, because different F'_i can be made of from any independent linear combination of F_i .³⁾

Let us write ψ as representing an n -dimensional ($n \geq N$) contravariant vector (with upper index) on which the $n \times n$ $SU(N)$ matrices act. Then ψ is transformed to ψ' by U .

$$\psi = \begin{bmatrix} \xi^1 \\ \xi^2 \\ \vdots \\ \xi^n \end{bmatrix}, \quad \psi' = \begin{bmatrix} \xi^{1'} \\ \xi^{2'} \\ \vdots \\ \xi^{n'} \end{bmatrix}, \quad n \geq N \quad (\text{G.5a})$$

$$\psi \rightarrow \psi' = U\psi \quad (\text{G.5b})$$

$$\xi^{a'} = U^a_b \xi^b, \quad U = e^{-i\theta_i F_i} \quad (\text{G.5c})$$

where simultaneous appearance of indices means sum over the indices up to n . The collection of n matrices with $n = N$ is called a fundamental representation and adjoint (or regular) representation if $n = N^2 - 1$.

G.1.1

Adjoint Representation

The representation matrices with $n = N^2 - 1$ can be expressed in terms of the structure constants.

$$U(N^2 - 1) = [F_a]_{bc} = -i f_{abc}, \quad (a, b, c = 1 \sim N^2 - 1) \quad (\text{G.6})$$

Proof: Insert $A = [F_a]$, $B = [F_b]$, $C = [F_c]$ in the Jacobi identities

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0 \quad (\text{G.7})$$

use Eq. (G.3) and $f_{abc} = -f_{bac}$, then we obtain

$$f_{abd} f_{cde} + f_{bcd} f_{ade} + f_{cad} f_{bde} = 0 \quad (\text{G.8})$$

Then insertion of Eq. (G.6) into the above equality leads to

$$[F_a, F_c] = i f_{acd} F_d \quad (\text{G.9})$$

Therefore, F_a defined in Eq. (G.6) is an adjoint representation of the $SU(N)$. $\square$

2) An example of global structure is "connectedness". For example, $SU(2)$ and $O(3)$ have the same structure constants, but $SU(2)$ is simply connected and $O(3)$ is doubly connected and there is a one-to-two correspondence between the two.

3) For instance, instead of angular momentum J_i , we can use $J_{\pm} = J_x \pm J_y$, which satisfies $[J_+, J_-] = 2J_z$, $[J_z, J_{\pm}] = \pm J_{\pm}$ (G.4)

If the Lagrangian or the equation of motion is invariant under $SU(N)$ transformation, F_k is a conserved quantity. The number of matrices that can be diagonalized simultaneously is called the rank of the group. The rank of $SU(N)$ is $N - 1$. Matrices that are made of F_k and commute with all F_k 's are called Casimir operators. Particles which can be identified as the representation vector of a group, can be specified and classified by the eigenvalues of the Casimir operators and F_i 's that can be diagonalized simultaneously.

Representation matrix of the complex conjugate vector ψ^* is given by U^* . Since its generator $-F_k^*$ satisfies the same commutation relation Eq. (G.3) also, ψ^* is also a representation vector and is called a conjugate representation. If ψ represents a particle, then ψ^* represents its antiparticle. Therefore, the quantum numbers of antiparticles are minus those of particles. If we define a covariant vector (with lower indices) by

$$\psi^\dagger = (\psi^*)^T = (\xi^{1*}, \xi^{2*}, \dots, \xi^{N*}) \equiv (\xi_1, \xi_2, \dots, \xi_N) \quad (G.10)$$

It transforms as

$$\xi_a \rightarrow \xi_a' = \sum U^{a*}_b \xi_b^* = \sum U_{ab} \xi_b = \sum \xi_b U_{ba}^\dagger \equiv (U^\dagger)^b_a \xi_b \quad (G.11a)$$

$$\text{or } \psi^\dagger \rightarrow \psi^{\dagger'} = \psi^\dagger U^\dagger \quad (G.11b)$$

If a representation matrix $U(A)$ of any element A can be expressed as

$$U(A) = \begin{bmatrix} U_1(A) & 0 & 0 \\ 0 & U_2(A) & 0 \\ 0 & 0 & \ddots \end{bmatrix} \quad (G.12)$$

It is said reducible and written as

$$U = U_1 \oplus U_2 \oplus \dots \oplus U_k \quad (G.13)$$

G.1.2

Direct Product

When there are two representations $U(A)$, $U(B)$ of dimensionality n_A , n_B which have elements $A = \{A_j\}$, $B = \{B_k\}$. Product of basic elements ξ_j , η_k of each representation $\xi_j \eta_k$ spans $(n_A \times n_B)$ -dimensional space. The representation in this space $U(AB)$ is expressed as

$$U(AB)(\xi_j \eta_k) = \xi_j' \eta_k' = (U(A)\xi_j)(U(B)\eta_k) \quad (G.14)$$

and is called a direct product representation. It is written as

$$U(AB) = U(A) \otimes U(B) \quad (G.15)$$

In general the direct product is reducible.

Example G.1

Each representation of the angular momentum $J = J_a$ and $J = J_b$ has $2J_a + 1$ and $2J_b + 1$ elements. The representation of the direct product $J_a \otimes J_b$ has $(2J_a + 1)(2J_b + 1)$ elements. It is reducible and decomposed to $J = J_1 \oplus J_2 \oplus \cdots \oplus J_n$ where $J_1 = J_a + J_b$, $J_2 = J_a + J_b - 1, \dots, J_n = |J_a - J_b|$.

G.2 $SU(3)$

G.2.1

Structure Constants

It is customary to adopt Pauli matrices for $SU(2)$ ($F_i = \sigma_i/2$) and Gell-Mann's λ matrices for $SU(3)$ ($F_i = \lambda_i/2$). Although $SU(3)$ itself is a general mathematical framework, it will be convenient for us to identify the three light quarks $q^i = (u, d, s)$ as the $SU(3)$ fundamental representation state vector and express them as

$$u = \xi^1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad d = \xi^2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad s = \xi^3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (\text{G.16})$$

Since $u \leftrightarrow d$ [= $SU(2)$ transformation] is a part of $SU(3)$ transformation, we define λ_i which operates only on u and d .

$$\lambda_i = \begin{bmatrix} \tau_i & 0 \\ 0 & 0 \end{bmatrix}, \quad i = 1 \sim 3 \quad (\text{G.17a})$$

Then λ_i 's are isospin operators. In a similar manner, we define $SU(2)$ operators "V spin" that act only on $u \leftrightarrow s$, and "U spin" that act only on $d \leftrightarrow s$.

$$\begin{aligned} \lambda_4 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \\ V_3 &= \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1/2 \end{bmatrix} \\ \lambda_6 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \\ U_3 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & -1/2 \end{bmatrix} \end{aligned} \quad (\text{G.17b})$$

We have defined nine matrices, but

$$I_3 - V_3 + U_3 = 0 \quad (\text{G.17c})$$

and are not independent. We define

$$\lambda_8 \equiv \sqrt{3}Y = \frac{2}{\sqrt{3}}(U_3 + V_3) = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \quad (\text{G.17d})$$

The eigenvalue of the λ_8 is called hypercharge and is denoted as “Y”. The number of diagonal matrices is two, which is the rank of $SU(3)$. The eight λ_i satisfy the following commutation relations

$$\left[\frac{\lambda_i}{2}, \frac{\lambda_j}{2} \right] = i f_{ijk} \frac{\lambda_k}{2} \quad (\text{G.18a})$$

$$\left\{ \frac{\lambda_i}{2}, \frac{\lambda_j}{2} \right\} = \frac{1}{3} \delta_{ij} + i d_{ijk} \frac{\lambda_k}{2} \quad (\text{G.18b})$$

f_{ijk} , d_{ijk} are totally antisymmetric or symmetric with respect to $i j k$ permutations and have values listed in Table G.1

Table G.1 Structure constants of $SU(3)$.

$f_{123} = 1$
$f_{147} = f_{246} = f_{257} = f_{345} = f_{516} = f_{637} = \frac{1}{2}$
$f_{458} = f_{678} = \frac{\sqrt{3}}{2}$
$d_{118} = d_{228} = d_{338} = -d_{888} = \frac{1}{\sqrt{3}}$
$d_{146} = d_{157} = d_{256} = d_{344} = d_{355} = \frac{1}{2}$
$d_{247} = d_{366} = d_{377} = -\frac{1}{2}$
$d_{448} = d_{558} = d_{668} = d_{778} = -\frac{1}{2\sqrt{3}}$

Some Definitions of Terminology Let $F_a(n)$ be an n -dimensional representation matrix,

$$\text{Tr}[F_a(n)F_b(n)] = T_R(n)\delta_{ab} \quad (\text{G.19a})$$

$$\sum_{a=1}^{N^2-1} F_a(n)^2 = C_2(n)\mathbf{1} \quad (\text{G.19b})$$

By convention, the normalization of $SU(N)$ matrix is chosen to be $T_F \equiv T_R(n = N) = 1/2$. With this choice $SU(N)$ matrices satisfy relations

$$T_F \equiv T_R(N_F) = \frac{1}{2}, \quad T_R(N_A) = N \quad (\text{G.20a})$$

$$C_F \equiv C_2(N_F) = \frac{N^2 - 1}{2N}, \quad C_A \equiv C_2(N_A) = N \quad (\text{G.20b})$$

where suffix F denotes fundamental ($n = N_F = N$) and A denotes adjoint ($n = N_A = N^2 - 1$). A general method to obtain $C_2(n)$ is given in Sect. G.2.3.

V, U Spin Defining raising or lowering operators of I , V , U spins by

$$\begin{aligned} I_{\pm} &= \frac{\lambda_1 \pm \lambda_2}{2}, \quad V_{\pm} = \frac{\lambda_4 \pm \lambda_5}{2}, \quad U_{\pm} = \frac{\lambda_6 \pm \lambda_7}{2} \\ 2I_3 &= [I_+, I_-], \quad 2V_3 = [V_+, V_-], \quad 2U_3 = [U_+, U_-] \end{aligned} \quad (\text{G.21})$$

In the quark model, $q = (u, d, s)$ has quantum numbers

$$\begin{aligned} I_3(q) &= (1/2, -1/2, 0) \\ V_3(q) &= (1/2, 0, -1/2) \\ U_3(q) &= (0, 1/2, -1/2) \\ Y(q) &= (1/3, 1/3, -2/3) \end{aligned} \quad (\text{G.22})$$

If we place the three quarks in a I_3 - Y plane and represent each q as a vector from the origin (referred to as a weight diagram), when we make representations of direct products of the state vectors, multiplication of u means an increase of I_3 , Y by $1/2$, $1/3$, respectively, which translates to an addition of the u -vector in the I_3 - Y plane. Similarly, the multiplication of d, s means the addition of d, s vectors. Actions of $I_{\pm}, V_{\pm}, U_{\pm}$ are depicted in Figure G.2. For instance the action of I_+ changes d to u , which means moving a point in the I_3 - Y plane by 1 unit to the right. Conversely, I_- changes u to d and moves a point by 1 unit to the left. Similarly, the action of $V_{\pm}$ is to change $s(u)$ to $u(s)$ or to move a point diagonally, as is depicted in Fig. G.2a.

G.2.2

Irreducible Representation of a Direct Product

a. Meson Octet Let us consider a composite system, spin 0 mesons, made of quarks and antiquarks. In the terminology of $SU(3)$ group representation, we are making direct products of q^i and $\bar{q}_j$, or $\Phi_j^i = q^i \bar{q}_j$. By a unitary transformation

$$\begin{aligned} q &\rightarrow q' = Uq, \quad \bar{q} \rightarrow \bar{q}' = \bar{q}U^\dagger \\ \text{or } [\Phi] &\equiv [\Phi]_j^i \rightarrow \Phi' = U\Phi U^\dagger \end{aligned} \quad (\text{G.23})$$

which means a product $\text{Tr}(\Phi) = \bar{q}q = \bar{u}u + \bar{d}d + \bar{s}s$ is invariant under $SU(3)$ transformation.

$$\therefore \eta' \equiv \frac{1}{\sqrt{3}}(\bar{u}u + \bar{d}d + \bar{s}s) \quad (\text{G.24})$$

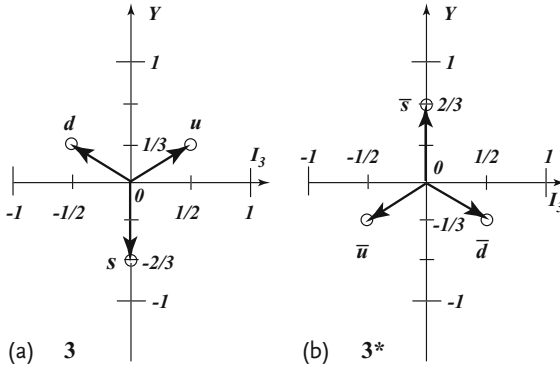


Figure G.1 (a) (u, d, s) make a triplet $\mathbf{3}$ and have $(I_3, Y) = (1/2, 1/3), (-1/2, 1/3), (0, -2/3)$. They are expressed as vectors from the origin. (b) $(\bar{u}, \bar{d}, \bar{s})$ make $\mathbf{3}^*$. Vectors $(\bar{u}, \bar{d}, \bar{s})$ have opposite quantum numbers to those of (u, d, s) .

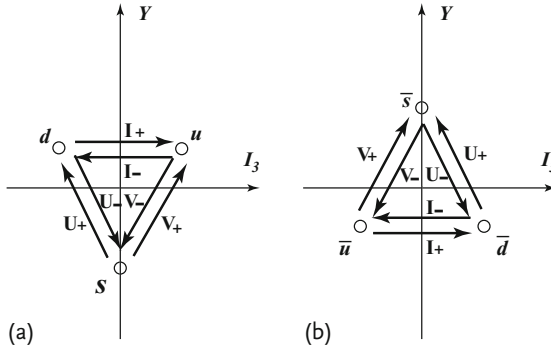


Figure G.2 $I_{\pm}, V_{\pm}, U_{\pm}$ move the state in the direction the arrows indicate.

is a vector that transforms to itself and hence makes singlet. The remaining eight vectors mix together by the transformation and cannot be separated further and hence makes an multiplet of 8 members called octet. They are expressed as

$$P_j^i = q^i \bar{q}_j - \frac{1}{3} \sum q^i \bar{q}_i = \Phi_j^i - \frac{1}{3} \text{Tr}[\Phi_j^i] \quad (\text{G.25})$$

The conclusion of the above discussion is that the direct representation made of three q 's and three $\bar{q}$'s can be decomposed into groups of one and eight members. We express it as

$$\mathbf{3} \otimes \mathbf{3}^* = \mathbf{1} \oplus \mathbf{8} \quad (\text{G.26})$$

Let us derive the same equation from a slightly different point of view. Base vectors of the direct product $\mathbf{3} \otimes \mathbf{3}^*$ are nine vectors, which are obtained by multiplying u, d, s in $\mathbf{3}$ and $\bar{u}, \bar{d}, \bar{s}$ in $\mathbf{3}^*$. Take the example of multiplying u by $\bar{d}$. As u has $(I_3, Y) = (1/2, 1/3)$ and $\bar{d}$, $(1/2, -1/3)$, the product $u\bar{d}$ has $(I_3, Y) = (1, 0)$. In

the I_3 - Y plane, it is obtained by summing the two vectors u and $\bar{d}$, as shown in Fig. G.3a. Other operations are similar. To repeat, first start from the origin. We place three vectors of (u, d, s) as in Fig. G.3b. Then add another three vectors of $\bar{u}, \bar{d}, \bar{s}$ from each of the three apexes, and we obtain a total of nine states. Three of them, $u\bar{u}, d\bar{d}, s\bar{s}$, come back to where they started, i.e. there is a triple point at $I_3 = Y = 0$. One combination of them $\eta' = \text{Tr}[\Phi]/\sqrt{3}$ forms a singlet, as we already explained (or according to a general rule Sect. G.2.2.c) and should be separated. The result is $3 \otimes 3^* = 1 \oplus 8$ as depicted in Fig. G.3c.

b. Baryon Octet and Decuplet Baryons are made of three quarks. We first make two-quark composites made of nine product vectors $T^{ij} = \xi^i \xi^j$, and see their transformation properties. They are, in general, reducible, but we can separate them into symmetric and antisymmetric parts:

$$T^{ij} = \frac{1}{2}(\xi^i \xi^j - \xi^j \xi^i) + \frac{1}{2}(\xi^i \xi^j + \xi^j \xi^i) = A^{ij} + S^{ij} \quad (\text{G.27})$$

As the permutation symmetry does not change by the transformation, A^{ij} and S^{ij} transform to themselves. The nine product vectors have been decomposed into a three-dimensional antisymmetric group and a six-dimensional symmetric group. Namely

$$3 \otimes 3 = 3^* \oplus 6 \quad (\text{G.28})$$

In the vectorial representation, they are expressed as Figure G.4a. That the three A^{ij} belongs to 3^* and not to 3 can be easily seen by looking at their quantum num-

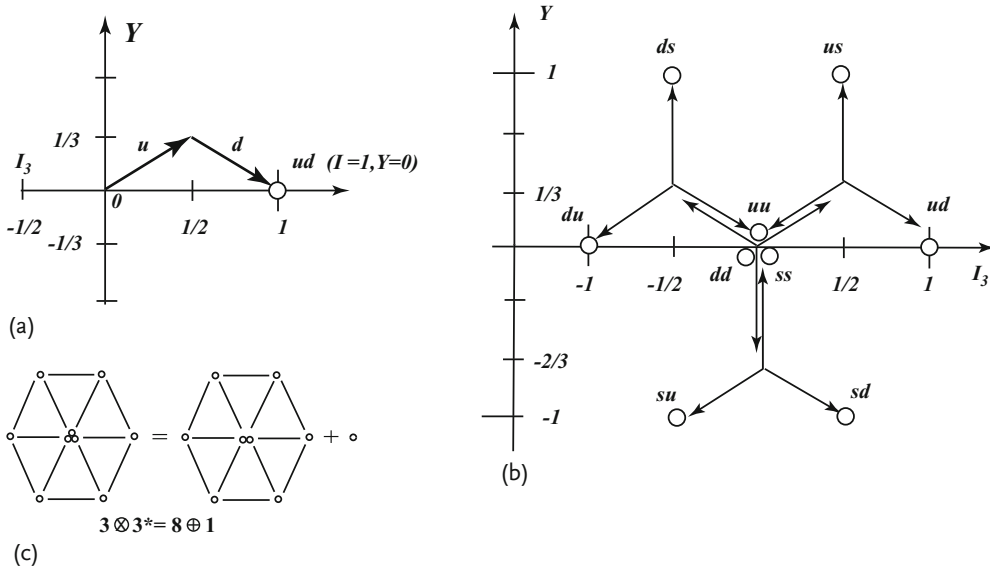


Figure G.3 (a) As $u, \bar{d}$ has $(I_3, Y) = (1/2, 1/3), (1/2, -1/3)$ $u\bar{d}$ can be obtained as the vector sum of the two states and has $(1, 0)$ in the I_3 - Y plane. (b) Applying the same operation to all combinations of $(u, d, s) \times (\bar{u}, \bar{d}, \bar{s})$, a total of nine states are obtained. (c) $3 \otimes 3^* = 8 \oplus 1$.

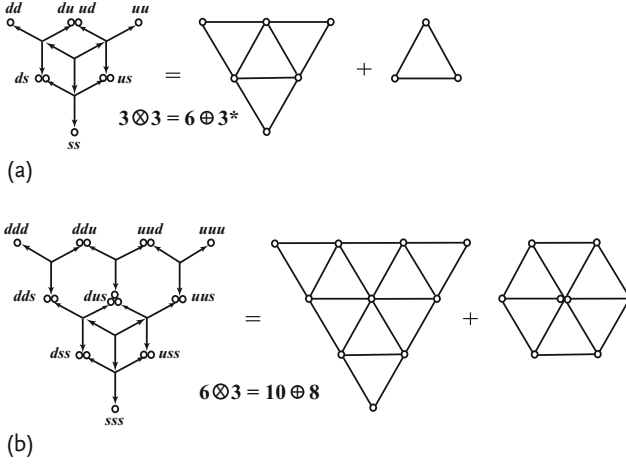


Figure G.4 Baryon multiplets. (a) $3 \otimes 3 = 6 \oplus 3^*$, (b) $6 \otimes 3 = 10 \oplus 8$.

bers which are the same as $(\bar{u}, \bar{d}, \bar{s})$. Mathematically, it can be shown as follows. Define

$$\eta_i \equiv \varepsilon_{ijk} A^{jk} \quad (\text{G.29})$$

then

$$B = \eta_i \xi^i \quad (\text{G.30})$$

is invariant under $SU(3)$ transformation, because

$$\begin{aligned} B \xrightarrow{U} B' &= \eta'_i \xi'^i = \varepsilon_{ijk} \xi'^i \xi'^j \xi'^k = \varepsilon_{ijk} U_l^i U_m^j U_n^k \xi^l \xi^m \xi^n \\ &= \varepsilon_{lmn} (\det U) \xi^l \xi^m \xi^n = \varepsilon_{lmn} \xi^l \xi^m \xi^n = B \end{aligned} \quad (\text{G.31})$$

From Eq. (G.31), the vector η_i 's transformation matrix is shown to be $(U^{-1})^T = U^*$, which means η_i belongs to 3^* . Multiplying $T^{ij} = 3^* \oplus 6$ by ξ^k in 3 further, we get from the first term of Eq. (G.28) ($A^{ij} \xi^k$)

$$3^* \otimes 3 = 1 \oplus 8 \quad (\text{G.32})$$

just like we obtained meson singlet and octet. But unlike the meson case, the singlet term in Eq. (G.32) is a totally antisymmetric combination of ξ^i 's made from $A^{ij} \xi^k$ and the remaining 8 terms constitute an octet. Product vectors of the second term of S^{ij} and ξ^k can be separated to totally symmetric part of which there are 10 members and the remaining 8 members which form another octet. Namely

$$6 \otimes 3 = 10 \oplus 8 \quad (\text{G.33})$$

To summarize, we have derived

$$\begin{aligned} 3 \otimes 3 \otimes 3 &= 1 \oplus 8 \oplus 8 \oplus 10 \\ \text{symmetry property : } A \quad MA \quad MS \quad S \end{aligned} \quad (\text{G.34})$$

Here, the singlet **1** consists of a totally antisymmetric and the decuplet **10** of totally symmetric combinations of $\xi^{i'}$ s. The two octets **8** occupy the same position in $I_3 - Y$ plane, but they have different permutation symmetry. As we have seen, one is symmetric (MS = mixed symmetry) and the other is antisymmetric (MA = mixed antisymmetric) with respect to the first two indices. Weight diagrams of these multiplets are depicted in Figure G.4.

c. General Rules of the Irreducible Weight Diagrams A way to decompose a product representation into a sum of irreducible representations is given by using the Young diagram (Sect. G.2.4). There is a theorem concerning the structure of irreducible weight diagrams.

■ Theorem G.1

Any weight diagram of irreducible representations must obey the following rules (proof omitted).

1. A weight diagram makes layers of convex polygons, generally irregular hexagons, with triangles in a special case (Figure G.4 and Figure G.5).
2. Multiplicity of points on the sides and apexes of the outermost layer is one.
3. Multiplicity of those points on the second layer is two and increases by one as we go into inner layers.
4. When the polygon of the inner layers becomes a triangle, the multiplicity stops to increase and stay at the same value no matter how deep we go into inner layers. As a special case, the multiplicity of all the points on triangle weight diagram is one.

d. Baryon Wave Functions To construct wave functions of **1**, **8_{MA}**, **8_{MS}**, **10_S**, let us first consider those without the s -quark. Since, u, d have isospin $1/2$, the wave functions can be constructed using the Clebsch–Gordan coefficients of $SU(2)$. As $1/2 \otimes 1/2 = \mathbf{0} \oplus \mathbf{1}$ with wave functions

$$I = 0 : \quad |0, 0\rangle = \frac{1}{\sqrt{2}}(ud - du), \quad I = 1 : \quad \begin{cases} |1, +1\rangle &= uu \\ |1, 0\rangle &= \frac{1}{\sqrt{2}}(ud + du) \\ |1, -2\rangle &= dd \end{cases} \quad (\text{G.35})$$

They have the desired property being antisymmetric and symmetric with exchange of the first pair of quarks. Adding the third quark to the wave functions, we can make $I = 1/2$ from the first $I = 0$ and $I = 1/2$ and $3/2$ from the second $I = 1$. Using Clebsch–Gordan coefficients we obtain

$$I = 1/2_{MA} : \quad \begin{cases} |1/2, +1/2\rangle &= \frac{1}{\sqrt{2}}(ud - du)u \\ |1/2, -1/2\rangle &= \frac{1}{\sqrt{2}}(ud - du)d \end{cases} \quad (\text{G.36a})$$

$$I = 1/2_{MS} : \quad \begin{cases} |1/2, +1/2\rangle &= \frac{1}{\sqrt{6}}\{2uud - (ud + du)u\} \\ |1/2, -1/2\rangle &= -\frac{1}{\sqrt{6}}\{2ddu - (ud + du)d\} \end{cases} \quad (\text{G.36b})$$

$$I = 3/2_S : \begin{cases} |3/2, +3/2\rangle &= uuu \\ |3/2, +1/2\rangle &= \frac{1}{\sqrt{3}}(uud + udu + duu) \\ |3/2, -1/2\rangle &= \frac{1}{\sqrt{3}}(udd + dud + ddu) \\ |3/2, -3/2\rangle &= ddd \end{cases} \quad (\text{G.36c})$$

There is no totally antisymmetric combination. It can only be made using all three u, d, s . Obviously, $I = 1/2_{MA}$, $I = 1/2_{MS}$, $I = 3/2$ wave functions are members of $\mathbf{8}_{MA}$, $\mathbf{8}_{MS}$, $\mathbf{10}_S$. Since totally symmetric wave functions of $\mathbf{10}$ can be made easily, we list other wave functions in Table G.2.

There are some ambiguity in constructing the neutral ($I_3 = Y = 0$) wave functions. As the isospin is a good quantum number, we make Σ^0 from Σ^+ by applying I_- and Λ^0 by making orthogonal to Σ^0 .

G.2.3

Tensor Analysis

a. Dimension of an Irreducible Representation A general product vector or a tensor which is a product of contra-variant vectors ξ^i and q covariant vectors ξ_j is written as $T_{(q)}^{(p)} = T_{lmn\dots}^{ijk\dots}$ and transforms like

$$(T_{lmn\dots}^{ijk\dots})' = U_a^i U_b^j \dots U_l^{\dagger c} U_m^{\dagger d} \dots T_{cd\dots}^{ab\dots} \quad (\text{G.37})$$

It is called a contra-variant tensor of rank p if $q = 0$, and a covariant tensor of rank q if $p = 0$ and a mixed tensor if $p, q \neq 0$. We write the whole collection of the tensors as (p, q) . The weight diagram of (p, q) is a irregular hexagon which has

Table G.2 Baryon wave functions with mixed symmetry ϕ_{MA} , ϕ_{MS} and total antisymmetry ϕ_A .

	ϕ_{MA}	ϕ_{MS}
p	$\frac{1}{\sqrt{2}}(ud - du)u$	$\frac{1}{\sqrt{6}}\{2uud - (ud + du)u\}$
n	$\frac{1}{\sqrt{2}}(ud - du)d$	$-\frac{1}{\sqrt{6}}\{2ddu - (ud + du)d\}$
Σ^+	$\frac{1}{\sqrt{2}}(us - su)u$	$\frac{1}{\sqrt{6}}\{2uus - (us + su)u\}$
Σ^0	$\frac{1}{2}\{(usd + dsu) - s(ud + du)\}$	$\frac{1}{\sqrt{12}}\{2(ud + du)s - (usd + dsu) - s(ud + du)\}$
Σ^-	$\frac{1}{\sqrt{2}}(ds - sd)d$	$\frac{1}{\sqrt{6}}\{2dds - (ds + sd)d\}$
Λ^0	$\frac{1}{\sqrt{12}}\{2(ud - du)s + (usd - dsu) - s(ud - du)\}$	$\frac{1}{2}\{(usd - dsu) + s(ud - du)\}$
Ξ^0	$\frac{1}{\sqrt{2}}(su - us)s$	$\frac{1}{\sqrt{6}}\{2ssu - (us + su)s\}$
Ξ^-	$\frac{1}{\sqrt{2}}(sd - ds)s$	$\frac{1}{\sqrt{6}}\{2ssd - (ds + sd)s\}$
	ϕ_A	
Λ_1^0	$\frac{1}{\sqrt{6}}\{s(ud - du) + (dsu - usd) + (ud - du)s\}$	

upper and lower sides of length p and q (see Figure G.5). If p or q is 0, then the hexagon becomes a triangle.

We can prove that $\delta_j^i = \delta_{ij}$ is a mixed tensor of $p = q = 1$, and that ε_{ijk} , ε^{ijk} are covariant and contra-variant tensors of rank 3. Therefore, the product of a tensor by them is also a tensor. In general, the tensor (p, q) in Eq. (G.37) is reducible. By making contractions with δ_j^i , ε_{ijk} , ε^{ijk} we can make tensors of lower rank. For instance,

$$R_{bc\dots}^{jk\dots} = \delta_i^a T_{abc\dots}^{ijk\dots} = T_{ibc\dots}^{ijk\dots} \quad (\text{G.38a})$$

$$S_{abcd\dots}^{k\dots} = \varepsilon_{aij} T_{bcd\dots}^{ijk\dots} \quad (\text{G.38b})$$

are tensors of $(p-1, q-1)$, $(p-2, q+1)$. If the tensor before contraction is already traceless or symmetrized, the above operation gives zero result which means it is no longer reducible. Therefore, a tensor belonging to an irreducible representation is symmetrized with respect to upper or lower indices and traceless. Therefore, it satisfies a condition

$$T_{ilm\dots}^{ij\dots} = 0 \quad (\text{G.39})$$

As the upper and lower indices are already symmetrized, all we need is to take a trace of the first index. Expressing the representation of this tensor space as $D^N(p, q)$, where N denotes the number of independent elements and is given by

$$N = \frac{1}{2}(p+1)(q+1)(p+q+2) \quad (\text{G.40})$$

Proof: We arrange upper indices like

$$\underbrace{11\dots 22\dots 33\dots}_{\substack{a \quad b \\ p}} \quad (\text{G.41})$$

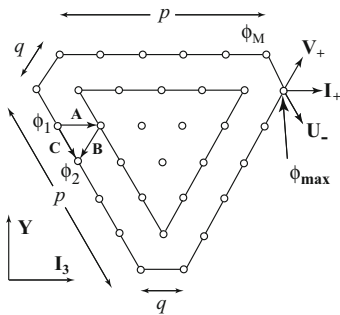


Figure G.5 The weight diagram of (p, q) is an irregular hexagon which has upper and lower sides of length p and q . Multiplicity of points at periphery is one, and increases by one as one goes one step inside until one reaches a triangle. From there it stays constant. For

multiple points, the configuration of the wave function is not unique and depends on the path one chooses. For a single multiplicity point, the configuration is unique. For instance, the same ϕ_2 can be obtained from ϕ_1 through path $A \rightarrow B$ or C .

For a ($0 \leq a \leq p$), b can take $0 \sim p - a$ positions which means $(p - a + 1)$ ways to choose, giving N_u ways to arrange upper indices.

$$N_u = \sum_{a=0}^p (p - a + 1) = \frac{1}{2}(p + 1)(p + 2) \quad (\text{G.42})$$

Likewise there are $N_d = (q + 1)(q + 2)/2$ ways to arrange lower indices, giving a total of

$$N_u N_d = \frac{1}{4}(p + 1)(p + 2)(q + 1)(q + 2) \quad (\text{G.43})$$

There are as many traceless conditions Eq. (G.39) as $N_u N_d$ with p, q replaced with $p - 1, q - 1$, so the number of independent combinations is

$$\begin{aligned} N &= \frac{1}{4}(p + 1)(p + 2)(q + 1)(q + 2) - \frac{1}{4}p(p + 1)q(q + 1) \\ &= \frac{1}{2}(p + 1)(q + 1)(p + q + 2) \end{aligned} \quad (\text{G.44})$$

□

A few examples:

$$\begin{aligned} D^1(0, 0) &= \mathbf{1}, & D^3(1, 0) &= \mathbf{3}, & D^3(0, 1) &= \mathbf{3}^*, \\ D^8(1, 1) &= \mathbf{8}, & D^6(2, 0) &= \mathbf{6}, & D^6(0, 2) &= \mathbf{6}^*, \\ D^{10}(3, 0) &= \mathbf{10}, & D^{10}(0, 3) &= \mathbf{10}^*, \dots \end{aligned} \quad (\text{G.45})$$

Next, we will show that the multiplicity of the outermost layer is one. As

$$\begin{aligned} I_3 &= \frac{1}{2}[n(u) + n(d) - n(\bar{u}) - n(\bar{d})] \\ Y &= \frac{1}{3}[n(u) + n(d) - 2n(s) - n(\bar{u}) - n(\bar{d}) + 2n(\bar{s})] \\ p &= n(u) + n(d) + n(s) \\ q &= n(\bar{u}) + n(\bar{d}) + n(\bar{s}) \end{aligned} \quad (\text{G.46})$$

the value of $n(q)$ is uniquely defined when one specifies p (or q), I_3 and Y if q (or p) = 0.

b. Value of the Casimir Operators How to derive $C_2 \equiv C_2(n) = \mathbf{F}^2 = \sum_k (F_k)^2$. First we rewrite the Casimir operator in terms of I, V, U spin operators.

$$\begin{aligned} \mathbf{F}^2 &= \frac{1}{2}\{I_+, I_-\} + I_3^2 + \frac{1}{2}\{V_+, V_-\} + \frac{1}{2}\{U_+, U_-\} + \frac{3}{4}Y^2 \\ &= (I_- I_+ + V_- V_+ + U_- U_+) + I_3^2 + 2I_3 + \frac{3}{4}Y^2 \end{aligned} \quad (\text{G.47})$$

As the polygon is always convex in the $I_3 - Y$ plane, for the element $\phi_{\max}$ with the highest value of I_3

$$I_+ \phi_{\max} = V_+ \phi_{\max} = U_- \phi_{\max} = 0 \quad (\text{G.48})$$

Therefore, if we know the values of I_3 and Y of the $\phi_{\max}$,

$$\langle \mathbf{F}^2 \rangle = \langle I_3 \rangle^2 + 2\langle I_3 \rangle + \frac{3}{4}Y^2 \quad (\text{G.49})$$

We can obtain the maximum value of I_3 by assuming all the p quarks are u and all the q antiquarks are $\bar{d}$, which gives $I_3 = (p + q)/2$. Then $Y = (p - q)/3$, giving

$$\langle \mathbf{F}^2 \rangle = \frac{1}{3}(p^2 + pq + q^2) + (p + q) \quad (\text{G.50})$$

Some examples are:

$$\begin{aligned} C_2(1) &= \langle \mathbf{F}^2 \rangle_1 = (p = q = 0) &= 0 \\ C_2(3) &= \langle \mathbf{F}^2 \rangle_3 = (p = 1, q = 0) = \langle \mathbf{F}^2 \rangle_{3^*} = \frac{4}{3} \\ C_2(6) &= \langle \mathbf{F}^2 \rangle_6 = (p = 2, q = 0) &= \frac{10}{3} \\ C_2(8) &= \langle \mathbf{F}^2 \rangle_8 = (p = q = 1) &= 3 \\ C_2(10) &= \langle \mathbf{F}^2 \rangle_{10} = (p = 3, q = 0) &= 6 \end{aligned} \quad (\text{G.51})$$

G.2.4

Young Diagram

We describe here, a convenient method to decompose a $SU(N)$ product representation to a sum of irreducible representations without proof.

G.2.4.1 Rules for Making Young Diagrams

1. Young diagram makes a correspondence between an N -dimensional fundamental representation with a box, and its complex representation N^* with a column of $N - 1$ boxes.

$$\mathbf{N} = \square \quad \bar{\mathbf{N}} = \left. \begin{array}{c} \square \\ \square \\ \vdots \\ \square \end{array} \right\} (N - 1 \text{ boxes}) \quad (\text{G.52})$$

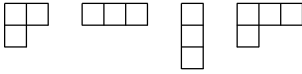
For example

$$\begin{aligned} SU(2) : \quad \square &= \mathbf{2} \text{ or } \mathbf{2}^* \\ SU(3) : \quad \square &= \mathbf{3}, \quad \begin{array}{c} \square \\ \square \end{array} = \mathbf{3}^* \end{aligned} \quad (\text{G.53})$$

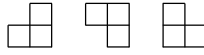
2. Any irreducible representation is expressed by columns and rows of n boxes. The arrangement of the n boxes has to be convex toward the upper left. It means if the number of boxes in the first row is λ_1 , that of the second row $\lambda_2, \dots$ and that of the last row λ_m , then

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m, \quad \lambda_1 + \lambda_2 + \dots + \lambda_m = n \quad (\text{G.54})$$

Allowed arrangements



Forbidden arrangements



3. Any permutation of boxes in a row is meant to be symmetric and that in a column to be antisymmetric. Because of this property, the maximum size of the column is limited to $N - 1$, because in $SU(N)$ the maximum number of elements that are antisymmetric is N and there is only one combination to make N elements antisymmetric. If a column of $m (\geq N)$ boxes does not exist, the dimension is the same, so we may omit it. Pictorial examples are, for $N = 4$

$$N \left\{ \begin{array}{c} \square \\ \square \\ \square \\ \square \end{array} \right\} = \bullet \quad \text{One-dimensional representation} \quad \begin{array}{c} \square \square \\ \square \square \\ \square \square \\ \square \end{array} = \begin{array}{c} \square \square \\ \square \end{array} \quad (G.55)$$

4. For any representation, there is a conjugate representation. Combined together, they make a $N \times N$ representation. For instance, for $SU(3)$

$$(a) \begin{array}{c} \square \\ \square \end{array} \quad (b) \begin{array}{c} \square \square \\ \square \square \\ \square \end{array} \quad (c) \begin{array}{c} \square \square \\ \square \square \\ \square \end{array} = (d) \begin{array}{c} \square \square \\ \square \end{array} = (a) \quad (G.56)$$

As we can see from (b), (c) is conjugate to (a), but applying the condition Eq. (G.55), (c)=(d), which coincides with (a), in other words, $\mathbf{8} = \mathbf{8}^*$ in $SU(3)$.

G.2.4.2 Decomposition of Product Representations of A and B

1. Insert 1 in all the boxes of the first row of the representation **B**, 2 in all the boxes of the second row, and continue to m in the m -th (last) row.
2. Starting from the first row of **B**, pick any number of boxes and attach them to **A** in all possible ways subject to conditions
 - i. Keep the condition “convex toward the upper left”.
 - ii. In any row, the number in any box must not decrease from left to right.
 - iii. In any column, all the boxes must have a different number and it must increase downwards.
- iv. Draw a line starting from the box at the extreme right to the first box on the first row. Move to the second row and enter the box again at the extreme right and continue, until the line passes all the boxes. At any box on the line, count the number (n_i) of boxes on the path which contains the number i . The rule to obey is $n_i \geq n_{i+1}$.

Example G.2

Correct prescription

$$\begin{array}{|c|c|} \hline & \\ \hline & \\ \hline \end{array} \otimes \begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline & 1 & \\ \hline 2 & & \\ \hline \end{array} \oplus \begin{array}{|c|c|} \hline & \\ \hline 1 & \\ \hline 2 & \\ \hline \end{array}$$

Incorrect prescription

$$\begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline \end{array} \quad \begin{array}{|c|} \hline & \\ \hline 2 & \\ \hline 1 & \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline & 2 & 1 \\ \hline & & \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline & 1 & 2 \\ \hline & & \\ \hline \end{array}$$

(a) (b) (c) (d)

Figure G.6 Diagrams (a) and (b) do not satisfy condition (i). Diagram (c) does not satisfy condition (ii) and (d) does not satisfy (iv), because on a line entering the first row from right and stopping at the first box, $n_1 (= 0) < n_2 (= 1)$.

Example G.3

$$\begin{array}{|c|c|} \hline & \\ \hline & \\ \hline \end{array} \otimes \begin{array}{|c|c|} \hline 1 & 1 \\ \hline 2 & \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline & 1 & 1 \\ \hline 2 & & \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline & 1 & 1 \\ \hline & 2 & \\ \hline & & \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline & & 1 \\ \hline & 1 & 2 \\ \hline & & \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline & & 1 \\ \hline & 1 & 1 \\ \hline & 2 & \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline & & 1 \\ \hline & 1 & 2 \\ \hline & & \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline & & 1 \\ \hline & 2 & \\ \hline & & \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline & & 1 \\ \hline & 1 & 2 \\ \hline & & \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline & & 1 \\ \hline & 1 & 2 \\ \hline & & \\ \hline \end{array} \oplus \dots$$

$$\therefore 8 \otimes 8 = 27 \oplus 10 \oplus 10^* \oplus 8 \oplus 8 \oplus 1 \quad (\text{G.57})$$

There are two Young diagrams corresponding to the same 8 , but their permutation properties are different, hence they correspond to different representations.

G.2.4.3 Dimension of the Irreducible Representation

The dimension of a representation is given by

$$D = \frac{F}{H} \quad (\text{G.58})$$

Calculation of F : Insert N in the box at the upper left corner. Increase the number by one as you move to the right ($N + 1, N + 2, \dots$). Decrease it by one as you go downward, for instance, insert $N - 1$ in the first box in the second row. See Fig. G.7a. D is given as product of all the numbers in the boxes.

N	N+1	N+2	N+3
N-1	N	N+1	
N-2	N-1	N	
N-3			

$$F = N(N+1)(N+2)(N+3) \times (N-1)N(N+1) \times (N-2)(N-1)N \times (N-3)$$

$$D = F/H$$

$h_1 = 4$
 $h_2 = 3$
 $h_3 = 2$
 $h_4 = 1$
 $h_5 = 1$

 $H = h_1 \times h_2 \times h_3 \times h_4 \times h_5$

Figure G.7 The dimension of an irreducible representation using the Young diagram is given by F/H .

Calculation of H : Draw a line in any row starting from the extreme right, bend it at any box with a right angle downward and exit from below. This is called a hook. Let h be the number of boxes in the path of the hook; H is given as the product of all possible h 's (Fig. G.7b).

Example G.4

Dimension of  in $SU(3)$ is given in Figure G.8.

$$\begin{array}{|c|c|} \hline 3 & 4 \\ \hline 2 & \\ \hline \end{array} \quad F=3 \times 4 \times 2 \quad H=3 \times 1 \times 1 \quad \begin{array}{|c|c|c|} \hline \bullet & & \bullet \\ \hline & & \\ \hline \bullet & & \\ \hline \end{array} \quad \begin{array}{l} h_1 = 3 \\ h_2 = 1 \\ h_3 = 1 \end{array}$$

$$D=F/H=8$$

Figure G.8 An example of $SU(3)$ representation. $F = 3 \times 4 \times 2 = 24$, $H = 3 \times 1 \times 1 = 3$. Therefore, $D = F/H = 8$.

Example G.5

$SU(3)$:

$$\begin{array}{l}
 3^* \otimes 3 = 1 \oplus 8 \qquad 3 \otimes 3 = 6 \oplus 3^* \qquad 6 \otimes 3 = 10 \oplus 8 \\
 \begin{array}{|c|} \hline \\ \hline \end{array} \otimes \begin{array}{|c|} \hline \\ \hline \end{array} = \begin{array}{|c|} \hline \\ \hline \end{array} \oplus \begin{array}{|c|c|} \hline & \\ \hline \end{array} \quad \begin{array}{|c|} \hline \\ \hline \end{array} \otimes \begin{array}{|c|} \hline \\ \hline \end{array} = \begin{array}{|c|c|} \hline & \\ \hline \end{array} \oplus \begin{array}{|c|} \hline \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline & \\ \hline \end{array} \otimes \begin{array}{|c|} \hline \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline & & \\ \hline \end{array} \oplus \begin{array}{|c|c|} \hline & \\ \hline \end{array}
 \end{array}$$

$SU(4)$: Hadron multiplets that are composed of four quarks are easily constructed using the Young diagram.

$$\begin{array}{ll}
 \text{Mesons :} & 4 \otimes 4^* = 1 \oplus 15 \\
 \text{Baryons :} & 4 \otimes 4 \otimes 4 = 4 \oplus 20 \oplus 20 \oplus 20 \\
 \text{Symmetry} & \quad \quad \quad A \quad MS \quad MA \quad S
 \end{array} \tag{G.59}$$

The symmetry property is the same as that of $SU(3)$. The reason what was a singlet in $SU(3)$ now is a quartet, is there are four ways to make a totally antisymmetric three-quark wave function out of four quarks. An octet and decuplet in $SU(3)$ have become **20** adding baryons containing charm quarks.

$SU(5)$

$$\begin{array}{ll}
 5 \otimes 5^* & = 1 \oplus 25 \\
 5 \otimes 5 & = 10 \oplus 15 \\
 & \quad \quad \quad A \quad S \\
 5^* \otimes 10 & = 5 \oplus 45 \\
 5 \otimes 10 & = 40 \oplus 40 \oplus 10^* \\
 10 \otimes 10 & = 5^* \oplus 45^* \oplus 50 \\
 10^* \otimes 10 & = 1 \oplus 24 \oplus 75
 \end{array} \tag{G.60}$$

$SU(6)$

$$\begin{array}{ll}
 \text{Mesons :} & 6 \otimes 6^* = 1 \oplus 35 \\
 \text{Baryons :} & 6 \otimes 6 \otimes 6 = 20 \oplus 70 \oplus 70 \oplus 56 \\
 \text{Symmetry} & \quad \quad \quad A \quad MS \quad MA \quad S
 \end{array} \tag{G.61}$$

Appendix H

Mass Matrix and Decaying States

H.1

The Decay Formalism

We describe here the time evolution of unstable particles in the Weisskopf–Wigner approximation [385, 386]. The arguments here closely follow those of [220, 283]. We consider a system that is described by a Hamiltonian

$$H = H_0 + H_W \quad (\text{H.1})$$

where H_0 is the Hamiltonian of the strong interaction and free part of the weak interaction while H_W is that of the weak interaction that induces transition and is assumed to be a small perturbation to H_0 . The eigenstate of H_0 consists of i degenerate discrete states $|A_j\rangle$ and a series of continuum states $|B_n\rangle$ to which A can decay weakly. Since we want to apply the formalism to the neutral K meson system, we set $i = 2$, $|A_{1(2)}\rangle = |K^0(\bar{K}^0)\rangle$ and $|B_n\rangle = 2\pi, 3\pi, \pi^\pm \ell^\mp \nu, \dots$

$$\begin{aligned} H_0|A_j\rangle &= E_0|A_j\rangle \\ H_0|B_n\rangle &= E_n|B_n\rangle \end{aligned} \quad (\text{H.2})$$

As an initial state at $t = 0$ we consider a superposition of the discrete states $|A\rangle$.

$$|t = 0\rangle = a_1(0)|A_1\rangle + a_2(0)|A_2\rangle \quad (\text{H.3})$$

The time evolution of the state vector $|t\rangle$ obeys the Schrödinger equation.

$$i \frac{d}{dt}|t\rangle_S = (H_0 + H_W)|t\rangle_S \quad (\text{H.4})$$

where the subscript S denotes that the states are those in the Schrödinger picture. They have time dependence $|t\rangle_S = e^{-iEt}|t = 0\rangle_S$. In the following discussions, it is more convenient to work in the interaction picture:

$$|t\rangle \equiv |t\rangle_I = e^{iH_0t}|t\rangle_S \quad (\text{H.5})$$

For simplicity, we omit the subscript I for the interaction picture. The state $|t\rangle$ obeys the equation

$$i \frac{d}{dt}|t\rangle = H_W(t)|t\rangle, \quad H_W(t) = e^{iH_0t} H_W e^{-iH_0t} \quad (\text{H.6})$$

We write

$$\begin{aligned} |t=0\rangle &= |t=0\rangle_s \\ |t\rangle &= a_1(t)|A_1\rangle + a_2(t)|A_2\rangle + \sum_n b_n(t)|B_n\rangle \end{aligned} \quad (\text{H.7})$$

Inserting Eq. (H.7) in Eq. (H.6) we obtain

$$\begin{aligned} i \frac{\partial a_j(t)}{\partial t} \Big|_{j=1,2} &= \sum_{k=1,2} \langle A_j | H_W | A_k \rangle a_k(t) \\ &\quad + \sum_n e^{i(E_0 - E_n)t} \langle A_j | H_W | B_n \rangle b_n(t) \end{aligned} \quad (\text{H.8a})$$

$$\begin{aligned} i \frac{\partial b_m(t)}{\partial t} &= \sum_{k=1,2} e^{i(E_m - E_0)t} \langle B_m | H_W | A_k \rangle a_k(t) \\ &\quad + \sum_n e^{i(E_m - E_n)t} \langle B_m | H_W | B_n \rangle b_n(t) \end{aligned} \quad (\text{H.8b})$$

Equations (H.8) are exact. We impose initial conditions such that

$$\begin{aligned} a_1(0) &= a_1, \quad a_2(0) = a_2 \\ b_m(0) &= 0 \end{aligned} \quad (\text{H.9})$$

Weisskopf–Wigner’s first approximation (1) is that with the initial conditions we imposed the weak interactions in the final states are negligible, which is equivalent to discarding the second term in Eq. (H.8b). This means we regard pions and muons, which are the decay products of the K mesons, as stable. With this approximation, Eqs. (H.8) can be solved. We obtain

$$b_m(t) = -i \sum_{k=1,2} \int_0^t dt' e^{i(E_m - E_0)t'} \langle m | H_W | k \rangle a_k(t') \quad (\text{H.10a})$$

$$\begin{aligned} a_j(t) &= a_j - i \sum_{k=1,2} \int_0^t dt' \langle j | H_W | k \rangle a_k(t') \\ &\quad - \sum_{n,k} \int_0^t dt' \int_0^{t'} dt'' e^{i(E_0 - E_n)(t' - t'')} \langle j | H_W | n \rangle \langle n | H_W | k \rangle a_k(t'') \end{aligned} \quad (\text{H.10b})$$

where we abbreviated $|A_k\rangle, |B_n\rangle$ simply as $|k\rangle, |n\rangle$. Unlike Eqs. (H.8), Eqs. (H.10) are no longer mutually coupled. They can be solved by means of a Laplace transform. Define a new variable

$$\tilde{a}_j(s) = \int_0^\infty dt e^{-st} a_j(t) \quad (\text{H.11})$$

insert it in Eq. (H.10b), and we obtain

$$\tilde{a}_j(s) = \frac{a_j}{s} - \frac{i}{s} \sum_k W_{jk}(s) \tilde{a}_k \quad (\text{H.12a})$$

$$W_{jk}(s) = \langle j | H_W | k \rangle + \sum_n \frac{\langle j | H_W | n \rangle \langle n | H_W | k \rangle}{E_0 - E_n + is} \quad (\text{H.12b})$$

Using matrix notation

$$\begin{aligned} \mathbf{a}(0) &= \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad \mathbf{a}(t) = \begin{bmatrix} a_1(t) \\ a_2(t) \end{bmatrix}, \quad \tilde{\mathbf{a}}(s) = \begin{bmatrix} \tilde{a}_1(s) \\ \tilde{a}_2(s) \end{bmatrix} \\ \mathbf{W}(s) &= [W_{jk}(s)] \end{aligned} \quad (\text{H.13})$$

Equations (H.12) can be cast in a form

$$[s + i \mathbf{W}(s)] \tilde{\mathbf{a}}(s) = \mathbf{a}(0) \quad (\text{H.14a})$$

$$\therefore \tilde{\mathbf{a}}(s) = [s + i \mathbf{W}(s)]^{-1} \mathbf{a}(0) \quad (\text{H.14b})$$

Then the inverse Laplace transform may be applied to yield

$$\begin{aligned} \mathbf{a}(t) &= \frac{1}{2\pi i} \int_{s_0-i\infty}^{s_0+i\infty} ds e^{st} \tilde{\mathbf{a}}(s) = \frac{1}{2\pi i} \int_{s_0-i\infty}^{s_0+i\infty} ds \frac{e^{st}}{[s + i \mathbf{W}(s)]} \mathbf{a}(0) \\ &= \frac{1}{2\pi i} \int_{-\infty}^{+\infty} d\gamma \frac{e^{(i\gamma + s_0)t}}{[\gamma - is_0 + \mathbf{W}(i\gamma + s_0)]} \mathbf{a}(0) \end{aligned} \quad (\text{H.15})$$

The value of s_0 has to be chosen in such a way that the integration path in the complex s -plane lies to the right of all singularities of $[s + i \mathbf{W}(s)]^{-1}$. From Eq. (H.12b) we see that $\mathbf{W}$ is analytic except at those points where

$$E_0 - E_n + is = 0 \quad (\text{H.16})$$

which are on the imaginary axis. Therefore we can set $s_0 > 0$ and infinitesimally small and rewrite it as ϵ . Then the question is from where does the main contribution to the integral come from?

For $H_W = 0$ we have $\mathbf{W}(s) = 0$ and the singularity of the integrand is a pole at $s = 0$. In other words, the main contribution comes from the vicinity of $s = 0$. In this case we obtain

$$\mathbf{a}(t) = \mathbf{a} \quad \text{for } t \geq 0 \quad (\text{H.17})$$

as is necessary, since the states $|A_k\rangle$ are stable when $H_W = 0$. When it is switched on, this pole moves but for small perturbation, it will stay close to the imaginary axis. Moreover, $\mathbf{W}(s)$ is expected to vary appreciably only when s changes by an amount of $\sim E_0 - E_n$, i.e. over a range of energy determined by the strong interaction H_0 . The second approximation (2) of the Weisskopf–Wigner approach is to

replace $\mathbf{W}(s)$ in the vicinity of $s = 0$, $\text{Res} > 0$, by a constant

$$\begin{aligned}\mathbf{W}(s) &\rightarrow \mathbf{W} \equiv \lim_{\epsilon \rightarrow 0} \mathbf{W}(\epsilon) \\ \mathbf{W}_{jk} &= \langle j | H_{\mathbf{W}} | k \rangle + \sum_n \frac{\langle j | H_{\mathbf{W}} | n \rangle \langle n | H_{\mathbf{W}} | k \rangle}{E_0 - E_n + i\epsilon}\end{aligned}\quad (\text{H.18})$$

This is an approximation valid for time t very long compared with strong interaction scale. Then using the residue theorem to evaluate Eq. (H.15), we obtain

$$\mathbf{a}(t) = \frac{1}{2\pi i} \oint ds \frac{e^{st}}{s + i\mathbf{W}} \mathbf{a}(0) = e^{-i\mathbf{W}t} \mathbf{a}(0) \quad t > 0 \quad (\text{H.19})$$

Returning to the Schrödinger picture and denoting $\boldsymbol{\Phi} = e^{-iH_0 t} \mathbf{a}$, it satisfies the Schrödinger equation with the Hamiltonian replaced with the mass matrix

$$i \frac{d}{dt} \boldsymbol{\Phi}(t) = \mathbf{A} \boldsymbol{\Phi}(t) \quad (\text{H.20})$$

where the mass matrix $\mathbf{A}$ is defined by

$$\mathbf{A} = H_0 + \mathbf{W} \equiv \mathbf{M} - \frac{i}{2} \boldsymbol{\Gamma} \quad (\text{H.21a})$$

$$\mathbf{M} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^\dagger), \quad \boldsymbol{\Gamma} = i(\mathbf{A} - \mathbf{A}^\dagger) \quad (\text{H.21b})$$

Both $\mathbf{M}$ and $\boldsymbol{\Gamma}$ are hermitian but $\mathbf{A}$ is not. The matrix elements are expressed as

$$M_{jk} = E_0 \delta_{jk} + \langle j | H_{\mathbf{W}} | k \rangle + \text{P} \sum_n \frac{\langle j | H_{\mathbf{W}} | n \rangle \langle n | H_{\mathbf{W}} | k \rangle}{E_0 - E_n} \quad (\text{H.22a})$$

$$\Gamma_{jk} = 2\pi \sum_n \delta(E_0 - E_n) \langle j | H_{\mathbf{W}} | n \rangle \langle n | H_{\mathbf{W}} | k \rangle \quad (\text{H.22b})$$

where P means principal value. Γ_{jk} are determined solely by matrix elements of $H_{\mathbf{W}}$ for energy conserving process, i.e. they are on-shell amplitudes.

Comparing Eqs. (H.22) with Eq. (6.38), we see that the mass matrix is essentially an unperturbed strong Hamiltonian plus the weak transition operator i.e. $\mathbf{W} = \hat{T}_{\mathbf{W}}$. To the first order in weak interaction, it reduces to $H_0 + H_{\mathbf{W}}$. If $H_{\mathbf{W}}$ is a $\Delta S = 1$ operator, it does not induce mixing between K^0 and $\bar{K}^0$. It is a second order effect as expressed in Eqs. (H.22). The difference between $\hat{T}_{\mathbf{W}}$ here and the general transition operator $\hat{T}$ in Eq. (6.38) is that the former is valid only for time long compared to strong interaction scale and that only continuum states to which the discrete states can decay appear in the intermediate states.

Equation (H.20) looks like a Schrödinger equation but $\mathbf{A}$ is not hermitian. Within a sub-space consisting of K^0 and $\bar{K}^0$ which are degenerate eigenstates of H_0 , the transition operator $\hat{T}_{\mathbf{W}} = \mathbf{W}$ gives modifications due to the weak interaction process. The antihermitian part of $\hat{T}_{\mathbf{W}}$ describes the exponential decay of the states. If all the Γ_{jk} 's were zero, the system would evolve without decay. But the off-diagonal

elements of $\hat{T}_W$ still cause mixing of K^0 and $\bar{K}^0$ and the eigenstates of the mass matrix are mixture of both. Denoting the eigenstates and eigenvalues as $|\sigma\rangle$ and λ_σ , the time evolution of the states is given by

$$|\sigma(t)\rangle = |\sigma(0)\rangle e^{-i\lambda_\sigma t}, \quad \lambda_\sigma = m_\sigma - i\frac{\Gamma_\sigma}{2} \quad (\text{H.23})$$

and represents exponentially decaying states with mean life $\tau = 1/\Gamma_\sigma$.

Note that the exponential decay law is only an approximation. If $\mathbf{W}$ in Eq. (H.15) is expanded in Taylor series, inclusion of higher powers of γ results in corrections to the exponential decay law. It can be shown that the approximation fails both for very short and for very long times [146, 228, 229].

Appendix I

Answers to the Problems

Chapter 2: Particles and Fields

Problem 2.2: For a static solution $\partial_0 = 0$. Applying $-\nabla^2 + \mu^2$ to $\varphi(\mathbf{r}) = \int \frac{d^3 p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{r}} \tilde{\varphi}(\mathbf{p})$ and using $\delta^3(\mathbf{r}) = \int d^3 p e^{i\mathbf{p}\cdot\mathbf{r}} / (2\pi)^3$, we have $\tilde{\varphi}(\mathbf{p}) = g / (p^2 + \mu^2)$. Then

$$\begin{aligned} \varphi(\mathbf{r}) &= g \int \frac{d^3 p}{(2\pi)^3} \frac{e^{i\mathbf{p}\cdot\mathbf{r}}}{p^2 + \mu^2} = \frac{1}{(2\pi)^2} \int_0^\infty p^2 dp \int_{-1}^1 dz \frac{e^{iprz}}{p^2 + \mu^2} \\ &= \frac{g}{4\pi r} \frac{1}{2\pi i} \int_{-\infty}^\infty e^{ipr} \left[\frac{1}{p - i\mu} + \frac{1}{p + i\mu} \right] = g \frac{e^{-\mu r}}{4\pi r} \end{aligned}$$

In deriving the last equation, we have closed the contour through the upper half-circle of the complex p plane.

Problem 2.3: $r_W = \hbar c / m_W \simeq 200 \text{ MeV} \cdot 10^{-13} \text{ cm} / 81 \text{ GeV} \simeq 2.5 \times 10^{-16} \text{ cm}$.

Problem 2.4: Assuming the size of the Crab Nebula is $L = 400 \text{ ly} = 400 \times 0.946 \times 10^{16} \text{ m}$, $m_\gamma < \hbar c / 100 \text{ ly} \simeq 2 \times 10^{-7} \text{ eV m} / 4 \times 10^{18} \text{ m} \simeq 0.5 \times 10^{-25} \text{ eV}$. Assuming an intergalactic distance of $1 \text{ Mpc} \sim 3.26 \times 10^6 \text{ ly}$, $m_\gamma < 0.4 \times 10^{-29} \text{ eV}$.

Problem 2.8: $\lambda = \frac{4\pi E}{\Delta m^2} = 4\pi \frac{E}{\hbar c} \left(\frac{\hbar c}{|\Delta m| c^2} \right)^2 = 2.4 \frac{E(\text{MeV})}{\delta m^2(\text{eV})^2} (\text{m})$

Chapter 3: Lorentz Invariance

Problem 3.1: Applying Lorentz boost to the longitudinal component

$$\begin{aligned} x^{0'} &= \gamma x^0 + \beta \gamma (\hat{\boldsymbol{\beta}} \cdot \mathbf{x}), \quad x_{\parallel}' = \beta \gamma x^0 + \gamma (\hat{\boldsymbol{\beta}} \cdot \mathbf{x}) \\ \therefore \mathbf{x}' &= \mathbf{x}_{\parallel}' + \mathbf{x}_{\perp} = \hat{\boldsymbol{\beta}} [\beta \gamma x^0 + \gamma (\hat{\boldsymbol{\beta}} \cdot \mathbf{x})] + [\mathbf{x} - \hat{\boldsymbol{\beta}} (\hat{\boldsymbol{\beta}} \cdot \mathbf{x})] \\ &= \mathbf{x} + \hat{\boldsymbol{\beta}} \gamma x^0 + (\gamma - 1) (\hat{\boldsymbol{\beta}} \cdot \mathbf{x}) \hat{\boldsymbol{\beta}} \end{aligned}$$

Problem 3.2: (3.28e): A relativistic wave equation for a scalar field is given by the Klein–Gordon equation. It has a plane wave solution $\sim e^{ik_\mu x^\mu}$, which should be a Lorentz scalar. Then $k_\mu x^\mu$ is also a scalar, which in turn means k_μ is a Lorentz vector. Another way of saying that $k_\mu x^\mu$ is a Lorentz scalar is it is a phase, i.e. a pure number which does not depend on the coordinate frame.

(3.28g): Since $(\rho c; \mathbf{j})$ is a vector and $\partial_\mu \partial^\mu$ is a Lorentz scalar, the Maxwell equation $\partial_\mu F^{\mu\nu} = j^\nu$ tells that $(A^\mu = \phi/c; \mathbf{A})$ must also be a vector.

(3.28h): $\partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} \stackrel{(3.24b)}{=} M^\nu_\mu \partial_\nu$
which shows that ∂_μ is a covariant vector.

Problem 3.6:

$$\begin{aligned}\eta'' &= \frac{1}{2} \ln \frac{1 + \beta''}{1 - \beta''} = \frac{1}{2} \ln \frac{1 + (\beta + \beta')/(1 + \beta\beta')}{1 - (\beta + \beta')/(1 + \beta\beta')} \\ &= \frac{1}{2} \ln \frac{(1 + \beta)(1 + \beta')}{(1 - \beta)(1 - \beta')} = \eta + \eta'\end{aligned}$$

Problem 3.7(c):

$$\begin{aligned}g_{\mu\nu} &= g_{\rho\sigma} L^\rho_\mu L^\sigma_\nu = g_{\rho\sigma} (\delta^\rho_\mu + \epsilon^\rho_\mu) (\delta^\sigma_\nu + \epsilon^\sigma_\nu) \\ &\simeq g_{\rho\sigma} \delta^\rho_\mu \delta^\sigma_\nu + g_{\rho\sigma} \delta^\rho_\mu \epsilon^\sigma_\nu + g_{\rho\sigma} \epsilon^\rho_\mu \delta^\sigma_\nu \\ &= g_{\mu\nu} + \epsilon_{\mu\nu} + \epsilon_{\nu\mu}, \quad \therefore \quad \epsilon_{\mu\nu} = -\epsilon_{\nu\mu}\end{aligned}$$

Chapter 4: Dirac Equation

Problem 4.5: $\bar{\psi} \gamma^\mu \psi = \psi^\dagger (1, \alpha_x, \alpha_y, \alpha_z) \psi \cdot e^{\alpha_x \eta_x/2} = \cosh(\eta_x/2) + \alpha_x \sinh(\eta_x/2)$.
The transformation property of the 0-th component is

$$\begin{aligned}\bar{\psi}' \gamma^0 \psi' &= \psi^\dagger [\cosh(\eta_x/2) + \alpha_x \sinh(\eta_x/2)] [\cosh(\eta_x/2) + \alpha_x \sinh(\eta_x/2)] \psi \\ &= \psi^\dagger [\{\cosh^2(\eta_x/2) + \sinh^2(\eta_x/2)\} + 2 \sinh(\eta_x/2) \cosh(\eta_x/2) \alpha_x] \psi \\ &= \psi^\dagger [\cosh \eta_x + \sinh \eta_x \alpha_x] \psi = \gamma \bar{\psi} \gamma^0 \psi + \beta \gamma \bar{\psi} \gamma^1 \psi \\ \bar{\psi}' \gamma^1 \psi' &= \psi^\dagger [\cosh(\eta_x/2) + \alpha_x \sinh(\eta_x/2)] \alpha_x [\cosh(\eta_x/2) + \alpha_x \sinh(\eta_x/2)] \psi \\ &= \psi^\dagger [\{\cosh^2(\eta_x/2) + \sinh^2(\eta_x/2)\} \alpha_x + 2 \sinh(\eta_x/2) \cosh(\eta_x/2) \psi] \psi \\ &= \psi^\dagger [\cosh \eta_x \alpha_x + \sinh \eta_x] \psi = \beta \gamma \bar{\psi} \gamma^0 \psi + \gamma \bar{\psi} \gamma^1 \psi \\ \bar{\psi}' \gamma^2 \psi' &= \bar{\psi} \gamma^2 \psi, \quad \bar{\psi}' \gamma^3 \psi' = \bar{\psi} \gamma^3 \psi\end{aligned}$$

which reproduces Eq. (3.62).

Problem 4.6:

$$\begin{aligned}\bar{u}_r(p) u_s(p) &= (\xi_r^\dagger \sqrt{p \cdot \bar{\sigma}}, \xi_r^\dagger \sqrt{p \cdot \sigma}) \begin{bmatrix} \sqrt{p \cdot \bar{\sigma}} \xi_s \\ \sqrt{p \cdot \sigma} \xi_s \end{bmatrix} \\ &= \xi_r^\dagger (\sqrt{p \cdot \bar{\sigma}} p \cdot \sigma + \sqrt{p \cdot \sigma} p \cdot \bar{\sigma}) \xi_s \stackrel{(4.60)}{=} 2m \xi_r^\dagger \xi_s = 2m \delta_{rs} \\ u_r^\dagger(p) u_s(p) &= (\xi_r^\dagger \sqrt{p \cdot \sigma}, \xi_r^\dagger \sqrt{p \cdot \bar{\sigma}}) \begin{bmatrix} \sqrt{p \cdot \bar{\sigma}} \xi_s \\ \sqrt{p \cdot \sigma} \xi_s \end{bmatrix} \\ &= \xi_r^\dagger (p \cdot \sigma + p \cdot \bar{\sigma}) \xi_s \stackrel{(4.58)}{=} 2E \delta_{rs}\end{aligned}$$

Others can be proved similarly.

Problem 4.8:

$$\begin{aligned}
\sum_{r=1,2} u_r(\mathbf{p}) \bar{u}_r(\mathbf{p}) &= N \begin{bmatrix} \sqrt{p \cdot \bar{\sigma}} \xi_r \\ \sqrt{p \cdot \bar{\sigma}} \xi_r \end{bmatrix} N \left(\xi_r^\dagger \sqrt{p \cdot \bar{\sigma}}, \xi_r^\dagger \sqrt{p \cdot \bar{\sigma}} \right) \\
&= N^2 \begin{bmatrix} \sqrt{p \cdot \bar{\sigma}} & \mathbf{0} \\ \mathbf{0} & \sqrt{p \cdot \bar{\sigma}} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{p \cdot \bar{\sigma}} & \mathbf{0} \\ \mathbf{0} & \sqrt{p \cdot \bar{\sigma}} \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{p \cdot \sigma} p \cdot \bar{\sigma} & p \cdot \sigma \\ p \cdot \bar{\sigma} & \sqrt{p \cdot \bar{\sigma}} p \cdot \sigma \end{bmatrix} \\
&= \begin{bmatrix} m & p \cdot \sigma \\ p \cdot \bar{\sigma} & m \end{bmatrix} = \not{p} + m
\end{aligned}$$

Alternative proof: Using $\mathcal{A}_+ = \sum_r u_r(\mathbf{p}) \bar{u}_r(\mathbf{p})$ and Eqs. (4.64), one can directly prove Eqs. (4.67).

Problem 4.10:

(a) $\bar{u}(p') \gamma^1 u(p)$: We show for the case of γ^1 . In the nonrelativistic limit $E = E', |\mathbf{p}| = |\mathbf{p}'| = p$, $\mathbf{p}' = p \mathbf{n}'$, $\mathbf{p} = p \mathbf{n}$.

$$\begin{aligned}
\bar{u}_r(\mathbf{p}') \gamma^1 u_s(\mathbf{p}) &= N \left(\xi_r^\dagger \sqrt{p' \cdot \bar{\sigma}}, \xi_r^\dagger \sqrt{p' \cdot \sigma} \right) \begin{bmatrix} \mathbf{0} & \sigma_1 \\ -\sigma & \mathbf{0} \end{bmatrix} \begin{bmatrix} \sqrt{p \cdot \bar{\sigma}} \xi_s \\ \sqrt{p \cdot \bar{\sigma}} \xi_s \end{bmatrix} \\
&= N^2 \xi_r^\dagger [-\sqrt{p' \cdot \sigma} \sigma_1 \sqrt{p \cdot \bar{\sigma}} + \sqrt{p' \cdot \bar{\sigma}} \sigma_1 \sqrt{p \cdot \sigma}] \xi_s \\
&= N^2 \xi_r^\dagger \left[\left(\sqrt{\frac{E+m}{2}} + \sqrt{\frac{E-m}{2}} \sigma \cdot \mathbf{n}' \right) \sigma_1 \right. \\
&\quad \times \left. \left(\sqrt{\frac{E+m}{2}} + \sqrt{\frac{E-m}{2}} \sigma \cdot \mathbf{n} \right) - (\dots - \dots) \right] \xi_s \\
&= N^2 \xi_r^\dagger [p(\sigma \cdot \mathbf{n}' \sigma_1 + \sigma_1 \sigma \cdot \mathbf{n})] \xi_s \\
&= \xi_r^\dagger [(p' + p) + i \sigma \times (p' - p)]_1 \xi_s
\end{aligned}$$

where we have used $\sqrt{p \cdot \sigma} = \sqrt{m} \exp \left[\frac{1}{2} \eta \sigma \cdot \mathbf{n} \right] = \sqrt{\frac{E+m}{2}} - \sqrt{\frac{E-m}{2}} \sigma \cdot \mathbf{n}$

(b) $\bar{u}(p') u(p)$: We keep terms only to lowest order in the 3-momenta. Thus up to $O(p^2, p'^2)$

$$\begin{aligned}
p &= (m, \mathbf{p}), \quad p' = (m, \mathbf{p}'), \quad p \cdot \sigma = m - \sigma \cdot \mathbf{p}, \\
p \cdot \bar{\sigma} &= m + \sigma \cdot \mathbf{p}
\end{aligned} \tag{I.1}$$

$$\begin{aligned}
\bar{u}(p') u(p) &= \xi^\dagger(\sqrt{p' \cdot \bar{\sigma}}, \sqrt{p' \cdot \sigma}) \begin{bmatrix} \sqrt{p \cdot \bar{\sigma}} \xi \\ \sqrt{p \cdot \bar{\sigma}} \xi \end{bmatrix} \\
&= \xi^\dagger [\sqrt{(p' \cdot \bar{\sigma})(p \cdot \sigma)} + \sqrt{(p' \cdot \sigma)(p \cdot \bar{\sigma})}] \xi \\
&= \xi^\dagger [\sqrt{(m + \sigma \cdot \mathbf{p}')(m - \sigma \cdot \mathbf{p})} + \sqrt{(m - \sigma \cdot \mathbf{p}')(m + \sigma \cdot \mathbf{p})}] \xi \\
&= m \xi^\dagger \left[1 - \frac{\sigma \cdot (\mathbf{p} - \mathbf{p}')}{2m} + 1 + \frac{\sigma \cdot (\mathbf{p} - \mathbf{p}')}{2m} + O(p^2) \right] \xi \simeq 2m \xi_s^\dagger \xi_r
\end{aligned}$$

(c) $\langle p' | \gamma^5 | p \rangle$: First use the fact that

$$\begin{aligned}
 u(p) &= \frac{\not{p}}{m} u(p), \quad \bar{u}(p') = \bar{u}(p') \frac{\not{p}'}{m} \\
 \therefore \quad \bar{u}(p') \gamma^5 u(p) &= \frac{1}{2m} [\bar{u}(p') \not{p}' \gamma^5 u(p) + \bar{u}(p') \gamma^5 \not{p} u(p)] \\
 &= \frac{1}{2m} \bar{u}(p') \gamma^5 (\not{p} - \not{p}') u(p) \\
 \gamma^0 \gamma^5 (\not{p} - \not{p}') &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & (p \cdot \sigma - p' \cdot \sigma) \\ (p \cdot \bar{\sigma} - p' \cdot \bar{\sigma}) & 0 \end{bmatrix} \\
 &= \begin{bmatrix} (p \cdot \bar{\sigma} - p' \cdot \bar{\sigma}) & 0 \\ 0 & -(p \cdot \sigma - p' \cdot \sigma) \end{bmatrix} \\
 \bar{u}(p') \gamma^5 (\not{p} - \not{p}') u(p) &= u(p')^\dagger \gamma^0 \gamma^5 (\not{p} - \not{p}') u(p) \\
 &= \xi^\dagger (\sqrt{p' \cdot \sigma}, \sqrt{p' \cdot \bar{\sigma}}) \begin{bmatrix} (p \cdot \bar{\sigma} - p' \cdot \bar{\sigma}) & 0 \\ 0 & -(p \cdot \sigma - p' \cdot \sigma) \end{bmatrix} \begin{bmatrix} \sqrt{p \cdot \bar{\sigma}} \xi \\ \sqrt{p \cdot \sigma} \xi \end{bmatrix} \\
 &= \xi^\dagger [\sqrt{p' \cdot \sigma} (p \cdot \bar{\sigma} - p' \cdot \bar{\sigma}) \sqrt{p \cdot \bar{\sigma}} - \sqrt{p' \cdot \bar{\sigma}} (p \cdot \sigma - p' \cdot \sigma) \sqrt{p \cdot \sigma}] \xi
 \end{aligned}$$

Using $(p \cdot \bar{\sigma})(p \cdot \sigma) = (p \cdot \sigma)(p \cdot \bar{\sigma}) = m^2$, we can organize the above equation to make

$$= 2m \xi^\dagger [\sqrt{(p' \cdot \sigma)(p \cdot \bar{\sigma})} - \sqrt{(p' \cdot \bar{\sigma})(p \cdot \sigma)}] \xi$$

up to here the equations are exact. Using the approximations Eq. (I.1)

$$\begin{aligned}
 \bar{u}(p') \gamma^5 u(p) &= \frac{1}{2m} \bar{u}(p') \gamma^5 (\not{p} - \not{p}') u(p) \\
 &= m \xi^\dagger \left[\left(1 + \frac{\boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{p}')}{2m} \right) - \left(1 - \frac{\boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{p}')}{2m} \right) + O(p^2) \right] \xi \\
 &\simeq \xi_s^\dagger \boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{p}') \xi_r
 \end{aligned}$$

Problem 4.12: Using Eq. (4.120)

$$\psi_L = (N_1 + iN_2)_L = \xi - \sigma_2 \eta^*, \quad \psi_R = (N_1 + iN_2)_R = i(\eta - \sigma_2 \xi^*)$$

Then applying Eq. (4.122) and Eq. (4.123), we obtain

$$\begin{aligned}
 i(\partial_0 - \boldsymbol{\sigma} \cdot \nabla) \psi_L &= i(\partial_0 - \boldsymbol{\sigma} \cdot \nabla) \xi - i(\partial_0 - \boldsymbol{\sigma} \cdot \nabla)(\sigma_2 \eta^*) \\
 &= mi(\eta - \sigma_2 \xi^*) = m\psi_R
 \end{aligned}$$

Similarly

$$i(\partial_0 + \boldsymbol{\sigma} \cdot \nabla) \psi_R = m\psi_L$$

This is the Dirac equation Eq. (4.51) expressed in two-component spinors.

Chapter 5: Field Quantization

Problem 5.1:

1. Complex field:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi^\dagger)} &= \partial^\mu \varphi, \quad \frac{\partial \mathcal{L}}{\partial \varphi^\dagger} = -m^2 \varphi \\ \therefore \quad \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi^\dagger)} - \frac{\partial \mathcal{L}}{\partial \varphi^\dagger} &= (\partial_\mu \partial^\mu + m^2) \varphi = 0\end{aligned}$$

2. Electromagnetic field:

$$\begin{aligned}\frac{\partial \left(\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right)}{\partial(\partial_\mu A_\nu)} &= F^{\mu\nu}, \quad \frac{\partial \mathcal{L}}{\partial A_\nu} = -q j^\nu, \\ \therefore \quad \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} - \frac{\partial \mathcal{L}}{\partial A_\nu} &= -\partial_\mu F^{\mu\nu} + q j^\nu = 0\end{aligned}$$

3. Dirac field:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} &= \bar{\psi} i \gamma^\mu, \\ \frac{\partial \mathcal{L}}{\partial \psi} &= -m \bar{\psi} \rightarrow \bar{\psi} (i \gamma^\mu \overleftarrow{\partial}_\mu + m) = 0\end{aligned}$$

Problem 5.3:

$$\begin{aligned}\frac{d^4 x'}{d^4 x} &= (\text{Jacobian of } x^\mu \rightarrow x'^\mu) = \det \left(\frac{\partial x'^\mu}{\partial x^\nu} \right) \\ &= \det(\delta_\nu^\mu + \partial_\nu \delta x^\mu) = 1 + \partial_\mu (\delta x^\mu) \\ \therefore \quad \int (d^4 x' - d^4 x) \mathcal{L} &= \int d^4 x \mathcal{L} \partial_\mu (\delta x^\mu) = \int d^4 x [\partial_\mu (\mathcal{L} \delta x^\mu) - \partial_\mu \mathcal{L} \delta x^\mu]\end{aligned}$$

The first term vanishes because it is a surface integral.

Problem 5.5: The second term of Eq. (5.47) can be written as

$$\begin{aligned}\partial_\mu T_\nu^\mu \delta x^\nu &= \partial_\mu T_\nu^\mu \epsilon_\rho^\nu x^\rho = \partial_\mu T^{\mu\nu} \epsilon_{\nu\rho} x^\rho = [\partial_\mu (T^{\mu\nu} x^\rho) - T^{\rho\nu}] \epsilon_{\nu\rho} \\ &= -[\partial_\mu (T^{\mu\nu} x^\rho) - T^{\rho\nu}] \epsilon_{\rho\nu} \xrightarrow{\rho \leftrightarrow \nu} -[\partial_\mu (T^{\mu\rho} x^\nu) - T^{\nu\rho}] \epsilon_{\nu\rho} \\ &= \frac{1}{2} [\partial_\mu (T^{\mu\nu} x^\rho - T^{\mu\rho} x^\nu) - T^{\rho\nu} + T^{\nu\rho}] \epsilon_{\nu\rho} = 0\end{aligned}$$

Problem 5.6: Differentiating the Hamiltonian density

$$\frac{\partial \mathcal{H}}{\partial t} = \frac{\partial \phi^\dagger}{\partial t^2} \frac{\partial^2 \phi}{\partial t} + \nabla \phi^\dagger \cdot \nabla \frac{\partial \phi}{\partial t} + m^2 \phi^\dagger \frac{\partial \phi}{\partial t} + \{\phi \leftrightarrow \phi^\dagger\}$$

Changing the first term by use of the Euler equation

$$\begin{aligned} \frac{\partial \phi^\dagger}{\partial t} (\nabla^2 - m^2) \phi + \nabla \phi^\dagger \cdot \nabla \dot{\phi} + m^2 \phi^\dagger \dot{\phi} + \{\phi \leftrightarrow \phi^\dagger\} \\ = \nabla \cdot \{\phi^\dagger \nabla \dot{\phi} + \dot{\phi}^\dagger \nabla \phi\} = -\nabla \cdot \mathcal{P} \end{aligned}$$

It is a surface integral and vanishes at infinity. It also proves problem (2). Next differentiating $\mathcal{P}$

$$\begin{aligned} -\frac{\partial \mathcal{P}}{\partial t} &= \frac{\partial^2 \phi^\dagger}{\partial t^2} \nabla \phi + \nabla \phi^\dagger \frac{\partial^2 \phi}{\partial t^2} + \frac{\partial \phi^\dagger}{\partial t} \nabla \frac{\partial \phi}{\partial t} + \nabla \frac{\partial \phi^\dagger}{\partial t} \frac{\partial \phi}{\partial t} \\ &= (\nabla^2 - m^2) \phi^\dagger \nabla \phi + \nabla \phi^\dagger (\nabla^2 - m^2) \phi + \nabla \left(\frac{\partial \phi^\dagger}{\partial t} \frac{\partial \phi}{\partial t} \right) \end{aligned}$$

Partially integrating the ∇^2 part

$$\begin{aligned} &= [\partial_i \phi^\dagger \nabla \phi + \{\phi^\dagger \leftrightarrow \phi\}]_{x_i=-\infty}^{x_i=\infty} - \nabla (\nabla \phi^\dagger \cdot \nabla \phi - m^2 \phi^\dagger \phi) \\ &\quad + \nabla \left(\frac{\partial \phi^\dagger}{\partial t} \frac{\partial \phi}{\partial t} \right) \end{aligned}$$

They are all surface integrals and vanish.

Problem 5.9: If the operator $e^{ia\hat{p}}$ satisfies

$$O(\hat{q} + a) = e^{ia\cdot\hat{p}} O(\hat{q}) e^{-ia\cdot\hat{p}} \quad (\text{I.2})$$

then, for infinitesimal displacement ε ,

$$O(\hat{q}) + \varepsilon \frac{d}{d\hat{q}} O(\hat{q}) + \dots = O(\hat{q}) + i\varepsilon[\hat{p}, O(\hat{q})] + \dots$$

By putting $O = \hat{q}$, we obtain $[\hat{q}, \hat{p}] = i$. Namely, space translation operator is the momentum operator. The above equation means $\frac{d}{d\hat{q}} O(\hat{q}) = i[\hat{p}, O]$. In the Lorentz invariant formalism, translation operator is defined by

$$O(x + a) = e^{iP_\mu x^\mu} O(x) e^{iP_\mu x^\mu}$$

Namely the space part is defined with opposite sign compared to equation (I.2). Therefore to move $x^i \rightarrow x^i + a^i$ we have to use $-P^i = P_i$. Hence

$$\nabla_i O_H = -i[P^i, O_H] = i[P_i, O_H]$$

Problem 5.12: Inserting

$$\begin{aligned}
\phi(x) &= \int \frac{d^3 k}{(2\pi)^3 \sqrt{2\omega}} \left[c_k e^{i k x - i \omega t_x} + d_k^\dagger e^{-i k x + i \omega t_x} \right] \\
\pi(y) &= \partial_t \phi^\dagger(y) = \int \frac{i d^3 k}{(2\pi)^3} \sqrt{\frac{\omega}{2}} \left[c_k^\dagger e^{-i k y + i \omega t_y} - d_k e^{+i k y - i \omega t_y} \right] \\
[\phi(x, t), \pi(y, t)] &= \iint \frac{i d^3 k d^3 p}{(2\pi)^6} \frac{1}{2} \left\{ [c_k, c_p^\dagger] e^{i k x - i p y - i(\omega_k - \omega_p)t} \right. \\
&\quad - [d_k^\dagger, d_p] e^{-i k x - i p y + i(\omega_k - \omega_p)t} - [c_k, d_p] e^{i k x - i p y - i(\omega_k - \omega_p)t} \\
&\quad \left. + [d_k^\dagger, d_p^\dagger] e^{-i k x - i p y + i(\omega_k - \omega_p)t} \right\}
\end{aligned}$$

The third and fourth term vanish and the first and second term give $\mathbf{k} = \mathbf{p}$ by integration

$$= \int \frac{i d^3 k}{2(2\pi)^3} \{ e^{i \mathbf{k} \cdot (\mathbf{x} - \mathbf{y})} + e^{-i \mathbf{k} \cdot (\mathbf{x} - \mathbf{y})} \} = i \delta^3(\mathbf{x} - \mathbf{y})$$

Problem 5.14: Prove by induction. We ignore the normalization factor since that do not change before and after the calculations. Using

$$a_k a_p^\dagger = \delta_{pk} + a_p^\dagger a_k, \quad a_k^\dagger a_p^\dagger = a_p^\dagger a_k^\dagger, \quad a_k |0\rangle = 0,$$

we can prove for the one particle state

$$\begin{aligned}
H|p\rangle &= N \sum \omega_k a_k^\dagger a_k a_p^\dagger |0\rangle \\
&= N \sum \omega_k a_k^\dagger (\delta_{pk} + a_p^\dagger a_k) |0\rangle = \omega_p |p\rangle
\end{aligned}$$

Next operating the Hamiltonian on $|n\rangle \equiv |p_1, p_2, \dots\rangle = a_{p_1}^\dagger |p_2, \dots\rangle = a_{p_1}^\dagger |n-1\rangle$ and assume $H|n-1\rangle = (\omega_2 + \dots + \omega_n)|n-1\rangle$, then

$$\begin{aligned}
H|n\rangle &= \sum \omega_k a_k^\dagger a_k a_{p_1}^\dagger |n-1\rangle = \sum \omega_k a_k^\dagger (\delta_{p_1 k} + a_{p_1}^\dagger a_k) |n-1\rangle \\
&= \omega_1 a_{p_1}^\dagger |n-1\rangle + a_{p_1}^\dagger H|n-1\rangle \\
&= \omega_1 |n\rangle + (\omega_2 + \dots) a_{p_1}^\dagger |n-1\rangle \\
&= (\omega_1 + \omega_2 + \dots) |n\rangle
\end{aligned}$$

Chapter 6: Scattering Matrix**Problem 6.1:** Using

$$\begin{aligned}
\langle f | H_I^\varepsilon(t_1) H_I^\varepsilon(t_2) | i \rangle &= \sum \langle f | H_{IS} | n \rangle e^{-i(E_n - E_f)t_1 - \varepsilon | t_1 |} \\
&\quad \times \langle n | H_{IS} | i \rangle e^{-i(E_i - E_n)t_2 - \varepsilon | t_2 |}
\end{aligned}$$

The expectation value of the second order S-matrix becomes

$$\begin{aligned}
 \langle f|S|i\rangle &= (-i)^2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \langle f|H_{IS}^{\varepsilon}(t_1)H_{IS}^{\varepsilon}(t_2)|i\rangle \\
 &= -\sum \langle f|H_{IS}|n\rangle \langle n|H_{IS}|i\rangle \\
 &\quad \times \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 e^{-i(E_n-E_f)t_1-\varepsilon|t_1|-i(E_i-E_n)t_2-\varepsilon|t_2|} \\
 &= -\sum \langle f|H_{IS}|n\rangle \langle n|H_{IS}|i\rangle \\
 &\quad \times \int_{-\infty}^{\infty} dt_1 e^{i(E_f-E_i)t_1-\varepsilon|t_1|} \int_0^{\infty} d\tau e^{i(E_i-E_n)\tau-\varepsilon\tau} \\
 &= 2\pi i \delta(E_f-E_i) \sum \frac{\langle f|H_{IS}|n\rangle \langle n|H_{IS}|i\rangle}{E_i-E_n+i\varepsilon}
 \end{aligned}$$

where we changed the integration variable $t_2 \rightarrow \tau = t_1 - t_2$ in going from the second to the third line.

Problem 6.2: From Eq. (6.51), $dE_{\text{LAB}} = \beta\gamma d\cos\theta^*$, $d\Omega^* = d\cos\theta^* d\phi^*$. Therefore $\frac{dN}{d\Omega^*} \propto \frac{dN}{d\cos\theta^*} \propto \frac{dN}{dE_{\text{LAB}}}$

Problem 6.3:

$$\begin{aligned}
 s &= (p_1 + p_2)^2 = p_1^2 + p_2^2 + 2(p_1 \cdot p_2) \\
 t &= (p_1 - p_3)^2 = p_1^2 + p_3^2 - 2(p_1 \cdot p_3) \\
 u &= (p_1 - p_4)^2 = p_1^2 + p_4^2 - 2(p_1 \cdot p_4)
 \end{aligned}$$

Using $p_i^2 = m_i^2$, $p_2 - p_3 - p_4 = -p_1$, we obtain $s + t + u = \sum m_i^2$.

Problem 6.4: Using the flux factor $F = 4E_1 E_2 v_{12} = 4E_1 M(p_1/E_1) = 4p_1 M$ and the phase space volume $dLIPS$, the cross section is given by

$$\begin{aligned}
 d\sigma &= |\mathcal{M}_{fi}|^2 dLIPS_{\text{LAB}}/F \\
 &= (2\pi)^4 \delta(E_1 + M - E_3 - E_4) \delta^3(\mathbf{p}_1 - \mathbf{p}_3 - \mathbf{p}_4) \frac{|\mathcal{M}_{fi}|^2}{F} \frac{d^3 p_3 d^3 p_4}{(2\pi)^6 4E_3 E_4}
 \end{aligned}$$

$\int d^3 p_4$ gives 1, at the same time changing the momentum $\mathbf{p}_4$ in E_4 . To carry out $\int d^3 p_3$, we put the content of the δ function as G:

$$\begin{aligned}
 G &\equiv E_3 + E_4 = p_3 + \sqrt{(\mathbf{p}_1 - \mathbf{p}_3)^2 + M^2} \\
 &= p_3 + \sqrt{p_1^2 + p_3^2 - 2p_1 p_3 \cos\theta + M^2} \\
 \therefore \frac{\partial G}{\partial p_3} &= 1 + \frac{p_3 - p_1 \cos\theta}{E_4} = \frac{1}{E_4} (E_4 + p_3 - p_1 \cos\theta) \\
 &= (E_1 + M - p_1 \cos\theta)/E_4 = [M + p_1(1 - \cos\theta)]/E_4
 \end{aligned}$$

Now we introduce $q = p_1 - p_3 = p_4 - p_2$ and use $p_2^2 = p_4^2 = M^2$, $p_1^2 = p_3^2 = 0$.

$$\begin{aligned} p_4^2 &= (p_4 - p_2 + p_2)^2 = (q + p_2)^2 = q^2 + 2(p_2 \cdot q) + p_2^2 \\ \therefore -q^2 &= 2(p_2 \cdot q) = 2M(E_1 - E_3) = 2M(p_1 - p_3) \\ &= -(p_1 - p_3)^2 = 2p_1 p_3 (1 - \cos \theta) \\ \therefore p_1(1 - \cos \theta) &= M(p_1 - p_3)/p_3 \end{aligned}$$

Inserting the above equalities, we obtain $\partial G/\partial p_3 = p_1 M/p_3 E_4$ and

$$\begin{aligned} \frac{d L I P S_{\text{LAB}}}{F} &= \frac{1}{4 p_1 M} \left(\frac{\partial G}{\partial p_3} \right)^{-1} \frac{p_3^2 d \Omega_{\text{LAB}}}{16 \pi^2 E_3 E_4} = \frac{p_3^2 d \Omega_{\text{LAB}}}{64 \pi^2 (M p_1)^2} \\ \therefore \left. \frac{d \sigma}{d \Omega} \right|_{\text{LAB}} &= \left(\frac{p_3}{p_1} \right)^2 \frac{|\mathcal{M}_{fi}|^2}{64 \pi^2 M^2} \end{aligned}$$

Problem 6.6:

1. Proof of Eq. (6.142a). Using $\gamma^5 \gamma^5 = 1$, $\gamma^5 \gamma^\mu = -\gamma^\mu \gamma^5$

$$\begin{aligned} \text{Tr}[\gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_{2n+1}}] &= \text{Tr}[\gamma^5 \gamma^5 \gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_{2n+1}}] = \text{Tr}[\gamma^5 \gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_{2n+1}} \gamma^5] \\ &= -\text{Tr}[\gamma^5 \gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_{2n+1}} \gamma^5] = (-1)^{2n+1} \text{Tr}[\gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_{2n+1}}] \\ &= -\text{Tr}[\gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_{2n+1}}], \quad \therefore \text{Tr}[\gamma^{\mu_1} \gamma^{\mu_2} \dots \gamma^{\mu_{2n+1}}] = 0 \end{aligned}$$

2. Proof of Eq. (6.142b) Since $\text{Tr}[A] = \text{Tr}[A^T]$, $C \gamma^{\mu T} C^{-1} = -\gamma^\mu$, we insert $C C^{-1}$ between all γ 's.

$$\begin{aligned} \text{Tr}[\gamma^1 \gamma^2 \dots \gamma^n] &= \text{Tr}[\gamma^{nT} \dots \gamma^{2T} \gamma^{1T}] = \text{Tr}[C \gamma^{nT} C^{-1} \dots C \gamma^{1T} C^{-1}] \\ &= (-1)^n \text{Tr}[\gamma^n \dots \gamma^2 \gamma^1] \end{aligned}$$

3. Proof of Eq. (6.142c). Taking the trace of both sides of $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}$, and use $\text{Tr}[\gamma^\mu \gamma^\nu] = \text{Tr}[\gamma^\nu \gamma^\mu]$
4. Proof of Eq. (6.142d). The first step is

$$\text{Tr}[\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma] = \text{Tr}[\{2g^{\mu\nu} - \gamma^\nu \gamma^\mu\} \gamma^\rho \gamma^\sigma]$$

Using Eq. (6.142c)

$$\begin{aligned} &= 8g^{\mu\nu} g^{\rho\sigma} - \text{Tr}[\gamma^\nu \gamma^\mu \gamma^\rho \gamma^\sigma] = 8g^{\mu\nu} g^{\rho\sigma} - \text{Tr}[\gamma^\nu \{2g^{\mu\rho} - \gamma^\rho \gamma^\mu\} \gamma^\sigma] \\ &= 8g^{\mu\nu} g^{\rho\sigma} - 8g^{\mu\rho} g^{\nu\sigma} + \text{Tr}[\gamma^\nu \gamma^\rho \gamma^\mu \gamma^\sigma] \\ &= 8g^{\mu\nu} g^{\rho\sigma} - 8g^{\mu\rho} g^{\nu\sigma} + \text{Tr}[\gamma^\nu \gamma^\rho \{2g^{\mu\sigma} - \gamma^\sigma \gamma^\mu\}] \\ &= 8g^{\mu\nu} g^{\rho\sigma} - 8g^{\mu\rho} g^{\nu\sigma} + 8g^{\mu\sigma} g^{\nu\rho} - \text{Tr}[\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma] \end{aligned}$$

In deriving the last equation, we used invariance of the trace under cyclic permutation.

$$\therefore \text{Tr}[\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma] = 4g^{\mu\nu} g^{\rho\sigma} - 4g^{\mu\rho} g^{\nu\sigma} + 4g^{\mu\sigma} g^{\nu\rho}$$

Problem 6.7: Using Eq. (6.148)

$$\begin{aligned}
 \bar{u}_s(p_f) \gamma^0 u_r(p_i) &= u_s(p_f)^\dagger u_r(p_i) \\
 &\sim \xi_s^\dagger \left[\left(1 - \frac{\boldsymbol{\sigma} \cdot \mathbf{p}_f}{E_f + m} \right) \left(1 - \frac{\boldsymbol{\sigma} \cdot \mathbf{p}_i}{E_i + m} \right) \right. \\
 &\quad \left. + \left(1 + \frac{\boldsymbol{\sigma} \cdot \mathbf{p}_f}{E_f + m} \right) \left(1 + \frac{\boldsymbol{\sigma} \cdot \mathbf{p}_i}{E_i + m} \right) \right] \xi_r \\
 &= 2\xi_s^\dagger \left[1 + \frac{p^2}{(E + m)^2} (\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}_f)(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}_i) \right] \xi_r \\
 &= 2\xi_s^\dagger \left[\left(1 + \frac{p^2}{(E + m)^2} (\hat{\mathbf{p}}_f \cdot \hat{\mathbf{p}}_i) \right) + \frac{p^2}{(E + m)^2} i \boldsymbol{\sigma} \cdot \hat{\mathbf{p}}_f \times \hat{\mathbf{p}}_i \right] \xi_r
 \end{aligned}$$

where we have used $\mathbf{p} = p\hat{\mathbf{p}}$, $p_f = p_i = p$, $E_f = E_i = E$. The first term is the spin non-flip and the second term is the spin flip term. Since $(\hat{\mathbf{p}}_f \cdot \hat{\mathbf{p}}_i) = \cos \theta$ and $|\hat{\mathbf{p}}_f \times \hat{\mathbf{p}}_i| = \sin \theta$,

$$\begin{aligned}
 A &= \text{spin non-flip term} = \left[(E + m) \left\{ 1 + \frac{p^2}{(E + m)^2} \cos \theta \right\} \right]^2 \\
 B &= \text{spin flip term} = \left[(E + m) \left\{ \frac{p^2}{(E + m)^2} \sin \theta \right\} \right]^2 \\
 \sum_{r,s=1,2} (A + B) &= 2 \times (E + m)^2 \left[1 + \frac{2p^2}{(E + m)^2} \cos \theta + \frac{p^4}{(E + m)^4} \right]
 \end{aligned}$$

Factor 2 arises because each term is counted twice, for non-flip term $r = s = 1, 2$ and flip term $r = 1, s = 2$ and $r = 2, s = 1$. Using $\cos \theta = 1 - 2 \sin^2(\theta/2)$

$$\begin{aligned}
 A + B &= 2(E + m)^2 \left[\left(1 + \frac{p^2}{(E + m)^2} \right)^2 - \frac{4p^2}{(E + m)^2} \sin^2 \frac{\theta}{2} \right] \\
 &= 8E^2 \left(1 - \frac{p^2}{E^2} \sin^2 \frac{\theta}{2} \right)
 \end{aligned}$$

Dividing by 2 for taking the average for the initials state and also by $4E^2$ for particle normalization recovers the Eq. (6.147).

Problem 6.8: Considering only the vacuum in the intermediate stage

$$\begin{aligned}
 &\langle f | i e \left[\phi^\dagger(x) \partial^\mu \phi(x) - \partial^\mu \phi^\dagger(x) \phi(x) \right] | i \rangle \\
 &= i e \left\{ \langle f | (\phi^\dagger(x) | 0 \rangle \langle 0 | \partial^\mu \phi(x) | i \rangle) - \langle f | \partial^\mu \phi^\dagger(x) | 0 \rangle \langle 0 | \phi(x) | i \rangle \right\} \\
 &= e \left\{ p_i^\mu e^{i(\mathbf{p}_i - \mathbf{p}_f) \cdot \mathbf{x} - i(E_i - E_f)t} + p_f^\mu e^{-i(\mathbf{p}_i - \mathbf{p}_f) \cdot \mathbf{x} + i(E_i - E_f)t} \right\} \\
 &= e \left\{ p_i^\mu e^{i(p_i - p_f)_\mu x^\mu} + p_f^\mu e^{-i(p_i - p_f)_\mu x^\mu} \right\}
 \end{aligned}$$

The first and the second term give the same contribution after integration over x . Comparing with Eq. (6.135), we see only $\bar{u}_s \gamma^\mu u_s$ is replaced with $2p^\mu$. Therefore, the only modification in the final cross section is to replace $(1 - v^2 \sin^2 \frac{\theta}{2})$ with 1.

Chapter 7: QED

Problem 7.1: Noting $q = p_3 - p_1$, $q^2 = p_3^2 + p_1^2 - 2(p_1 \cdot p_3) = 2\{m^2 - (p_1 \cdot p_3)\}$

$$\begin{aligned} \frac{1}{2} q_\mu L^{\mu\nu} &= (p_{3\mu} - p_{1\mu}) \left[p_3^\mu p_1^\nu + p_1^\mu p_3^\nu + \frac{q^2}{2} g^{\mu\nu} \right] \\ &= \left[\{p_3^2 - (p_1 \cdot p_3)\} p_1^\nu + \{(p_1 \cdot p_3) - p_1^2\} p_3^\nu + \frac{q^2}{2} q^\nu \right] \\ &= \frac{q^2}{2} p_1^\nu - \frac{q^2}{2} p_3^\nu + \frac{q^2}{2} q^\nu = 0 \end{aligned}$$

Problem 7.2: The invariant matrix squared is given by

$$|\overline{\mathcal{M}}|^2 = \frac{2e^4}{s^2} [(t - M^2)^2 + (u - M^2)^2 + 2sM^2]$$

This can be derived from Eq. (7.17) and Eq. (7.16b) by crossing $s \rightarrow u$, $t \rightarrow s$, $u \rightarrow u$ and neglecting the electron mass m^2 . In CM we have

$$s = 4E^2, \quad u = M^2 - 2E^2(1 + \beta \cos \theta), \quad t = M^2 - 2E^2(1 - \beta \cos \theta)$$

Then

$$\begin{aligned} \frac{s^2}{2e^4} |\overline{\mathcal{M}}|^2 &= [-2E^2(1 - \beta \cos \theta)]^2 + [-2E^2(1 + \beta \cos \theta)]^2 + 8M^2 E^2 \\ &= 8E^4(1 + \beta^2 \cos^2 \theta) + 8M^2 E^2 \\ &= 8E^2(E^2 + p^2 \cos^2 \theta + E^2 - P^2) \\ &= s^2[1 - (\beta^2/2) \sin^2 \theta] \\ \frac{d\sigma}{d\Omega} &= \frac{p_f}{p_i} \frac{1}{64\pi^2 s} |\overline{\mathcal{M}}|^2 = \frac{\alpha^2}{2s} \beta \left(1 - \frac{\beta^2}{2} \sin^2 \theta \right) \\ &\rightarrow \sigma_{\text{TOT}} = \frac{4\pi\alpha^2}{3s} \beta \left(\frac{3 - \beta^2}{2} \right) \end{aligned}$$

Problem 7.4:

$$k_\nu \mathcal{M}^{\mu\nu} = e^2 \bar{u}(p') \left[\gamma^\mu \frac{(\not{p}' + \not{k} + m)}{(p + k)^2 - m^2} \not{k} + \not{k} \frac{(\not{p}' - \not{k}' + m)}{(p - k')^2 - m^2} \gamma^\mu \right] u(p)$$

Using $p^2 = m^2$, $\not{k} \not{k} = k^2 = 0$, $\not{p} u(p) = m u(p)$, $\bar{u}(p') \not{p}' = m \bar{u}(p')$, $p - k' = p' - k$.

Each term in $[\dots]$ gives

$$\begin{aligned} \gamma^\mu \frac{\not{p}' + \not{k} + m}{(p + k)^2 - m^2} \not{k} u(p) &= \gamma^\mu \frac{2(p \cdot k) - \not{k} \not{p}' + m \not{k}}{2(p \cdot k)} u(p) = \gamma^\mu u(p) \\ \bar{u}(p') \not{k} \frac{\not{p}' - \not{k}' + m}{(p - k')^2 - m^2} \gamma^\mu &= \bar{u}(p') \not{k} \frac{\not{p}' - \not{k} + m}{(p' - k)^2 - m^2} \gamma^\mu \\ &= \bar{u}(p') \frac{2(p' \cdot k) - \not{p}' \not{k} + m \not{k}}{-2(p' \cdot k)} = -\bar{u}(p') \gamma^\mu \end{aligned}$$

The first and the second terms compensate each other.

Problem 7.5: We calculate interference terms only. Putting $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2$,

$$\begin{aligned} |\overline{\mathcal{M}_1 \mathcal{M}_2^*}| &= \frac{e^4}{4su} \sum_{r,s} [\bar{u}_r(p') \gamma^\nu (\not{p} + \not{k}) \gamma^\mu u_s(p)] \\ &\quad \times [\bar{u}_s(p) \gamma_\nu (\not{p}' - \not{k}') \gamma_\mu u_r(p')] \\ &= \frac{e^4}{4su} \text{Tr}[\not{p}' \gamma^\nu (\not{p} + \not{k}) \gamma^\mu \not{p} \gamma_\nu (\not{p}' - \not{k}') \gamma_\mu] \end{aligned}$$

Using $\gamma_\mu \not{A} \not{B} \not{C} \gamma^\mu = -2 \not{C} \not{B} \not{A}$, $\gamma_\mu \not{A} \not{B} \gamma^\mu = 4(A \cdot B)$

$$|\overline{\mathcal{M}_1 \mathcal{M}_2^*}| = -\frac{2e^4}{su} (p + k) \cdot (p - k') \text{Tr}[\not{p}' \not{p}]$$

noting $(p \cdot p') = -t/2$, $(k \cdot k') = -(t + Q^2)/2$

$$|\overline{\mathcal{M}_1 \mathcal{M}_2^*}| = \frac{2e^4}{su} t \{s + t + u + 2Q^2\} = \frac{2e^4}{su} t Q^2$$

Next we prove $|\overline{\mathcal{M}_2 \mathcal{M}_1^*}| = |\overline{\mathcal{M}_1 \mathcal{M}_2^*}|$. Since

$$|\overline{\mathcal{M}_2 \mathcal{M}_1^*}| = \frac{4e^4}{su} \text{Tr}[\gamma_\mu (\not{p}' - \not{k}') \gamma_\nu \not{p}' \gamma^\mu (\not{p} + \not{k}) \gamma^\nu \not{p}]$$

We only need to prove

$$\text{Tr}[\gamma^1 \gamma^2 \dots \gamma^n] = \text{Tr}[\gamma^n \dots \gamma^2 \gamma^1].$$

Since

$$\text{Tr}[A] = \text{Tr}[A^T], \quad C \gamma^{\mu T} C^{-1} = -\gamma^\mu,$$

we insert $C C^{-1}$ between all γ 's.

$$\begin{aligned} \text{Tr}[\gamma^1 \gamma^2 \dots \gamma^n] &= \text{Tr}[\gamma^{nT} \dots \gamma^{2T} \gamma^{1T}] = \text{Tr}[C \gamma^{nT} C^{-1} \dots C \gamma^{1T} C^{-1}] \\ &= (-1)^n \text{Tr}[\gamma^n \dots \gamma^2 \gamma^1] \end{aligned}$$

Problem 7.6: Derivation of the Klein–Nishina formula.

From the assumptions

$$\begin{aligned} \not{k} \not{\epsilon} &= -\not{\epsilon} \not{k}, \quad \not{k}' \not{\epsilon}' = -\not{\epsilon}' \not{k}', \quad \not{p} \not{\epsilon} = -\not{\epsilon} \not{p}, \quad \not{p}' \not{\epsilon}' = -\not{\epsilon}' \not{p}' \\ (\not{p}' + m) \not{\epsilon} u(p) &= \not{\epsilon} (-\not{p}' + m) u(p) = 0, \quad (\not{p}' + m) \not{\epsilon}' u(p) = 0 \end{aligned}$$

As we do not take photon sum

$$\begin{aligned} |\overline{\mathcal{M}}|^2 &= \frac{e^4}{2} \text{Tr} \left[(\not{p}' + m) \left\{ \frac{\not{\epsilon}' \not{\epsilon} \not{k}}{2k \cdot p} + \frac{\not{\epsilon} \not{\epsilon}' \not{k}'}{2k' \cdot p} \right\} \right. \\ &\quad \left. \times (\not{p} + m) \left\{ \frac{\not{k} \not{\epsilon} \not{\epsilon}'}{2k \cdot p} + \frac{\not{k}' \not{\epsilon}' \not{\epsilon}}{2k' \cdot p} \right\} \right] \\ &= \frac{e^4}{2} \left[\frac{I}{(2k \cdot p)^2} + \frac{II}{(2k \cdot p)(2k' \cdot p)} + \frac{III}{(2k' \cdot p)(2k \cdot p)} + \frac{IV}{(2k' \cdot p)^2} \right] \end{aligned}$$

$I V$ can be obtained from I by replacing with $\varepsilon \rightarrow \varepsilon', k \rightarrow -k'$. $I I = I I I$ with the same reason as the Problem 7.4. We only need to calculate I and $I I$. As the rest is straightforward but lengthy calculations, we only show results of each term.

$$\begin{aligned}
 I &= 8(k \cdot p)[(k' \cdot p) + 2(k \cdot \varepsilon')^2] \\
 I I + I I I &= 16[(k \cdot p)(k' \cdot p)\{2(\varepsilon \cdot \varepsilon')^2 - 1\} - (k \cdot \varepsilon')^2(k' \cdot p) \\
 &\quad + (k' \cdot \varepsilon)^2(k \cdot p)] \\
 I V &= I(\varepsilon \rightarrow \varepsilon', k \rightarrow -k') = 8(k' \cdot p)[(k \cdot p) - 2(k' \cdot \varepsilon)^2] \\
 |\overline{\mathcal{M}}|^2 &= \frac{e^4}{(k \cdot p)^2}(k \cdot p)[(k' \cdot p) + 2(k \cdot \varepsilon')^2] \\
 &\quad + \frac{e^4}{(k' \cdot p)^2}(k' \cdot p)[(k \cdot p) - 2(k' \cdot \varepsilon)^2] \\
 &\quad + 2\frac{e^4}{(k \cdot p)(k' \cdot p)}[(k \cdot p)(k' \cdot p)\{2(\varepsilon \cdot \varepsilon')^2 - 1\} - (k \cdot \varepsilon')^2(k' \cdot p) \\
 &\quad + (k' \cdot \varepsilon)^2(k \cdot p)]
 \end{aligned}$$

Using $(k \cdot p) = m\omega$, $(k' \cdot p) = m\omega'$

$$\begin{aligned}
 |\overline{\mathcal{M}}|^2 &= e^4 \left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} - 2 + 4(\varepsilon \cdot \varepsilon')^2 \right] \\
 \therefore \frac{d\sigma}{d\Omega} \Big|_{\text{LAB}} &= \left(\frac{k'}{k} \right)^2 \frac{|\overline{\mathcal{M}}|^2}{64\pi^2 m^2} = \frac{\alpha^2}{4m^2} \left(\frac{\omega'}{\omega} \right)^2 \\
 &\quad \times \left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} - 2 + 4(\varepsilon \cdot \varepsilon')^2 \right]
 \end{aligned}$$

Problem 7.7 (2):

$$\begin{aligned}
 S_{fi} - \delta_{fi} &= -i \int d^4x \langle \gamma(k) | A_\mu | 0 \rangle j^\mu(x) = -i \epsilon_\mu(\lambda) \hat{j}(k)^\mu \\
 dP_\gamma &= |S_{fi} - \delta_{fi}|^2 \frac{d^3k}{(2\pi)^3 2\omega} = d^3k (2\pi)^3 2\omega \\
 &\quad \times \sum_\lambda e^2 \left[\frac{(\epsilon(\lambda) \cdot p_f)}{(k \cdot p_f)} - \frac{(\epsilon(\lambda) \cdot p_i)}{(k \cdot p_i)} \right]^2
 \end{aligned}$$

Chapter 8: Renormalization

Chapter 9: Symmetries

Problem 9.2: To prove the Theorem 9.1, we put

$$f(t) = e^{tA} B e^{-tA}, \quad f(0) = B$$

differentiating by the parameter t

$$\frac{\partial f(t)}{\partial t} = e^{tA}[A, B]e^{-tA} = [A, f(t)]$$

This is a differential equation of the first rank and uniquely defines a function if the initial condition at $t = 0$ is given. By changing A to tA , the right hand side of Eq. (9.15) is shown to satisfy the same differential equation and the same initial value. So both sides of the Eq. (9.1) are the same functions.

To prove the Theorem 9.2, change $A \rightarrow tA$, $B \rightarrow tB$ and differentiate by t . The left hand side is

$$g(t) = e^{tA+tB}, \quad g'(t) = \frac{dg(t)}{dt} = (A + B)g(t), \quad g(0) = 1$$

Differentiating the right hand side $[= h(t)]$ after the above change

$$h'(t) = e^{tA} A e^{tB} e^{-\frac{1}{2}t^2[A, B]} + e^{tA} e^{tB} (B - t[A, B]) e^{-\frac{1}{2}t^2[A, B]}, \quad h(0) = 1$$

Using Eq. (9.15) and the condition

$$\begin{aligned} e^{tB} A &= \left(A + [tB, A] + \frac{1}{2}[tB, [tB, A]] + \dots \right) e^{tB} \\ &= (A + [tB, A]) e^{tB} \\ \therefore e^{tB}[A, B] &= ([A, B] + [[tB, A], B]) e^{tB} \\ &= [A, B] e^{tB} \end{aligned}$$

Similarly

$$\begin{aligned} e^{tA}[A, B] &= [A, B] e^{tA} \\ \therefore h'(t) &= [(A - t[A, B])e^{tA} + e^{tA} B] e^{tB} e^{-\frac{1}{2}t^2[A, B]} \end{aligned}$$

From Eq. (9.16b), $e^{tA} B = (B + t[A, B])e^{tA}$

$$\therefore h'(t) = (A + B)e^{tA} e^{tB} e^{-\frac{1}{2}t^2[A, B]} = (A + B)h(t)$$

Since $g(t)$ and $h(t)$ satisfy the same differential equations and initial conditions, both are the same function. The second equation can be proved by noting

$$\begin{aligned} e^A e^B &= e^{A+B} e^{\frac{1}{2}[A, B]}, \quad e^B e^A = e^{A+B} e^{-\frac{1}{2}[A, B]}, \\ \therefore e^A e^B &= e^B e^A e^{[A, B]} \end{aligned}$$

Problem 9.3:

Putting $T_\lambda^J = \langle \lambda_3 \lambda_4 | T^J(E) | \lambda_1 \lambda_2 \rangle$ and using the partial wave expansion

$$\begin{aligned} \sigma_{\text{TOT}} &= \int d\Omega \sum_{\lambda_3, \lambda_4} |f_{\lambda_3 \lambda_4, \lambda_1 \lambda_2}(\theta)|^2 \\ &= \int d\Omega \sum_{\lambda_3 \lambda_4} \left[\frac{1}{4p^2} \sum_{JJ'} (2J+1)(2J'+1) T_\lambda^J T_\lambda^{J'*} d_{\lambda\mu}^J d_{\lambda\mu}^{J'} \right] \\ &= \frac{\pi}{p^2} \sum_{J, \lambda_3 \lambda_4} (2J+1) |T_\lambda^J|^2 \end{aligned}$$

We have used $\int_{-1}^1 d \cos \theta d_{MM'}^J(\theta) d_{MM'}^{J'}(\theta) = 2\delta_{JJ'}/(2J+1)$ to derive the last equality. Then

$$\begin{aligned} \text{Im} f_{\lambda_3 \lambda_4, \lambda_1 \lambda_2}(0) &= \frac{1}{2p} \sum_J (2J+1) \text{Im} T_{\lambda}^J d_{\lambda \mu}^J(0) \\ &= \frac{1}{4p} \sum_{J, \lambda_3, \lambda_4} (2J+1) |T_{\lambda}^J|^2 = \frac{p}{4\pi} \sigma_{\text{TOT}} \end{aligned}$$

In going to the third equality we have used $\text{Im} T_{\lambda}^J = \text{Im}[(e^{2i\delta_J} - 1)/i] = 1 - \cos 2\delta_J$, $|T_{\lambda}^J|^2 = 2(1 - \cos 2\delta_J)$.

Problem 9.4: The Maxwell equation says $\partial_{\mu} \partial^{\mu} A^{\nu} = j^{\nu}$, the properties of A^{ν} and j^{ν} are equal. Since $j^{\mu} = (\rho, \rho \mathbf{v})$, $A'^0(-t) = A^0(t)$, $A'(-t) = -A(t)$. For the transformation properties of the quantized fields, refer Eq. (9.133) and Eq. (9.134). The properties of $\mathbf{E}$ and $\mathbf{B}$ can be obtained by considering $\mathbf{E} = -\nabla A^0 - \partial_t \mathbf{A}$, $\mathbf{B} = \nabla \times \mathbf{A}$.

Problem 9.6: Use Table 4.2 in Chap. 4. Or use

$$\begin{aligned} \frac{i}{2} \gamma^5 \sigma^{\mu\nu} &= -\frac{i}{4} \epsilon_{\rho\sigma}^{\mu\nu} \gamma^{\rho\sigma}, \\ \frac{1}{2} \epsilon_{\rho\sigma}^{\mu\nu} F^{\rho\sigma} &= \tilde{F}^{\mu\nu}, \quad \tilde{F}^{\mu\nu}(E, B) = F^{\mu\nu}(-B, -E) \\ \frac{i}{2} \gamma^5 \sigma^{\mu\nu} F_{\mu\nu} &= -\frac{1}{2} \sigma^{\mu\nu} \tilde{F}_{\mu\nu} = -i \boldsymbol{\alpha} \cdot \mathbf{B} - \boldsymbol{\sigma} \cdot \mathbf{E} \end{aligned}$$

Problem 9.7*: Define operators

$$\begin{aligned} \sum_k (b_k^{\dagger} - a_k^{\dagger})(b_k - a_k) &= \sum_k (N_k + T_k) = N + T, \\ N_k &= a_k^{\dagger} a_k + b_k^{\dagger} b_k, \quad T_k = -(a_k^{\dagger} b_k + b_k^{\dagger} a_k) \end{aligned}$$

Then $[N_k, T_q] = [N, T] = 0$ and

$$\begin{aligned} [N_k, a_q] &= -\delta_{qk} a_q, \quad [N_k, a_q^{\dagger}] = \delta_{qk} a_q^{\dagger}, \quad [N_k, b_q] = -\delta_{qk} b_q, \\ [N_k, b_q^{\dagger}] &= \delta_{qk} b_q^{\dagger} \\ [T_k, a_q] &= \delta_{qk} b_q, \quad [T_k, a_q^{\dagger}] = -\delta_{qk} b_q^{\dagger}, \quad [T_k, b_q] = \delta_{qk} a_q, \\ [T_k, b_q^{\dagger}] &= -\delta_{qk} a_q^{\dagger} \end{aligned}$$

Applying the above commutation relations to the formulas

$$\begin{aligned} \text{If } [A, [A, B]] &= [B, [B, A]] = 0 \rightarrow e^{A+B} = e^{-\frac{1}{2}[A, B]} e^A e^B = e^A e^B e^{-\frac{1}{2}[A, B]} \\ e^A B e^{-A} &= B + [A, B] + \frac{1}{2!} [A, [A, B]] + \dots \end{aligned}$$

we obtain

$$\begin{aligned}
 e^{i\alpha(N+T)} &= e^{i\alpha T} e^{i\alpha N} = e^{i\alpha N} e^{i\alpha T} \\
 e^{i\alpha N}(a_q, b_q) e^{-i\alpha N} &= (e^{-i\alpha} a_q, e^{-i\alpha} b_q), \\
 e^{i\alpha N}(b_q^\dagger, a_q^\dagger) e^{-i\alpha N} &= (e^{i\alpha} b_q^\dagger, e^{i\alpha} a_q^\dagger) \\
 e^{i\alpha T}(a_q, b_q) e^{-i\alpha T} &= (a_q \cos \alpha + i b_q \sin \alpha, b_q \cos \alpha + i a_q \sin \alpha) \\
 e^{i\alpha T}(a_q^\dagger, b_q^\dagger) e^{-i\alpha T} &= (a_q^\dagger \cos \alpha - i b_q^\dagger \sin \alpha, b_q^\dagger \cos \alpha - i a_q^\dagger \sin \alpha)
 \end{aligned}$$

$$\text{Choosing } \alpha = \pi/2, C = e^{i\frac{\pi}{2}(N+T)} = e^{i\frac{\pi}{2}N} e^{i\frac{\pi}{2}T} = e^{i\frac{\pi}{2}T} e^{i\frac{\pi}{2}N}$$

$$\begin{aligned}
 C a_q C^{-1} &= \exp\left[i\frac{\pi}{2}T\right] \exp\left[i\frac{\pi}{2}N\right] a_q \exp\left[-i\frac{\pi}{2}N\right] \exp\left[-i\frac{\pi}{2}T\right] \\
 &= \exp\left[i\frac{\pi}{2}T\right] (-i a_q) \exp\left[-i\frac{\pi}{2}T\right] = b_q \\
 C b_q^\dagger C^{-1} &= \exp\left[i\frac{\pi}{2}T\right] \exp\left[i\frac{\pi}{2}N\right] b_q^\dagger \exp\left[-i\frac{\pi}{2}N\right] \exp\left[-i\frac{\pi}{2}T\right] \\
 &= \exp\left[i\frac{\pi}{2}T\right] (i b_q^\dagger) \exp\left[-i\frac{\pi}{2}T\right] = a_q^\dagger \\
 \therefore C \varphi C^{-1} &= \varphi^\dagger, \quad \text{Similarly } C \varphi^\dagger C^{-1} = \varphi
 \end{aligned}$$

Problem 9.8: As the two π^0 's are identical particles, their wave function must be totally symmetric, which means $L = \text{even}$. The angular momentum conservation requires $J = L = 1$ and the two conditions are incompatible.

Problem 9.9: The positronium is a bound state of e^-e^+ and its $C = (-1)^{L+S}$. For $L = 0$, $C = +, -$ depending on $S = 0, 1$ which allows $2\gamma, 3\gamma$, respectively, but not vice versa.

Chapter 10: Path Integral I

Problem 10.2:

$$\begin{aligned}
 \frac{\delta S}{\delta q(t')} &= \lim_{\varepsilon} \frac{1}{\varepsilon} \int dt L(q(t) + \varepsilon \delta(t-t'), \dot{q}(t) + \varepsilon \frac{d}{dt} \delta(t-t')) - L(q(t), \dot{q}(t)) \\
 &= \int dt \left[\frac{\partial L}{\partial q} \varepsilon \delta(t-t') + \frac{\partial L}{\partial \dot{q}} \varepsilon \frac{d}{dt} \delta(t-t') \right] = \left. \frac{\partial L}{\partial q} \right|_{t=t'} - \left. \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right|_{t=t'}
 \end{aligned}$$

Problem 10.4: Writing down the solution as $x = A \cos \omega t + B \sin \omega t$ with the condition $x(t_i = 0) = q_i$, $x(t_f = T) = q_f$ gives

$$x = q_i \cos \omega t + \frac{q_f - q_i \cos \omega T}{\sin \omega T} \sin \omega t$$

Then calculate

$$S_c = \int_0^T dt \left[\frac{m \dot{x}^2}{2} - \frac{m \omega^2}{2} x^2 \right]$$

Problem 10.7: (4) The average energy of a photon at temperature T is

$$\begin{aligned}\bar{E} &= -\frac{\partial \ln Z}{\partial \beta} = \frac{\partial}{\partial \beta} \ln 2 \sinh \frac{\hbar \omega \beta}{2} = \frac{\hbar \omega}{2} \tanh \frac{\hbar \omega \beta}{2} \\ &= \frac{\hbar \omega}{2} + \frac{\hbar \omega}{e^{\hbar \omega \beta} - 1}\end{aligned}$$

The number of photons per unit phase space is $2Vd^3p/(2\pi\hbar)^3 = (8\pi/c^3)V\nu^2 d\nu$ where the factor 2 represents two polarization states. Expressing $\hbar\omega = h\nu$ and omitting the zero point energy (the first term), the average photon energy per unit frequency per unit volume is given by $(8\pi h/c^3)\nu^3 d\nu/(e^{h\nu/kT} - 1)$ which is the Planck's formula.

Chapter 11: Path Integral II

Problem 11.1: The function f is expressed as a polynomial of η_i 's, but from the conditions $\int d\eta = 0$ and anticommutativity of all the G -variables, the only surviving term in the integrand is one which includes all the $\eta_1 \cdots \eta_N$. As

$$\eta_1 \cdots \eta_N = \prod_{i=1}^N \left(\sum_{j_k} (U^{-1})_{i j_k} \xi_{j_k} \right)$$

The right hand side is a polynomial of ξ 's, but as any overlap of the same variable $(\xi)^2$ gives zero, only a term which includes all the $\xi_1 \cdots \xi_N$ survives. Therefore

$$\begin{aligned}\eta_1 \cdots \eta_N &= \sum \varepsilon^{j_1 j_2 \cdots j_N} (U^{-1})_{j_1 1} (U^{-1})_{j_2 2} \cdots (U^{-1})_{j_N N} \xi_1 \xi_2 \cdots \xi_N \\ &= (\det U^{-1}) \xi_1 \xi_2 \cdots \xi_N\end{aligned}$$

Writing the coefficient of the surviving polynomial as A ,

$$\begin{aligned}\int \prod_{i=1}^N d\eta_i f(\eta_1, \cdots, \eta_N) &= \int \prod_{i=1}^N d\eta_i (A \eta_1 \cdots \eta_N) \\ &= J_G \int \prod_{i=1}^N d\xi_i A \det(U^{-1}) \xi_1 \cdots \xi_N,\end{aligned}$$

On the other hand, because of Eq. (11.85b)

$$\int d\eta_1 \cdots d\eta_N \eta_1 \cdots \eta_N = \int d\xi_1 \cdots d\xi_N \xi_1 \cdots \xi_N$$

Comparison of the two equations gives $J_G = \det U$.

Problem 11.2: Consider an orthogonal transformation

$$\eta = O^T \xi, \quad \eta^T A \eta = \xi^T O A O^T \xi = \xi^T A' \xi$$

As n -dimensional integration survives only when the integrand includes all the η_i 's, only $n/2$ -th power of the expansion survives.

$$\begin{aligned} I_n &= \int d\eta_1 \cdots d\eta_n e^{-\frac{1}{2}\eta^T A \eta} = J \int d\xi \cdots d\xi_n e^{-\frac{1}{2}\xi^T A' \xi} \\ &= \frac{(-1)^{n/2}}{\left(\frac{n}{2}\right)!} \int d\xi_1 \cdots d\xi_n \left(\xi^T A' \xi\right)^{n/2} \end{aligned}$$

The equation follows because

1. Since $\mathbf{O}$ is orthogonal transformation Jacobian $J = \det \mathbf{O} = 1$.
2. In the expansion only terms that have exactly n factors can contribute to the integral.

Now any antisymmetric $(n \times n)$ -dimensional matrix $\mathbf{A}$ with even n can be brought with a unitary transformation to the form

$$\mathbf{A}' = \mathbf{U} \mathbf{A} \mathbf{U}^\dagger = \begin{bmatrix} 0 & a_1 & \cdots & 0 & 0 \\ -a_1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n/2} \\ 0 & 0 & \cdots & -a_{n/2} & 0 \end{bmatrix} \quad (\text{I.3})$$

for odd n the matrix has one more row and one more column with all zeros. Then the determinant is easily obtained from the above form (See proof of Eq. (I.3) below)

$$\det \mathbf{A} = \det \mathbf{A}' = \begin{cases} (a_1)^2 \cdots (a_{n/2})^2 & n = \text{even} \\ 0 & n = \text{odd} \end{cases}$$

Assuming $\mathbf{A}'$ has this form,

$$\begin{aligned} I_n &= \frac{1}{\left(\frac{n}{2}\right)!} \int d\xi_1 \cdots d\xi_n \left(-\frac{1}{2}\xi^T A' \xi\right)^{n/2} \\ &= \frac{1}{\left(\frac{n}{2}\right)!} \int d\xi_1 \cdots d\xi_n \left[-\frac{1}{2} \sum_{j=1}^{n/2} (\xi_{2j-1} A'_{2j-1,2j} \xi_{2j} - \xi_{2j} A'_{2j,2j-1} \xi_{2j-1})\right]^{n/2} \\ &= \frac{(-1)^{n/2}}{\left(\frac{n}{2}\right)!} \int d\xi_1 \cdots d\xi_n \left[\sum_{j=1}^{n/2} a_j \xi_{2j-1} \xi_{2j}\right]^{n/2} \\ &= \frac{(-1)^{n/2}}{\left(\frac{n}{2}\right)!} \int d\xi_1 \cdots d\xi_n \sum_{j_1} \cdots \sum_{j_{n/2}} \xi_{2j_1-1} \xi_{2j_1} \cdots \xi_{2j_{n/2}-1} \xi_{2j_{n/2}} a_{j_1} \cdots a_{j_{n/2}} \end{aligned}$$

There can be no two equal ξ 's in the integral. Since ξ appear always pairwise, they can be commuted to normal ordering $\xi_n \cdots \xi_1$ with a sign change $(-1)^{n/2}$.

The product $a_{j_1} \cdots a_{n/2}$ gives just $\sqrt{\det A'}$. Then the role of sum is just to multiply $(n/2)!$ with the same result. All things put together we obtain

$$\begin{aligned} I_n (n \text{ even}) &= \int d\eta_1 \cdots d\eta_n e^{-\frac{1}{2}\eta^T A \eta} \\ &= \int d\xi_1 \cdots \xi_n \xi_n \cdots \xi_1 \sqrt{\det A'} = \sqrt{\det A} \end{aligned}$$

The result is to be contrasted with the normal (or bosonic) Gaussian integral

$$\int dx_1 \cdots dx_n e^{-\frac{1}{2}x^T A x} = \frac{(2\pi)^{n/2}}{\sqrt{\det A}}$$

(Proof of Eq. (I.3)): First consider the simple case of $n = 2$.

$$A = \begin{bmatrix} 0 & A_{12} \\ -A_{12} & 0 \end{bmatrix}$$

For the general case, we use the fact that iA is a hermitian matrix. It can be diagonalized by a unitary transformation V

$$V i A V^\dagger = D$$

where D is real and diagonal whose elements are solutions to the secular equation

$$\det |iA - \lambda \mathbf{1}| = 0$$

Since $A^T = -A$, we have $\det |(iA - \lambda \mathbf{1})^T| = \det |-iA - \lambda \mathbf{1}| = 0$. Therefore, if λ is a solution, so is $(-\lambda)$ and D is of the form

$$D = \begin{bmatrix} a_1 & 0 & \cdots & \cdots & 0 \\ 0 & -a_1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & \cdots & \cdots & a_{n/2} & 0 \\ 0 & \cdots & \cdots & 0 & -a_{n/2} \end{bmatrix}$$

From this one sees that $\det A = i^{2n} \det D = (a_1 \cdots a_n)^2$. To put D in the form Eq. (I.3), we use the 2×2 matrix

$$S_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}$$

which has the property

$$(-i)S_2 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} S_2^\dagger = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Thus $S(-iD)S^\dagger = A'$ for

$$S = \begin{bmatrix} S_2 & 0 & \cdots & 0 \\ 0 & S_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & S_2 \end{bmatrix}$$

Then the matrix A' defined by

$$A' = S(-iD)S^\dagger = S[-iV(iA)V^\dagger]S^\dagger = (SV)A(SV)^\dagger = UAU^\dagger$$

reproduces Eq. (I.3). Since both A and A' is real so is U , namely, U is an orthogonal matrix and Eq. (I.3) is proved.

Problem 11.4: We start from the second order interaction Z_2 given in Eq. (11.168)

$$\begin{aligned} Z_2 \equiv & \frac{N}{2!} \left\{ \int d^4x \left(-\frac{1}{i} \frac{\delta}{\delta \eta(x)} \right) (-iq\gamma^\mu) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\mu(x)} \right) \right\}^2 \\ & \times Z_0[\bar{\eta}, \eta, J] \\ = & \frac{N}{2} \left\{ \int d^4y \left(-\frac{1}{i} \frac{\delta}{\delta \eta(y)} \right) (-iq\gamma^\rho) \left(\frac{1}{i} \frac{\delta}{\delta \bar{\eta}(y)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J^\rho(y)} \right) \right\} \\ & \times \left[- \int d^4x d^4z_1 d^4z_2 d^4z_3 \bar{\eta}(z_2) S_F(z_2 - x) (-iq\gamma^\mu) \right. \\ & \left. \times S_F(x - z_3) \eta(z_3) D_{F\mu\nu}(x - z_1) J^\nu(z_1) Z_0[\bar{\eta}, \eta, J] \right] \end{aligned} \quad (I.4)$$

To make an internal photon line, $\delta/\delta J^\mu(x)$ has to act on $J^\nu(z_1)$, producing $D_{F\mu\rho}(x - y)$. There are two ways to make the electron self energy contributions: (a) one pulls down another electron line and connects it to the outgoing electron. Namely, $\delta/\delta \eta(y)$ acts on $\eta(z_3)$, and $\delta/\delta \bar{\eta}(y)$ acts on $(\delta Z_0/\delta \bar{\eta})$. (b) the other connecting the new line $(\delta Z_0/\delta \eta)$ to the incoming electron of the first vertex. The process (a) is expressed as:

$$\begin{aligned} Z_{26a} = & -\frac{N}{2} \int d^4x d^4y d^4z_2 d^4w \{ \bar{\eta}(z_2) S_F(z_2 - x) (-iq\gamma^\mu) \\ & \times S_F(x - y) (-iq\gamma^\rho) S_F(y - w) \eta(w) D_{F\mu\rho}(x - y) \} Z_0 \end{aligned}$$

The process (b) is given

$$\begin{aligned} Z_{26b} = & -\frac{N}{2} \int d^4x d^4y d^4z_1 d^4w \{ \bar{\eta}(w) S_F(w - y) (-iq\gamma^\rho) \\ & \times S_F(y - x) (-iq\gamma^\mu) S_F(x - z_1) \eta(z_1) D_{F\mu\rho}(x - y) \} Z_0 \end{aligned}$$

By changing the integration variables we see that Z_{26b} is identical to Z_{26a} and that serves merely to cancel the factor 1/2. This is a correction to the electron propaga-

tor. Since the electron propagator is given by

$$\begin{aligned}
 \langle \Omega | T[\hat{\psi}(x)\hat{\bar{\psi}}(y)] | \Omega \rangle &= \left(\frac{1}{i}\right)^2 \frac{\delta^2}{\delta \eta(y) \delta \bar{\eta}(x)} Z[\bar{\eta}, \eta, J] \Big|_{\bar{\eta}=\eta=J=0} \\
 &= \left(\frac{1}{i}\right)^2 \frac{\delta^2}{\delta \eta(y) \delta \bar{\eta}(x)} \{Z_0 + Z_1 + (\cdots + Z_{26} + \cdots) + \cdots\} \Big|_{\bar{\eta}=\eta=J=0} \\
 &= i S_F(x-y) + \int d^4 w d^4 z S_F(x-w)(-i q \gamma^\mu) S_F(w-z)(-i q \gamma^\nu) \\
 &\quad \times S_F(z-y) D_{F\mu\nu}(w-z)
 \end{aligned}$$

Using the momentum representation of the propagators Eq. (11.65) and Eq. (11.119) and integrating over w and z , the second term becomes:

$$\begin{aligned}
 &\int \frac{d^4 p}{(2\pi)^4} e^{-i p \cdot (x-y)} \frac{1}{\not{p} - m + i\epsilon} \\
 &\quad \times \left[\int \frac{d^4 k}{(2\pi)^4} \frac{-g_{\mu\nu}}{k^2} (-i q \gamma^\mu) \frac{1}{\not{p} - \not{k} - m + i\epsilon} (-i q \gamma^\nu) \right] \frac{1}{\not{p} - m + i\epsilon}
 \end{aligned}$$

Then the momentum representation of the electron propagator including the second order correction is expressed as

$$\begin{aligned}
 i \tilde{S}_F(p) &= \frac{i}{\not{p} - m} + \frac{i}{\not{p} - m} i \Sigma_e(p) \frac{i}{\not{p} - m} \\
 i \Sigma_e(p) &= \int \frac{d^4 k}{(2\pi)^4} \frac{-i g_{\mu\nu}}{k^2} (-i q \gamma^\mu) \frac{i}{\not{p} - \not{k} - m + i\epsilon} (-i q \gamma^\nu)
 \end{aligned}$$

Problem 11.5: The reduction formula for the Compton scattering

$$\gamma(k\lambda) + e(pr) \rightarrow \gamma(k'\lambda') + e(p's)$$

where λ, λ' are the photon polarizations and r, s are electron spin orientations is given by

$$\begin{aligned}
 &\text{out} \langle k'\lambda' p's | k\lambda, pr \rangle_{\text{in}} \\
 &= \int d^4 x_1 d^4 x_3 d^4 x_2 d^4 x_4 \varepsilon_\nu^*(k', \lambda') f_{k'}(x_3) (\partial_\rho \partial^\rho)_{x_3} \bar{U}_{p'}(x_4) (i \overleftrightarrow{\not{\partial}} - m)_{x_4} \\
 &\quad \times \frac{\delta^4 Z[\bar{\eta}, \eta, J]}{\delta \eta(x_2) \delta J^\mu(x_1) \delta \bar{\eta}(x_4) \delta J^\nu(x_3)} \Big|_{\bar{\eta}=\eta=J=0} \\
 &\quad \times (-i \overleftarrow{\not{\partial}} - m)_{x_2} U_p(x_2) (\overleftarrow{\partial}_\sigma \partial^\sigma)_{x_1} \varepsilon_\mu(k, \lambda) f_k(x_1) \quad (I.5)
 \end{aligned}$$

The generating functional that contributes to the Compton scattering was given by Eq. (11.171):

$$\begin{aligned}
 Z_{\text{Compton}} &= i \int d^4 x d^4 y d^4 z_1 d^4 z_2 d^4 w_1 d^4 w_3 \left\{ J^\sigma(w_1) D_{F\sigma\rho}(w_1 - y) \right. \\
 &\quad \times D_{F\alpha\beta}(x - z_1) J^\beta(z_1) \bar{\eta}(w_3) S_F(w_3 - y) (-i q \gamma^\rho) \\
 &\quad \times S_F(y - x) (-i q \gamma^\alpha) S_F(x - z_3) \eta(z_3) \left. \right\} Z_0[\bar{\eta}, \eta, J]
 \end{aligned}$$

As the photon source J^μ has no directionality, namely it could serve either as a source or a sink, there are two ways to make the photon derivatives producing two distinct scattering processes as depicted in Figure I.1.

Application of the four-fold derivative in the reduction formula to Z_{Compton} yields:

$$\begin{aligned} \frac{\delta^4 Z_{\text{Compton}}[\bar{\eta}, \eta, J]}{\delta \eta(x_2) \delta J^\mu(x_1) \delta \bar{\eta}(x_4) \delta J^\nu(x_3)} \Big|_{\bar{\eta}=\eta=J=0} &= i \int d^4x d^4y \\ &\times \left[D_{F\nu\rho}(x_3 - y) D_{F\alpha\mu}(x - x_1) S_F(x_4 - y) (-iq\gamma^\rho) \right. \\ &\times S_F(y - x) (-iq\gamma^\alpha) S_F(x - x_2) \Big] + \left[D_{F\nu\alpha}(x_3 - x) D_{F\rho\mu}(y - x_1) \right. \\ &\times S_F(x_4 - y) (-iq\gamma^\rho) S_F(y - x) (-iq\gamma^\alpha) S_F(x - x_2) \Big] \end{aligned}$$

Inserting the above equation in Eq. (I.5), we obtain the scattering amplitude for the Compton scattering

$$\begin{aligned} \text{out} \langle k' \lambda', p' s | k \lambda, p r \rangle_{\text{in}} &= i \int d^4x d^4y \\ &\times \bar{U}_{p'}(y) \left[f_{k'}^*(y) f_k(x) \varepsilon_\nu^*(k' \lambda') (-iq\gamma^\nu) S_F(y - x) (-iq\gamma^\mu) \varepsilon_\mu(k, \lambda) \right. \\ &\left. + f_k^*(x) f_k(y) \varepsilon_\mu(k, \lambda) (-iq\gamma^\mu) S_F(y - x) (-iq\gamma^\nu) \varepsilon_\nu^*(k', \lambda') \right] U_p(x) \end{aligned}$$

Insertion of

$$\begin{aligned} f_k(x) &= e^{-ik \cdot x}, \quad U_p(x) = u_r(p) e^{-ip \cdot x}, \\ S_F(x - y) &= \int \frac{d^4p}{(2\pi)^4} e^{-ip \cdot (x-y)} \frac{\not{p} + m}{p^2 - m^2 + i\varepsilon} \end{aligned}$$

and the integration of x and y yields two δ functions and after integrating over p gives the final form:

$$\begin{aligned} \text{out} \langle k' \lambda', p' s | k \lambda, p r \rangle_{\text{in}} &= -(2\pi)^4 i \delta^4(k + p - k' - p') e^2 \bar{u}_s(p') \\ &\times \left[\varepsilon'^* \frac{\not{p} + \not{k} + m}{(p + k)^2 - m^2 + i\varepsilon} \varepsilon / + \varepsilon / \frac{\not{p} - \not{k}' + m}{(p - k')^2 - m^2 + i\varepsilon} \varepsilon'^* \right] u_r(p) \end{aligned}$$

where $\varepsilon / = \varepsilon_\mu(k, \lambda) \gamma^\mu$, $\varepsilon' / = \varepsilon_\nu(k', \lambda') \gamma^\nu$. The Compton scattering amplitude Eq. (7.51) has thus been reproduced.

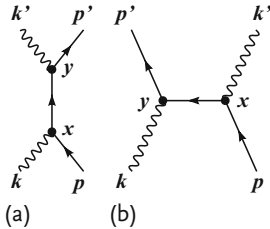


Figure I.1 Compton Scattering.

Chapter 12: Accelerators and Detectors

Problem 12.3:

As the mean n is large and the standard deviation $\sigma/n = 1/\sqrt{n}$ is small, r is large and distributes around n . Let

$$F(r) = \ln \left(\frac{n^r}{r!} e^{-n} \right) \simeq r \ln n - n - \left\{ \ln \sqrt{2\pi} + \left(r + \frac{1}{2} \right) \ln r - r \right\}$$

Expand $F(r)$ in Taylor series around $r = n$ and use up to quadratic terms in $(r - n)^2$

$$\begin{aligned} F(n) &= -\ln \sqrt{2\pi n}, \quad F'(n) = -\frac{1}{2n}, \quad F''(n) = -\left(\frac{1}{n} - \frac{1}{2n^2} \right) \\ F(r) &\simeq -\ln \sqrt{2\pi n} - \frac{1}{2n}(r - n) - \frac{1}{2} \left(\frac{1}{n} - \frac{1}{2n^2} \right) (r - n)^2 \\ &= -\ln \sqrt{2\pi n} - \frac{1}{2n}(r - n)(r - n + 1) + \frac{1}{4n^2}(r - n)^2 \end{aligned}$$

Neglecting the third term and approximating $n + 1 \simeq n$, we obtain

$$\begin{aligned} F(r) &= -\ln \sqrt{2\pi n} - \frac{1}{2n}(r - n)^2 \quad \text{or} \\ \frac{n^r}{r!} e^{-n} &\xrightarrow{n \rightarrow \text{large}} \frac{1}{\sqrt{2\pi n}} \exp \left[-\frac{(r - n)^2}{2n} \right] \end{aligned}$$

Chapter 13: Hadron Spectroscopy

Problem 13.1:

1. The total angular momentum in the final state is $J = L$ and must be zero. The parity of the final state is $P_1 P_2 (-1)^L = P_1 P_2 = P_1^2 = +1$. The last equality holds because the two particles are identical.
2. Let the orbital angular momentum of two particles ℓ and that of the relative motion between the CM of the two-particle and the third particle L . From the requirement $\mathbf{J} = \boldsymbol{\ell} + \mathbf{L} = 0$, $\ell = L$. Then the parity becomes $(-1)^3 (-1)^{\ell+L} = -$. The problem 2) and 3) were used to prove the parity violation in the weak decay of $K^0 \rightarrow 2\pi/3\pi$.
3. We set the z -axis along the photon direction in the rest frame of the initial particle. Then $J_z = L_z + S_z = S_z \neq 0$.
4. With the same setting as the problem 2), the angular momentum conservation requires $\ell + L = 1$. Then $\ell = L$, $L \pm 1$, therefore both parities are allowed.
5. $L + 1 = 1 \rightarrow L = 0, 1, 2$; $P = (-1)^2 (-1)^L$. $\therefore L = 1$ is allowed.
6. The orbital angular momentum L_i in the initial state can be considered as zero which means P is (-1) . As the final state consists of two identical bosons with $S = 0$. L must be even which means $P = +$ and the parity is violated.

Problem 13.2: The decay to 2γ is obviously of electromagnetic origin. With the same logic as used to determine the spin and parity of π^0 (see arguments in Sect. 13.2), the decay mode $\eta \rightarrow 2\gamma$ tells that the spin of η cannot be 1, and that C-parity is (+). Since $I = 0$ the G-parity is $G = +1$. However, the 3π final state has $G = -1$ which means $I = 1$. Therefore, the decay violates isospin and G-parity and cannot be strong. In fact, other branch modes have comparable branching ratio with the 2γ mode and are consistent that they are all induced by the electromagnetic interaction. As the electromagnetic interaction conserves parity, the spin parity of η is fixed to be $J^P = 0^-$.

Problem 13.3: From Eq. (13.34) Eq. (13.35)

$$\begin{aligned}
 {}^3H\pi^+ &= \left| \frac{1}{2}, -\frac{1}{2} \right\rangle |1, 1\rangle = \sqrt{\frac{1}{3}} \left| \frac{3}{2}, \frac{1}{2} \right\rangle |1, 0\rangle + \sqrt{\frac{2}{3}} \left| \frac{1}{2}, \frac{1}{2} \right\rangle |1, 0\rangle \\
 {}^3He\pi^0 &= \left| \frac{1}{2}, \frac{1}{2} \right\rangle |1, 0\rangle = \sqrt{\frac{2}{3}} \left| \frac{3}{2}, \frac{1}{2} \right\rangle |1, 0\rangle - \sqrt{\frac{1}{3}} \left| \frac{1}{2}, \frac{1}{2} \right\rangle |1, 0\rangle \\
 pd &= |1/2, 1/2\rangle \quad \therefore \quad \frac{T(pd \rightarrow {}^3H\pi^+)}{T(pd \rightarrow {}^3He\pi^0)} = -\frac{\sqrt{2/3} T_{1/2}}{\sqrt{1/3} T_{1/2}} = -\sqrt{2}
 \end{aligned}$$

Problem 13.4:

$$\begin{aligned}
 pp &= |1/2, -1/2\rangle |1/2, 1/2\rangle = |1, 1\rangle \\
 np &= |1/2, -1/2\rangle |1/2, 1/2\rangle = (|1, 0\rangle - |0, 0\rangle)/\sqrt{2} \\
 d\pi^+ &= |1, 1\rangle, \quad d\pi^0 = |1, 0\rangle \\
 \frac{T(pp \rightarrow d\pi^+)}{T(np \rightarrow d\pi^0)} &= \frac{T_1}{T_1/\sqrt{2}} = \sqrt{2} \rightarrow \frac{d\sigma(pp \rightarrow d\pi^+)}{d\sigma(np \rightarrow d\pi^0)} = 2
 \end{aligned}$$

Problem 13.6:

$$\begin{aligned}
 K^- n &= |1/2, -1/2\rangle |1/2, -1/2\rangle = |1, -1\rangle \\
 K^- p &= |1/2, -1/2\rangle |1/2, 1/2\rangle = (|1, 0\rangle + |0, 0\rangle)/\sqrt{2} \\
 K^0 \Xi^- &= |1/2, -1/2\rangle |1/2, -1/2\rangle = |1, -1\rangle \\
 K^+ \Xi^- &= |1/2, 1/2\rangle |1/2, -1/2\rangle = (|1, 0\rangle - |0, 0\rangle)/\sqrt{2} \\
 K^0 \Xi^0 &= |1/2, -1/2\rangle |1/2, 1/2\rangle = (|1, 0\rangle + |0, 0\rangle)/\sqrt{2} \\
 \therefore f(K^- n \rightarrow K^0 \Xi^-) &= T_1 \\
 f(K^- p \rightarrow K^0 \Xi^-) + f(K^- p \rightarrow K^0 \Xi^0) &= (T_1 - T_0)/2 + (T_1 + T_0)/2 \\
 &= T_1
 \end{aligned}$$

Problem 13.7:

$$\begin{aligned}
K^+ \Lambda &= |1/2, 1/2\rangle, \quad K^0 \Lambda^0 = |1/2, -1/2\rangle \\
\pi^+ n &= |1, 1\rangle |1/2, -1/2\rangle = \sqrt{1/3} |3/2, 1/2\rangle + \sqrt{2/3} |1/2, 1/2\rangle \\
\pi^- p &= |1, -1\rangle |1/2, +1/2\rangle = \sqrt{1/3} |3/2, -1/2\rangle - \sqrt{2/3} |1/2, -1/2\rangle \\
\therefore \frac{T(\pi^+ n \rightarrow K^+ \Lambda)}{T(\pi^- p \rightarrow K^0 \Lambda^0)} &= \frac{\sqrt{2/3}}{-\sqrt{2/3}}, \\
\rightarrow d\sigma(\pi^+ n \rightarrow K^+ \Lambda) &= d\sigma(\pi^- p \rightarrow K^0 \Lambda^0)
\end{aligned}$$

Problem 13.8:

(a) The real part and imaginary part of the scattering from Eq. (13.91) give

$$\operatorname{Re} T_\ell = \frac{\eta_\ell}{2} \sin 2\delta, \quad \operatorname{Im} T_\ell = \frac{1 - \eta_\ell}{2} + \eta_\ell \sin^2 \delta_\ell$$

$\eta = 1$ for elastic scattering and $\eta < 1$ at higher energy where inelastic scattering occurs and is a function of energy.

When $\delta = \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$,

$$\begin{aligned}
&\left(\operatorname{Re} T = \frac{\eta}{2}, \operatorname{Im} T = 0.5 \right), \quad \left(\operatorname{Re} T = 0, \operatorname{Im} T = \frac{1 + \eta}{2} \right), \\
&\left(\operatorname{Re} T = -\frac{\eta}{2}, \operatorname{Im} T = 0.5 \right),
\end{aligned}$$

and this means each trajectory circles around the center of the Argand diagram (counter-)clock wise when the force is attractive or repulsive. This is the case for a single amplitude. (b) When there is a non resonant repulsive force which does not change appreciably as a function of energy, the total amplitude is sum of the two (with proper normalization to lie within the circle). We consider, for simplicity, for the case when the total amplitude is the average of the two [see Fig. 1.2]. Then examples in Fig. 13.15 show various cases satisfying conditions (b–d).

Problem 13.9: Denote 3-momentum and kinetic energy of 3 pions as $\mathbf{k}_i$ and T_i which satisfy

$$\begin{aligned}
\mathbf{k}_1 + \mathbf{k}_2; \mathbf{k}_3 &= 0, \quad T_1 + T_2 + T_3 = m_K - 3m_\pi = Q \\
T_i &= k_i : \text{ultra-relativistic} \quad T_i = \frac{k_i^2}{2m_\pi} : \text{nonrelativistic}
\end{aligned}$$

We show the boundary in $(x, y) = \{(T_1 - T_3)/\sqrt{3}, T_2\}$ plane becomes a circle for nonrelativistic case.

$$\mathbf{k}_2 = -(\mathbf{k}_1 + \mathbf{k}_3) \rightarrow k_2^2 = k_1^2 + k_3^2 + 2k_1 k_3 \cos \xi$$

Therefore

$$T_2 = T_1 + T_3 + 2\sqrt{T_1 T_3} \cos \xi \rightarrow \left| \frac{T_2 - T_1 - T_3}{2\sqrt{T_1 T_3}} \right| \leq 1$$

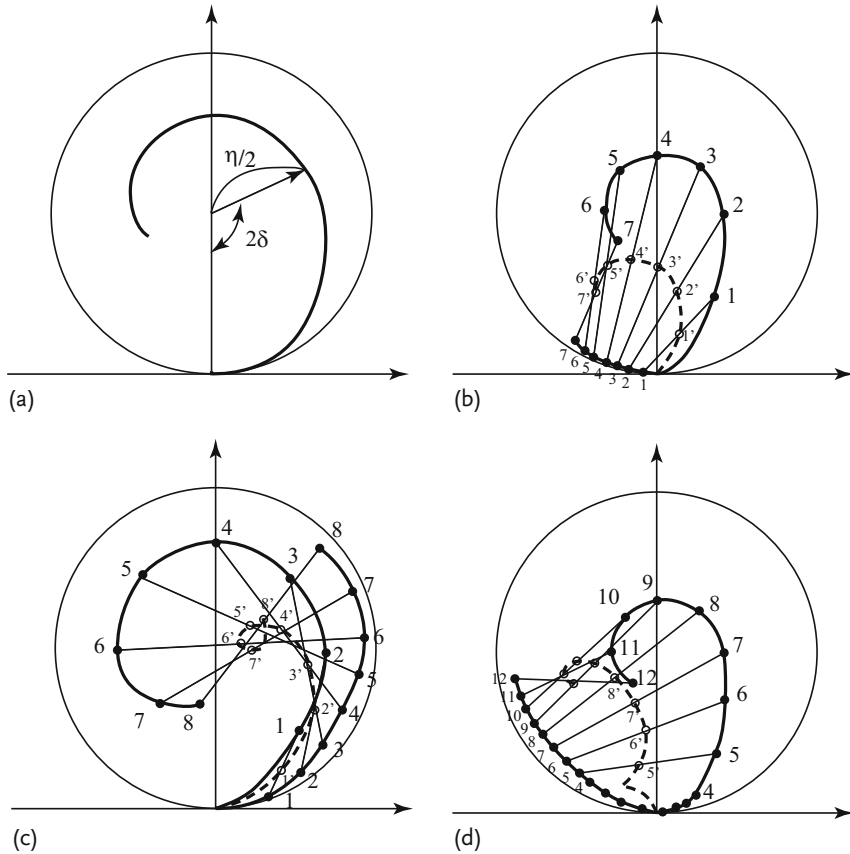


Figure I.2 Resonance behavior: (a) Single resonance amplitude. (b) Repulsive force + resonance, (c) Attractive force + resonance, (d) Repulsive force + resonance. (b–d) Points 1, 2, ... denote two amplitudes at the same energy. Points 1', 2', ... on dashed lines are

midpoints connecting 1-1', 2-2', etc. (d) is similar to (b), but here the resonant amplitude progresses slowly while the non-resonant amplitude swiftly and the relative velocity is reversed at high energy.

The boundary is obtained by putting $|\cos \xi| = 1$

$$T_1^2 + T_2^2 + T_3^2 - 2T_1T_2 - 2T_2T_3 - 2T_3T_1 = 0$$

Then

$$\begin{aligned} & \left(T_2 - \frac{Q}{3}\right)^2 + \left(\frac{T_1 - T_3}{\sqrt{3}}\right)^2 \\ &= \frac{T_1^2 + T_2^2 + T_3^2}{3} + \frac{2}{3}\{T_2(T_2 - Q) - T_1T_3\} + \frac{Q^2}{9} \\ &= \frac{1}{3}(T_1^2 + T_2^2 + T_3^2 - 2T_1T_2 - 2T_2T_3 - 2T_3T_1) + \frac{Q^2}{9} = \frac{Q^2}{9} \end{aligned}$$

Namely, in the (x, y) plane, it is a circle with center at $(x, y) = (0, Q/3)$ and radius $Q/3$.

Ultra-relativistic: From

$$|\cos \xi| = \frac{k_3^2 - k_1^2 - k_2^2}{2k_1 k_2} \leq 1 \rightarrow |k_1 - k_2| \leq k_3 \leq k_1 + k_2$$

$$T_i = k_i \rightarrow y = T_3 \geq |T_1 - T_2| = \sqrt{3}|x|$$

Therefore, y is above lines AC and BC. Next,

$$T_3 = Q - (T_1 + T_2) = Q - (k_1 + k_2) \leq Q - k_3 = Q - T_3, \rightarrow y = T_3 \leq \frac{Q}{2}$$

Namely, y is below the line AB.

Problem 13.12: Setting $\alpha = 2mV(r)$, the Schrödinger equation for radial wave function is given by

$$-u_k'' + \alpha u_k = k^2 u_k, \quad \alpha = 2mV(r) \quad (\text{I.6})$$

$$-v_k'' = k^2 v_k \quad (\text{I.7})$$

From Eq. (I.7), we have

$$v_k = \frac{\sin(kr + \delta_0)}{\sin \delta_0} = \cos kr + \sin kr \cot \delta_0 \simeq 1 + (k \cot \delta_0)r \quad (\text{I.8a})$$

$$\therefore \lim_{k \rightarrow 0} v_k = v_0 = 1 - r/a, \quad \lim_{k \rightarrow 0} k \cot \delta \equiv -1/a \quad (\text{I.8b})$$

where we have used $\lim_{k \rightarrow 0} \delta \propto k$. In Eq. (I.6) a solution u_0 for $k = 0$ satisfies

$$-u_0'' + \alpha u_0 = 0, \quad \therefore \alpha = u_0''/u_0, \quad (\text{I.9})$$

Making (I.6) $\times u_0$ and inserting Eq. (I.9)

$$-u_0 u_k'' + u_0 \alpha u_k = k^2 u_0 u_k, \quad \rightarrow [u_0' u_k - u_0 u_k']' = k^2 u_0 u_k \quad (\text{I.10})$$

Similarly

$$[v_0' v_k - v_0 v_k']' = k^2 v_0 v_k \quad (\text{I.11})$$

Making (I.11)–(I.10)

$$[v_0' v_k - v_0 v_k']' - [u_0' u_k - u_0 u_k']' = k^2 (v_0 v_k - u_0 u_k)$$

Integrating both sides

$$[v_0' v_k - v_0 v_k' - u_0' u_k + u_0 u_k']_0^\infty = \int_0^R (v_0 v_k - u_0 u_k)$$

Calculating l.h.s. explicitly using Eq. (I.8a) and Eq. (I.8b), we have

$$\frac{1}{a} + k \cot \delta_0 = k^2 \int_0^\infty dr (v_0 v_k - u_0 u_k) = k^2 \int_0^R dr (v_0 v_k - u_0 u_k)$$

where we used $u_0 = u_k = 0$ at $r = 0$ and $u_k = v_k$ at $r = \infty$. In the limit $k \rightarrow 0$, $v_k \rightarrow v_0$.

Chapter 14: Quark Model

Problem 14.1:

We show for the case of antisymmetric tensor. $A^{ij} \rightarrow A'^{ij}$

$$A'^{ij} = U^i_a U^j_b A^{ab} = -U^j_b U^i_a A^{ba} = -U^j_a U^i_b A^{ab} = -A'^{ji}$$

In going to the second last equality, we used that suffix for sum can be freely changed. For symmetric tensor S_{ij} , the proof is similar.

Problem 14.2: (I_3, Y) of $A^{12} = \frac{1}{\sqrt{2}}(ud - du)$, $A^{23} = \frac{1}{\sqrt{2}}(ds - sd)$, $A^{31} = \frac{1}{\sqrt{2}}(su - us)$, are $(0, 2/3)$, $(-1/2, -1/3)$, $(1/2, -1/3)$ which are the same quantum numbers as those of $\bar{s}, \bar{u}, \bar{d}$. Therefore, this three members belong to 3^* .

Problem 14.3:

$$\begin{aligned} \psi'_1 &= \varepsilon_{ijk} \xi'^i \xi'^j \xi'^k = \varepsilon_{ijk} U^i_a U^j_b U^k_c \xi^a \xi^b \xi^c \\ &= \det[U^i_a] \varepsilon_{abc} \xi^a \xi^b \xi^c = \varepsilon_{abc} \xi^a \xi^b \xi^c = \psi_1 \end{aligned}$$

Problem 14.4: Inserting quantum numbers (Y, I) = $(1, 1/2)$, $(0, 1)$, $(0, 0)$, $(-1, 1/2)$ of $(N, \Sigma, \Lambda, \Xi)$ into Eq. (14.31), we obtain

$$m_N = a + b + c/2, \quad m_\Sigma = a + 2c, \quad m_\Lambda = a, \quad m_\Xi = a - b + c/2.$$

Eliminating a, b, c leads to Eq. (14.30a). Equality (14.30b) for mesons can be obtained by changing $m_N \rightarrow m_K^2$, $m_\Xi \rightarrow m_K^2$, $m_\Lambda \rightarrow m_\eta^2$, $m_\Sigma \rightarrow m_\pi^2$.

Problem 14.5*: This takes an elaborate algebra.

Tensor Method: Baryons are composed of $q^i q^j q^k$, but as far as $SU(3)$ properties are concerned, it is convenient to use the equality Eq. (14.12) with meson octets substituted with baryon octets.

$$\begin{aligned} B_j^i &= \begin{bmatrix} -\frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda^0}{\sqrt{6}} & \Sigma^+ & p \\ -\Sigma^- & \frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda^0}{\sqrt{6}} & n \\ \Xi^- & -\Xi^0 & -\sqrt{\frac{2}{3}}\Lambda^0 \end{bmatrix} \\ \bar{B}_j^i &= (B_j^i)^* = \begin{bmatrix} -\frac{\bar{\Sigma}^0}{\sqrt{2}} + \frac{\bar{\Lambda}^0}{\sqrt{6}} & -\bar{\Sigma}^+ & \bar{\Xi}^+ \\ \bar{\Sigma}^- & \frac{\bar{\Sigma}^0}{\sqrt{2}} + \frac{\bar{\Lambda}^0}{\sqrt{6}} & -\bar{\Xi}^0 \\ \bar{p} & \bar{n} & -\sqrt{\frac{2}{3}}\bar{\Lambda}^0 \end{bmatrix} \end{aligned} \quad (I.12)$$

As the mass term is expressed as a tensor $\bar{B}_j^i B_b^a$, the solution is obtained if we can make a combination that transforms as a singlet and the 8-th member of an octet. As stated in Sect. 14.1.3, meson singlet and octet can be expressed as

$$\begin{aligned} M_0 &= \frac{u\bar{u} + d\bar{d} + s\bar{s}}{\sqrt{3}} = \frac{\delta_i^j}{\sqrt{3}} q^i \bar{q}_j \equiv S_a^b q^a \bar{q}_b \\ (M_8)_j^i &= q^i \bar{q}_j - \frac{1}{3} \delta_j^i q^a \bar{q}_a = \left(\delta_a^i \delta_j^b - \frac{1}{3} \delta_j^i \delta_a^b \right) q^a \bar{q}_b \equiv (O_j^i)_a^b q^a \bar{q}_b \end{aligned} \quad (I.13)$$

Therefore, to reduce an arbitrary tensor to one that transforms like the mass term, we make contractions of any (i, j) [or any pair of $(i, b), (a, j), (a, b)$] with S_j^i and $(O_3^3)_j^i$ and remove residual suffix $(a \cdots, b \cdots)$ dependence by using $\delta_b^a, \epsilon^{abc}, \epsilon_{abc}$. Consequently, the mass term takes a form

$$\begin{aligned} & A'\{\bar{B}_j^i B_i^j\} + B\{(O_3^3)_j^i \bar{B}_a^j B_i^a\} + C\{(O_3^3)_j^i \bar{B}_i^a B_a^j\} \\ &= A\{\bar{B}_j^i B_i^j\} + (O_3^3)_j^i [F'\{\bar{B}_a^j B_i^a - \bar{B}_i^a B_a^j\} + D'\{\bar{B}_a^j B_i^a + \bar{B}_i^a B_a^j\}] \end{aligned}$$

As the second term in $O_3^3 [\delta_3^3 \delta_b^b/3]$ has the same transformation property as a singlet, we move it to the singlet term and consider only the first term as the octet term, namely we are allowed to change $(O_3^3)_b^a \rightarrow \delta_3^a \delta_b^3$. In summary, the Hamiltonian mass term becomes

$$\begin{aligned} M &= A\{\bar{B}_j^i B_i^j\} + F\{\bar{B}_a^3 B_3^a - \bar{B}_3^a B_a^3\} + D\{\bar{B}_a^3 B_3^a + \bar{B}_3^a B_a^3\} \\ &= A\{\bar{p}p + \bar{n}n + \bar{\Sigma}^- \Sigma^+ + \bar{\Sigma}^0 \Sigma^0 + \bar{\Sigma}^+ \Sigma^- + \bar{\Lambda}^0 \Lambda^0 \\ &\quad + \bar{\Xi}^0 \Xi^0 + \bar{\Xi}^+ \Xi^-\} \\ &\quad + F\{\bar{p}p + \bar{n}n - \bar{\Xi}^0 \Xi^0 - \bar{\Xi}^+ \Xi^-\} + D\{\bar{p}p + \bar{n}n + \bar{\Xi}^0 \Xi^0 \\ &\quad + \bar{\Xi}^+ \Xi^- + \frac{4}{3}\bar{\Lambda}^0 \Lambda^0\} \end{aligned}$$

$$\begin{aligned} \therefore M_N &= A + F + D, \quad M_{\Xi} = A - F + D, \quad M_{\Sigma} = A, \\ M_{\Lambda} &= A + \frac{4}{3}D \end{aligned}$$

Eliminating A, F, D , we finally obtain

$$M_N + M_{\Xi} = \frac{1}{2}(M_{\Sigma} + 3M_{\Lambda})$$

Method Using U Spin: We consider to express the mass in terms of quantum numbers of $S U(3)$. Conventionally used (I_3, Y) are based on subgroup $S U(2)_I \times U(1)_Y$. As $\bar{\psi} \lambda_8 \psi$ has the quantum number Y and does not differentiate Λ from Σ^0 , we consider a different subgroup $S U(2)_U \times U(1)_Q$. $S U(2)_U$ is symmetry of d and s and is differentiated by U_3 . U_3 of (d, s) is $(+1/2, -1/2)$. (d, s) has U -spin $1/2$ and u is a U -spin singlet. The electric charge Q differentiate (d, s) and u . In terms of U_3 and Q , λ_8 is reexpressed as $\lambda_8 = \sqrt{3}Y = (2/\sqrt{3})(U_3 + V_3) = (4/\sqrt{3})(U_3 + 1/2 I_3)$ (see Eq. (G.17c) (G.17d) in Appendix G). Using (Q, U) . The octet members are classified as

$$\begin{aligned} (Q, U) &= (1, 1/2) \rightarrow (p, \Sigma^+), \quad (-1, 1/2) \rightarrow (\Sigma^-, \Xi^-), \\ (0, 1) &\rightarrow (n, \Sigma^0, \Lambda^0), \quad (0, 0) \rightarrow \Xi^0 \end{aligned}$$

Σ^0, Λ^0 are eigenstates of I_3 , but not of U_3 . To obtain $U_3 = 0$ eigenstate, we apply $U_- = (\lambda_6 - \lambda_7)/2$ to $U = U_3 = +1$ eigenstate i.e. $|n\rangle$. The operation is analogous to obtain $I_3 = 0$ state from $I = I_3 = +1$ state by applying $I_- = (\lambda_1 - \lambda_2)/2$.

I_- , U_- operations are to exchange ($u \rightarrow d$, $\bar{d} \rightarrow -\bar{u}$), ($d \rightarrow s$, $\bar{s} \rightarrow -\bar{d}$). Similar to the meson octet for which

$$\begin{aligned} |\eta^0\rangle &= |I : 0, 0\rangle = (u\bar{u} + d\bar{d} - 2s\bar{s})/\sqrt{6} \\ I_-|\pi^+\rangle &= I_-|u\bar{d}\rangle = -(u\bar{u} - d\bar{d}) = \sqrt{2}[-(u\bar{u} + d\bar{d})/\sqrt{2}] \\ &= \sqrt{2}|I : 1, 0\rangle \\ |U : 0, 0\rangle &= (-2u\bar{u} + d\bar{d} + s\bar{s})/\sqrt{6} \\ U_-|U : 1, 1\rangle &= U_-|K^0\rangle = U_-|d\bar{s}\rangle = \sqrt{2}[-(d\bar{d} + s\bar{s})/\sqrt{2}] \\ &= \sqrt{2}|U : 1, 0\rangle \end{aligned}$$

From the above equation, we can obtain

$$\begin{aligned} |U : 1, 0\rangle &= -\frac{1}{2}\pi^0 - \frac{\sqrt{3}}{2}\eta = -\frac{1}{2}|I : 1, 0\rangle - \frac{\sqrt{3}}{2}|I : 0, 0\rangle \\ |U : 0, 0\rangle &= \frac{\sqrt{3}}{2}\pi^0 - \frac{1}{2}\eta = \frac{\sqrt{3}}{2}|I : 1, 0\rangle - \frac{1}{2}|I : 0, 0\rangle \end{aligned}$$

Substituting corresponding baryon octets in the above equation, we have

$$|U : 1, 0\rangle = -\frac{1}{2}|I : 1, 0\rangle - \frac{\sqrt{3}}{2}|I : 0, 0\rangle = -\frac{1}{2}\Sigma^0 - \frac{\sqrt{3}}{2}\Lambda^0$$

As the mass of particles within a multiplet can be expressed as $a + bU_3 + cI_3$, neglecting I_3

$$\begin{aligned} M_n &= \langle U : 1, 1 | M | U : 1, 1 \rangle = a + b \\ \langle U : 1, 0 | M | U : 1, 0 \rangle &= \left\langle -\frac{1}{2}\Sigma^0 - \frac{\sqrt{3}}{2}\Lambda^0 \left| M \right| -\frac{1}{2}\Sigma^0 - \frac{\sqrt{3}}{2}\Lambda^0 \right\rangle \\ &= \frac{1}{4}M_{\Sigma^0} + \frac{3}{4}M_{\Lambda^0} = a \\ M_{\Xi^0} &= \langle U : 1, -1 | M | U : 1, -1 \rangle = a - b \end{aligned}$$

Eliminating a, b , we finally obtain

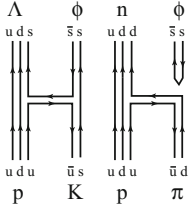
$$M_N + M_{\Xi} = \frac{1}{2}(M_{\Sigma} + 3M_{\Lambda})$$

Problem 14.6: Substituting Eq. (14.33) in Eq. (14.34a),

$$\begin{aligned} \begin{bmatrix} M_8^2 & M_{80}^2 \\ M_{08}^2 & M_0^2 \end{bmatrix} &= \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} M_\phi^2 & 0 \\ 0 & M_\omega^2 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \\ &= \begin{bmatrix} c^2 M_\phi^2 + s^2 M_\omega^2 & cs(-M_\phi^2 + M_\omega^2) \\ cs(-M_\phi^2 + M_\omega^2) & s^2 M_\phi^2 + c^2 M_\omega^2 \end{bmatrix} \end{aligned}$$

where $c = \cos \theta$, $s = \sin \theta$. On the other hand, from Eq. (14.35)

$$\begin{aligned} M_\phi^2 - M_8^2 &= M_0^2 - M_\omega^2 = s^2 M_\phi^2 + c^2 M_\omega^2 - M_\omega^2 = s^2(M_\phi^2 - M_\omega^2) \\ \therefore \sin^2 \theta &= \frac{M_\phi^2 - M_8^2}{M_\phi^2 - M_\omega^2} \end{aligned}$$

Figure I.3 Left: $K^- p \rightarrow \phi \Lambda$, Right: $\pi^- p \rightarrow \phi n$.

Problem 14.7: Using the quark model, $K^- = s\bar{u}$, $p = uud$, $\phi = s\bar{s}$, $\Lambda = uds$, $\pi^- = d\bar{u}$, $n = udd$, reactions can be described as in Fig. I.3. Namely in the process $K^- p \rightarrow \phi \Lambda$, quark lines are connected but in $\pi^- p \rightarrow \phi n$ are not.

Problem 14.8: The decay process of a vector particle into a lepton pair is one where a pair of $q\bar{q}$ annihilates into a photon, subsequently decaying to a lepton pair. This means the decay amplitude is proportional to the quark charge. Since pair annihilations $u\bar{u}$ and $d\bar{d}$ cannot be distinguished, the two amplitudes have to be added in the decay process of ρ , $\omega = (u\bar{u} \pm d\bar{d})/\sqrt{2}$ and $\phi = s\bar{s}$. They are proportional to

$$\left| \frac{1}{\sqrt{2}} \left\{ \frac{2}{3} - \left(-\frac{1}{3} \right) \right\} \right|^2 : \left| \frac{1}{\sqrt{2}} \left\{ \frac{2}{3} + \left(-\frac{1}{3} \right) \right\} \right|^2 : \left| -\frac{1}{3} \right|^2 = \frac{1}{2} : \frac{1}{18} : \frac{1}{9}$$

Problem 14.9: The decay rate Eq. (14.40) can be derived as follows. We consider a vector $V(J^P = 1^-)$ or a scalar $S(0^-)$ meson which is a composite of quarks (where $m_q \gg m_l$ is assumed) in the relative orbital angular momentum S-state. The cross section $q\bar{q} \rightarrow l\bar{l}$ is given by Eq. (C.17).

$$\sigma_{\text{TOT}} = \frac{4\pi\alpha^2}{3s} Q^2 \frac{\beta_f}{\beta_i} \left[1 + \frac{2 - \beta_i^2 - \beta_f^2}{2} + \frac{(1 - \beta_i^2)(1 - \beta_f^2)}{4} \right]$$

where

$$\beta_i = \sqrt{s - 4m_q^2}/\sqrt{s}, \quad \beta_f = \sqrt{s - 4m_l^2}/\sqrt{s}, \quad s = 4E_{\text{CM}}^2$$

and Q is the electric charge of the quark q . The cross section is a sum of spin triplet cross section σ_t and spin singlet σ_s , i.e. $\sigma = \frac{3}{4}\sigma_t + \frac{1}{4}\sigma_s$. For the vector meson decay $V \rightarrow l\bar{l}$, σ_s does not contribute. Hence the decay rate Γ in the limit $\beta_f \rightarrow 1$ ($s = 4m_q^2 \gg m_l^2$) is given by

$$\Gamma = (\text{flux}) \times \sigma_t = 2\beta_q |\psi(0)|^2 \frac{4}{3} \sigma = \frac{16\pi\alpha^2}{3} Q^2 \frac{|\psi(0)|^2}{m_V^2}$$

where we have put $s = m_V^2$, $\beta_i \rightarrow 0$ (nonrelativistic limit). $\psi(0)$ is the S-state wave function at the origin whose shape depends on the potential acting on the $q\bar{q}$ pair. it is smaller than Eq. (14.40) by a factor three. This is because the quark has three colors which shall be explained soon. Three identical diagrams contribute to the total decay amplitude, therefore the decay rate has to be multiplied by 9 and has to be divided by 3 because flux of each color is 1/3 on the average.

Problem 14.10: Performing cyclic replacement starting from

$$(ud - du)u(\uparrow\downarrow - \downarrow\uparrow) \uparrow = u \uparrow d \downarrow u \uparrow + d \downarrow u \uparrow u \uparrow - u \downarrow d \uparrow u \uparrow \\ - d \uparrow u \downarrow u \uparrow$$

One obtains:

$$\begin{aligned} & u \uparrow d \downarrow u \uparrow + d \downarrow u \uparrow u \uparrow - u \downarrow d \uparrow u \uparrow - d \uparrow u \downarrow u \uparrow \\ & + u \uparrow u \uparrow d \downarrow + u \uparrow d \downarrow u \uparrow - u \uparrow u \downarrow d \uparrow - u \uparrow d \uparrow u \downarrow \\ & + d \downarrow u \uparrow u \uparrow + u \uparrow u \uparrow d \downarrow - d \uparrow u \uparrow u \downarrow - u \downarrow u \uparrow d \uparrow \\ & = 2u \uparrow u \uparrow d \downarrow - u \uparrow u \downarrow d \uparrow - u \downarrow u \uparrow d \uparrow \\ & + 2u \uparrow d \downarrow u \uparrow - u \downarrow d \uparrow u \uparrow - u \uparrow d \uparrow d \downarrow \\ & + 2d \downarrow u \uparrow u \uparrow - d \uparrow u \uparrow u \downarrow - d \uparrow u \downarrow u \uparrow = \text{Eq. (14.52)} \end{aligned}$$

Problem 14.12: If one makes wave function of the neutron, one sees it has permutation symmetry $u \leftrightarrow d$. Therefore, the magnetic moment of the neutron μ_n can be obtained from μ_p by exchange of $u \leftrightarrow d$. Magnetic moments of other baryons having two identical quarks out of three can be obtained in a similar manner.

$$\begin{aligned} \mu_n &= \mu_p(u \leftrightarrow d), \quad \mu_{\Sigma^+} = \mu_p(d \rightarrow s), \quad \mu_{\Sigma^-} = \mu_n(u \rightarrow s) \\ \mu_{\Xi^0} &= \mu_n(d \rightarrow s), \quad \mu_{\Xi^-} = \mu_p(u \rightarrow s) \end{aligned}$$

As Ω^- is made of sss and has spin $s = 3/2$, its wave function is given by $s \uparrow s \uparrow s \uparrow$. Therefore the magnetic moment is given by $\mu_{\Omega^-} = 3\mu_s$.

The quark components of Σ^0 , Λ^0 are uds and the same for both, but they have isospin $I = 1, 0$ hence the wave functions are symmetric (antisymmetric). Then, by symmetry, the spin wave functions of $\Sigma^0(\Lambda^0)$ are also symmetric (antisymmetric). The structure of the spin wave function of Σ^0 then is the same as that of the proton in Eq. (14.52) and that of Λ^0 is antisymmetric and has the same structure as Problem 14.10.

$$\begin{aligned} |\Sigma^0\rangle &= uds \frac{1}{\sqrt{6}} \{2 \uparrow\uparrow\downarrow - (\uparrow\downarrow + \downarrow\uparrow) \uparrow\}, \\ |\Lambda^0\rangle &= uds \frac{1}{\sqrt{2}} \{(\uparrow\downarrow - \downarrow\uparrow) \uparrow\} \end{aligned}$$

As a preparation for calculating $\mu_{\Sigma\Lambda}$, let us use a method different given in the text. The Hamiltonian for the magnetic moment in the quark model is given by

$$H = \mu_u \sigma_z(u) + \mu_d \sigma_z(d) + \mu_s \sigma_z(s)$$

$\sigma_z(u)$ means that it acts only on u quark. Since $\sigma_z \downarrow = -\downarrow$, we have

$$\begin{aligned} H|\Sigma^0\rangle &= \frac{uds}{\sqrt{6}} [\mu_u \{2 \uparrow\uparrow\downarrow - (\uparrow\downarrow - \downarrow\uparrow) \uparrow\} \\ &+ \mu_d \{2 \uparrow\uparrow\downarrow - (-\uparrow\downarrow + \downarrow\uparrow) \uparrow\} \\ &+ \mu_s \{-2 \uparrow\uparrow\downarrow - (\uparrow\downarrow + \downarrow\uparrow) \uparrow\}] \end{aligned} \quad (\text{I.14})$$

By multiplying $\langle \Sigma^0 |$ from left and calculating the expectation values,

$$\mu_{\Sigma^0} = \frac{4}{6}\mu_u + \frac{4}{6}\mu_d + \mu_s \left(-\frac{4}{6} + \frac{1}{6} + \frac{1}{6} \right) = \frac{2}{3}(\mu_u + \mu_d) - \frac{1}{3}\mu_d$$

Similarly, multiplying $\langle \Lambda |$ by

$$\begin{aligned} H|\Lambda^0\rangle &= \frac{uds}{\sqrt{2}} \{ \mu_u(\uparrow\downarrow + \downarrow\uparrow) \uparrow + \mu_d(-\uparrow\downarrow - \downarrow\uparrow) \uparrow \\ &\quad + \mu_s(\uparrow\downarrow - \downarrow\uparrow) \uparrow \} \end{aligned}$$

we obtain $\mu_\Lambda = \mu_s$.

Multiplying Eq. (I.14) by $\langle \Lambda^0 |$, we have

$$\mu_{\Sigma\Lambda} = -\frac{1}{\sqrt{3}}\mu_u + \frac{1}{\sqrt{3}}\mu_d = \frac{-\mu_u + \mu_d}{\sqrt{3}}$$

Problem 14.13: The gluon has colors (8 species), hence $e\bar{e}$ cannot be converted to a gluon. As the gluon has the same C-parity (-1) as the photon, J/ψ having the same quantum number as the photon cannot be converted to 2 gluons. Hence, three gluons have to be exchanged.

Problem 14.14:

$$\begin{aligned} \mathbf{s}_i \cdot \mathbf{s}_j &= \frac{1}{2} \{ (\mathbf{s}_i + \mathbf{s}_j)^2 - \mathbf{s}_i^2 - \mathbf{s}_j^2 \} \\ &= \frac{1}{2} \{ S(S+1) - s_i(s_i+1) - s_j(s_j+1) \} \\ &= \begin{cases} \frac{1}{2} \{ 1(1+1) - 2 \times \frac{1}{2} (\frac{1}{2} + 1) \} = \frac{1}{4} & S = 1 \\ \frac{1}{2} \{ 0(0+1) - 2 \times \frac{1}{2} (\frac{1}{2} + 1) \} = -\frac{3}{4} & S = 0 \end{cases} \\ \sum_{i,j} (\mathbf{s}_i \cdot \mathbf{s}_j) &= \frac{1}{2} \{ (\mathbf{s}_1 + \mathbf{s}_2 + \mathbf{s}_3)^2 - \mathbf{s}_1^2 - \mathbf{s}_2^2 - \mathbf{s}_3^2 \} \\ &= \frac{1}{2} \left\{ S(S+1) - 3 \times \frac{1}{2} \left(\frac{1}{2} + 1 \right) \right\} \\ &= \begin{cases} \frac{1}{2} \left\{ \frac{3}{2} \left(\frac{3}{2} + 1 \right) - 3 \times \frac{1}{2} \left(\frac{1}{2} + 1 \right) \right\} = \frac{3}{4} & S = \frac{3}{2} \\ \frac{1}{2} \left\{ \frac{1}{2} \left(\frac{1}{2} + 1 \right) - 3 \times \frac{1}{2} \left(\frac{1}{2} + 1 \right) \right\} = -\frac{3}{4} & S = \frac{1}{2} \end{cases} \end{aligned}$$

Problem 14.15:

$$M = \sum_{i=1}^3 m_i + B \sum_{i,j} \frac{\mathbf{s}_i \cdot \mathbf{s}_j}{m_i m_j}$$

As Δ consists of u, d and has spin $s = 3/2$, the first term gives $3m_u$. Using Eq. (14.98b) the second term becomes $\frac{3}{4} \frac{B}{m_u^2}$. Therefore,

$$m(\Delta) = 3m_u + B \left(\frac{3}{4} \frac{1}{m_u^2} \right)$$

The equation for N can be obtained from the above formula by changing $\sum(\mathbf{s}_i \cdot \mathbf{s}_j) \rightarrow -3/4$ (Eq. (14.98b)), $m(N) = 3m_u + B \left(-\frac{3}{4} \frac{1}{m_u^2} \right)$

As Ξ^* consists of u or $d + 2s$ and has spin $s = 3/2$,

$$M = m_u + 2m_s + B \left\{ \frac{\mathbf{s}_s \cdot \mathbf{s}_s}{m_s^2} + \frac{2(\mathbf{s}_u \cdot \mathbf{s}_s)}{m_u m_s} \right\}.$$

Any quark pair makes spin $s = 1$, $(\mathbf{s}_s \cdot \mathbf{s}_s) = (\mathbf{s}_u \cdot \mathbf{s}_s) = 1/4$. Therefore,

$$m(\Xi^*) = m_u + 2m_s + B \left(\frac{1}{4m_s^2} + \frac{1}{2m_u m_s} \right)$$

Ξ has spin $s = 1/2$, but as the pair ss is symmetric, so is the spin wave function hence its spin has to be $s = 1$.

$$\begin{aligned} 2\mathbf{s}_u \cdot \mathbf{s}_s &= \frac{1}{2} \{ (\mathbf{s}_u + \mathbf{s}_s + \mathbf{s}_s)^2 - \mathbf{s}_u^2 - (\mathbf{s}_s + \mathbf{s}_s)^2 \} \\ &= \frac{1}{2} \left\{ \frac{1}{2} \left(\frac{1}{2} + 1 \right) - \frac{3}{4} - 1(1 + 1) \right\} = -1 \\ \therefore m(\Xi) &= m_u + 2m_s + B \left(-\frac{1}{m_u m_s} + \frac{1}{4m_s^2} \right) \end{aligned}$$

As Ω is made of sss , the mass can be obtained from that of $uuu = \Delta^{++}$ by replacing $u \rightarrow s$.

$$\therefore m(\Omega) = 3m_s + B \left(\frac{3}{4} \frac{1}{m_s^2} \right)$$

Problem 14.16:

$$M = m_1 + m_2 + \frac{C}{m_1 m_2} (\mathbf{s}_1 \cdot \mathbf{s}_2), \quad (\mathbf{s}_1 \cdot \mathbf{s}_2) = -3/4 (S = 0), \quad 1/4 (S = 1)$$

$$\phi = s\bar{s}, \quad S = 1 \longrightarrow m_\phi = 2m_s + C \frac{1}{4m_s^2}$$

$$K^* = s\bar{u}, \quad S = 1 \longrightarrow m_{K^*} = m_u + m_s + C \frac{1}{4m_u m_s}$$

$$\omega = (u\bar{u} + d\bar{d})/\sqrt{2}, \quad S = 1 \longrightarrow m_\omega = 2m_u + C \frac{1}{4m_u^2}$$

$$\rho, \quad S = 1 \longrightarrow m_\rho = m_\omega$$

As $\eta = (u\bar{u} + d\bar{d} - 2s\bar{s})/\sqrt{6}$, it has mass m_π with probability $1/3$ and has mass of a particle with $s\bar{s}$, $S = 0$. Since $\mathbf{s} \cdot \mathbf{s} = -3/4$

$$\begin{aligned} m_\eta &= \frac{1}{3} \left(2m_u - C \frac{3}{4m_u^2} \right) + \frac{2}{3} \left(2m_s - C \frac{3}{4m_s^2} \right) \\ &= \frac{2}{3} (m_u + 2m_s) - C \frac{3}{4} \left(\frac{1}{m_u^2} + \frac{1}{m_s^2} \right) \end{aligned}$$

$$K = u\bar{s}, \quad S = 0 \longrightarrow m_K = m_u + m_s - C \frac{3}{4m_u m_s}$$

$$\pi = ud, \quad S = 0 \longrightarrow m_\pi = 2m_u - C \frac{3}{4m_u^2}$$

Chapter 15: Weak Interaction

Problem 15.1: From Eq. (15.23), the most general Lagrangian for the beta decay can be expressed as

$$\begin{aligned}\mathcal{L} &= G_\beta \sum (\bar{p}\Gamma'_i n)[C_i(\bar{e}\Gamma_i \nu) + C'_i(\bar{e}\Gamma_i \gamma^5 \nu)] + h.c. \\ \Gamma_i &= 1(S), \gamma^\mu(V), \sigma^{\mu\nu}(T), \gamma^\mu \gamma^5(A), i\gamma^5(P) \\ \Gamma'_i &= 1(S), \gamma_\mu(V), \sigma_{\mu\nu}(T), \gamma_\mu \gamma^5(A), i\gamma^5(P) \\ h.c. &= G_\beta \sum (\bar{n}\gamma^0(\Gamma'_i)^\dagger \gamma^0 p)[C_i^*(\bar{\nu}\gamma^0 \Gamma_i^\dagger \gamma^0 e) \\ &\quad + (C'_i)^*(\bar{\nu}\gamma^0 \gamma^5 \Gamma_i^\dagger \gamma^0 e)]\end{aligned}$$

Using relations $\gamma^0 \Gamma_i^\dagger \gamma^0 = \Gamma_i$, $\gamma^0 \gamma^5 \gamma^0 = -\gamma^5$.

The hermitian conjugate becomes

$$h.c. = G_\beta \sum (\bar{n}\Gamma'_i p)[C_i^*(\bar{\nu}\Gamma_i e) - (C'_i)^*(\bar{\nu}\gamma^5 \Gamma_i e)] \quad (\text{h.c. of the first term})$$

P: P transformation changes (see Appendix F)

$$\begin{aligned}\bar{\psi}_1 \Gamma_i \psi_2 &\rightarrow \bar{\psi}_1 \gamma^0 \Gamma_i \gamma^0 \psi_2 = \varepsilon_{Pi} \bar{\psi}_1 \Gamma'_i \psi_2, \\ \bar{\psi}_1 \Gamma'_i \psi_2 &\rightarrow \bar{\psi}_1 \gamma^0 \Gamma'_i \gamma^0 \psi_2 = \varepsilon_{Pi} \bar{\psi}_1 \Gamma_i \psi_2 \\ \bar{\psi}_1 \Gamma_i \gamma^5 \psi_2 &\rightarrow \bar{\psi}_1 \gamma^0 \Gamma_i \gamma^0 \gamma^5 \gamma^0 \psi_2 = -\varepsilon_{Pi} \bar{\psi}_1 \Gamma'_i \gamma^5 \psi_2,\end{aligned}$$

ε_{Pi} denotes sign change depending on transformation property of Γ_i . For instance,

$$\varepsilon_{Pi} = +1, \quad \text{for } i = S, V, T, \quad -1 \quad \text{for } A, P.$$

Since $\varepsilon_{Pi}(\bar{p}\Gamma_i n)\varepsilon_{Pi}(\bar{e}\Gamma'_i \nu) = (\bar{p}\Gamma'_i n)(\bar{e}\Gamma_i \nu)$, we have

$$\begin{aligned}(\bar{p}\Gamma'_i n)[C_i(\bar{e}\Gamma_i \nu) + C'_i(\bar{e}\Gamma_i \gamma^5 \nu)] \\ \xrightarrow{P} (\varepsilon_{Pi} \bar{p}\Gamma_i n)[C_i(\varepsilon_{Pi} \bar{e}\Gamma'_i \nu) - C'_i(\varepsilon_{Pi} \bar{e}\Gamma'_i \gamma^5 \nu)] \\ = (\bar{p}\Gamma'_i n)[C_i(\bar{e}\Gamma_i \nu) - C'_i(\bar{e}\Gamma_i \gamma^5 \nu)]\end{aligned}$$

Namely, the sign change ε_{Pi} due to different Γ_i and exchange operation $\Gamma_i \leftrightarrow \Gamma'_i$ do not make any difference in the final Lagrangian. Only the sign change by $\gamma^5 \rightarrow \gamma^0 \gamma^5 \gamma^0 = -\gamma^5$ in the second term in $[\dots]$ transforms $C'_i \rightarrow -C'_i$. Therefore, we have shown that P invariance means $C'_i = 0$.

T: T transformation changes

$$\begin{aligned}C_i \bar{\psi}_1 \Gamma_i \psi_2 &\rightarrow C_i^* \bar{\psi}_1 B^{-1} \Gamma_i^* B \psi_2 = C_i^* \varepsilon_{Ti} \bar{\psi}_1 \Gamma'_i \psi_2 \\ \therefore (\bar{p}\Gamma'_i n)[C_i(\bar{e}\Gamma_i \nu) + C'_i(\bar{e}\Gamma_i \gamma^5 \nu)] &\xrightarrow{T} (\bar{p}B^{-1}(\Gamma'_i)^* B n) \\ &\quad \times [C_i^*(\bar{e}B^{-1} \Gamma_i^* B \nu) + (C'_i)^*(\bar{e}B^{-1} \Gamma_i^* B B^{-1}(\gamma^5)^* B \nu)]\end{aligned}$$

Inserting $B^{-1}\gamma^\mu B = \gamma_\mu$, $B^{-1}\gamma^5 B = \gamma^5$, the above equation becomes

$$\begin{aligned} &= (\varepsilon_{Ti} \bar{p} \Gamma_i n) [C_i^* (\varepsilon_{Ti} \bar{e} \Gamma_i' \nu) + (C_i')^* (\varepsilon_{Ti} \bar{e} \Gamma_i' \gamma^5 \nu)] \\ &= (\bar{p} \Gamma_i' n) [C_i^* (\bar{e} \Gamma_i \nu) + (C_i')^* (\bar{e} \Gamma_i \gamma^5 \nu)] \end{aligned}$$

This shows that T invariance means $C_i = C_i^*$, $C_i' = (C_i')^*$.

C: C transformation changes

$$\begin{aligned} C_i \bar{\psi}_1 \Gamma_i \psi_2 &\xrightarrow{C} -C_i \psi_1^T C^{-1} \Gamma_i C \bar{\psi}_2^T = -C_i \varepsilon_{Ci} \psi_1^T \Gamma_i^T \bar{\psi}_2^T \\ &= C_i \varepsilon_{Ci} \bar{\psi}_2 \Gamma_i \psi_1 \end{aligned}$$

$$\begin{aligned} \therefore (\bar{p} \Gamma_i' n) [C_i (\bar{e} \Gamma_i \nu) + C_i (\bar{e} \Gamma_i \gamma^5 \nu)] &\xrightarrow{C} (-p C^{-1} \Gamma_i' C \bar{n}) \\ &\times [-C_i (e C^{-1} \Gamma_i C \bar{\nu}) - C_i' (e C^{-1} \Gamma_i C C^{-1} \gamma^5 C \bar{\nu})] \end{aligned}$$

Considering $C = i\gamma^2\gamma^0$, $C^{-1}\Gamma_i C C^{-1}\gamma^5 C = \varepsilon_{Ci} \Gamma_i^T (\gamma^5)^T$.

The above equation becomes

$$\begin{aligned} &= (\varepsilon_{Ci} \bar{n} \Gamma_i' p) [C_i (\varepsilon_{Ci} \bar{\nu} \Gamma_i e) + C_i' (\varepsilon_{Ci} \bar{\nu} \gamma^5 \Gamma_i e)] \\ &= (\bar{n} \Gamma_i' p) [C_i (\bar{\nu} \Gamma_i e) + C_i' (\bar{\nu} \gamma^5 \Gamma_i e)] \end{aligned}$$

Comparing with Eq. (h.c. of the first term), it means $C_i = C_i^*$, $C_i' = -(C_i')^*$

Problem 15.2: Inserting Eq. (15.25c) into Eq. (15.18) in the text

$$\begin{aligned} d\Gamma &= G_\beta^2 (C_F^2 |\langle 1 \rangle|^2 + C_{GT}^2 |\langle \sigma \rangle|^2) (1 + \nu \cos \theta) \\ &\times \frac{1}{(2\pi)^5} p E (\Delta - E)^2 dE d\Omega_e d\Omega_\nu \end{aligned}$$

using

$$\int (1 + \nu \cos \theta) d\Omega_e d\Omega_\nu = 16\pi^2, \quad \int_0^\Delta E^2 (\Delta - E)^2 dE = \frac{\Delta^5}{30}$$

we obtain

$$\Gamma = \int d\Gamma = \frac{G_\beta^2 \Delta^5}{60\pi^3} (C_F^2 |\langle 1 \rangle|^2 + C_{GT}^2 |\langle \sigma \rangle|^2)$$

Problem 15.3: Only variables that are not integrated in the integrand are $\Delta = p_1 - p_4 = p_2 + p_3$. Lorentz invariance requires the integral is a tensor of rank 2 and can be put in a form

$$S^{\mu\nu} = A g^{\mu\nu} + B \Delta^\mu \Delta^\nu$$

In order to obtain coefficients A, B, we calculate $C = S^{\mu\nu} g_{\mu\nu}$, $D = S^{\mu\nu} \Delta_\mu \Delta_\nu$

$$\begin{aligned} p_2^\mu p_3^\nu g_{\mu\nu} &= p_2 \cdot p_3 = \frac{1}{2} (\Delta^2 - p_2^2 - p_3^2) = \frac{\Delta^2}{2} \\ p_2^\mu p_3^\nu \Delta_\mu \Delta_\nu &= (p_2 \cdot \Delta)(p_3 \cdot \Delta) = (p_2 \cdot p_3)^2 = \frac{\Delta^2}{4} \end{aligned}$$

In CM of the particles 2 and 3, $(E_2 + E_3) = 2p$. Then $\int \frac{d^3 p_2 d^3 p_3}{4 E_2 E_3} \delta^4(\Delta - p_2 - p_3) = \frac{\pi}{2}$

$$(A g^{\mu\nu} + B \Delta^\mu \Delta^\nu) g_{\mu\nu} = 4A + B \Delta^2 = \frac{\pi \Delta^2}{4}$$

$$(A g^{\mu\nu} + B \Delta^\mu \Delta^\nu) \Delta_\mu \Delta_\nu = (A + B \Delta^2) \Delta^2 = \frac{\pi \Delta^4}{8}$$

$$\therefore A = \frac{\pi \Delta^2}{24}, \quad B = \frac{\pi}{12}$$

Problem 15.4: Denoting energy momentum of the muon in its rest frame as $(m_\mu, 0)$ and that of the electron as (E, p) , Eq. (15.82) is written as

$$d\Gamma = \frac{G_\mu^2 (p_{1\mu} S^{\mu\nu} p_{4\nu})}{2 m_\mu \pi^5} \frac{d^3 p_4}{E_4},$$

$$p_{1\mu} S^{\mu\nu} p_{4\nu} = \frac{\pi}{24} \{ (p_1 \cdot p_4) \Delta^2 + 2(p_1 \cdot \Delta)(p_4 \cdot \Delta) \} \quad (\text{I.15})$$

In the rest frame of the muon,

$$p_1 \cdot p_4 = m_\mu E, \quad \Delta^2 = (m_\mu - E)^2 - p^2 = m_\mu^2 - 2m_\mu E + m_e^2,$$

$$p_1 \cdot \Delta = p_1^2 - p_1 \cdot p_4 = m_\mu^2 - m_\mu E,$$

$$p_4 \cdot \Delta = p_1 \cdot p_4 - p_4^2 = m_\mu E - m_e^2$$

Substituting the above relations in Eq. (I.15)

$$d\Gamma = \frac{G_\mu^2}{48\pi^4} \{ 3(m_\mu^2 + m_e^2)E - 4m_\mu E^2 - 2m_\mu m_e^2 \} \frac{d^3 p}{E} \quad (\text{I.16})$$

Putting $m_e^2 = 0$, $p = E = \frac{m_\mu}{2} x$, $\frac{d^3 p}{E} = 4\pi p dp$, we finally obtain

$$d\Gamma = \frac{G_\mu^2 m_\mu^5}{96\pi^3} x^2 (3 - 2x) dx$$

Problem 15.5: We insert the following relations in Eq. (I.16).

$$p = \frac{m_\mu}{2} x, \quad E = \frac{m_\mu}{2} \sqrt{x^2 + 4\omega}, \quad \omega = \left(\frac{m_e}{m_\mu} \right)^2$$

$$x_{\max} = 1 - \omega$$

$$d\Gamma = \frac{G_\mu^2 m_\mu^5}{96\pi^3} \{ 3(1 + \omega) \sqrt{x^2 + 4\omega} - 2(x^2 + 4\omega) - 4\omega \} \frac{x^2 dx}{\sqrt{x^2 + 4\omega}}$$

By making use of equalities;

$$\int \frac{dx}{\sqrt{x^2 + c}} = \log |x + \sqrt{x^2 + c}|$$

$$\int \sqrt{x^2 + c} dx = \frac{1}{2} \left\{ x \sqrt{x^2 + c} + c \log |x + \sqrt{x^2 + c}| \right\}$$

$$\int x^2 \sqrt{x^2 + c} dx = \frac{1}{8} \left\{ 2x(x^2 + C)^{\frac{3}{2}} - cx \sqrt{x^2 + c} - c^2 \log |x + \sqrt{x^2 + c}| \right\}$$

we obtain

$$\Gamma = \frac{G_\mu^2 m_\mu^5}{192\pi^3} f\left(\frac{m_e^2}{m_\mu^2}\right), \quad f(\omega) = 1 - 8\omega + 8\omega^3 - \omega^4 - 12\omega^2 \log \omega$$

Problem 15.7: We start from Eq. (15.114a)

$$\mathcal{M}(\pi^+ \rightarrow \pi^0 e^+ \nu) = \frac{G_\beta}{\sqrt{2}} \sqrt{2}(k'^\mu + k^\mu) \bar{u}(p_e) \gamma_\mu (1 - \gamma^5) v(p_\nu)$$

Inserting $k'^\mu + k^\mu = (m_{\pi^+} + m_{\pi^0})\delta_{\mu 0}$

$$\begin{aligned} & |(k'^\mu + k^\mu) \bar{u}(p_e) \gamma_\mu (1 - \gamma^5) v(p_\nu)|^2 \\ &= (m_{\pi^+} + m_{\pi^0})^2 \delta_{\mu 0} \delta_{\nu 0} \text{Tr}[(\not{p}_e + m_e) \gamma^\mu (1 - \gamma^5) \not{p}_\nu \gamma^\nu (1 - \gamma^5)] \\ &= 2(m_{\pi^+} + m_{\pi^0})^2 \delta_{\mu 0} \delta_{\nu 0} \text{Tr}[\not{p}_e \gamma^\mu \not{p}_\nu \gamma^\nu] \\ &= 8(m_{\pi^+} + m_{\pi^0})^2 [2E_e E_\nu - (p_e \cdot p_\nu)] \\ &= 8(m_{\pi^+} + m_{\pi^0})^2 E_e E_\nu (1 + v_e \cos \theta) \end{aligned}$$

$$\begin{aligned} \Gamma &= (2\pi)^4 \delta^4(\Delta - p_e - p_\nu) \frac{1}{2m_{\pi^+}} \\ &\times \sum_{\text{spin}} |\mathcal{M}|^2 \frac{d^3 p_\pi}{(2\pi)^3 2\omega} \frac{d^3 p_e}{(2\pi)^3 2E_e} \frac{d^3 p_\nu}{(2\pi)^3 2E_\nu} \\ &= \frac{G_\beta^2}{(2\pi)^5} \frac{8(m_{\pi^+} + m_{\pi^0})^2}{16m_{\pi^+} m_{\pi^0}} \int \delta(\Delta - E_e - E_\nu) E_e E_\nu (1 + v_e \cos \theta) \\ &\times \frac{p_e^2 d p_e d\Omega_e}{E_e} \frac{p_\nu^2 d p_\nu d\Omega_\nu}{E_\nu} \\ &= \frac{G_\beta^2}{\pi^3} \frac{(m_{\pi^+} + m_{\pi^0})^2}{4m_{\pi^+} m_{\pi^0}} \int p_e E_e (\Delta - E_e)^2 dE_e \\ &= \frac{G_\beta^2 \Delta^5}{30\pi^3} \frac{(1 - \Delta/(2m_{\pi^+}))^2}{1 - \Delta/m_{\pi^+}} R(m_e/\Delta) \end{aligned}$$

$$\begin{aligned} R(x) &= 30 \int_x^1 d\gamma \gamma (1 - \gamma)^2 \sqrt{\gamma^2 - x^2} \\ &= \sqrt{1 - x^2} \left(1 - \frac{9}{2}x^2 - 4x^4\right) + \frac{15}{2}x^4 \ln \frac{1 + \sqrt{1 - x^2}}{x} \end{aligned}$$

The $\cos \theta$ term vanishes when integrated over solid angle. To calculate $R(x)$, use

$$\begin{aligned} \int d\gamma \gamma \sqrt{\gamma^2 - 1} &= (1/3)(\gamma^2 - 1)^{3/2} \\ \int d\gamma \gamma^2 \sqrt{\gamma^2 - 1} &= \frac{1}{8} \left[2\gamma(\gamma^2 - 1)^{3/2} + \gamma(\gamma^2 - 1)^{1/2} - \ln \left| \gamma + \sqrt{\gamma^2 - 1} \right| \right] \\ \int d\gamma \gamma^3 \sqrt{\gamma^2 - 1} &= (1/5)(\gamma^2 - 1)^{5/2} + (1/3)(\gamma^2 - 1)^{3/2} \end{aligned}$$

Problem 15.8: As π is a boson having $J^P = 0^-$, its total wave function is symmetric. In order to have $J = 0$ with two pions, the angular orbital momentum has to be zero, making the spatial part of the wave function symmetric. This means isospin part of the wave function has to be symmetric, too. As $I = 1$ wave function is antisymmetric, $\pi^+\pi^0$ can be only in $I = 2$ state. Therefore, $\Delta I = 1/2$ rule forbids transition $K^+ \rightarrow \pi^+\pi^0$.

Problem 15.9:

1. From the previous problem, K_S, K_L can decay only to $I = 0$. $I = 0$ wave function can be constructed using the Clebsch–Gordan coefficients as

$$|I : 0, 0\rangle = \frac{1}{\sqrt{3}}(\pi^+\pi^- - \pi^0\pi^0 + \pi^-\pi^+)$$

and decay rates to $\pi^+\pi^-$, $\pi^0\pi^0$, $\pi^-\pi^+$ are equal = $1/3$. But as observationally $\pi^+\pi^-$ and $\pi^-\pi^+$ cannot be distinguished, the decay rate to $\pi^+\pi^-$ is twice as that to $\pi^0\pi^0$. Note that the decay $K_L \rightarrow 2\pi$ does not occur if CP is invariant.

2. We assume that the spurion s has $I = 1/2$, $I_3 = -1/2$. Combinations with the spurion are

$$\begin{aligned} |s\Xi^- \rangle &= |1, -1\rangle \\ |s\Xi^0 \rangle &= \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \left| \frac{1}{2}, +\frac{1}{2} \right\rangle = \frac{1}{\sqrt{2}}(|1, 0\rangle + |0, 0\rangle) \\ \langle s\Xi^- | T | \Lambda\pi^- \rangle &= \langle 1, -1 | T | 1, -1 \rangle = T_1 \\ \langle s\Xi^0 | T | \Lambda\pi^0 \rangle &= \frac{1}{\sqrt{2}}(\langle 1, 0 | + \langle 0, 0 |) T | 1, 0 \rangle = \frac{1}{\sqrt{2}} T_1 \\ \therefore \frac{\Gamma(\Xi^- \rightarrow \Lambda\pi^-)}{\Gamma(\Xi^0 \rightarrow \Lambda\pi^0)} &= \left| \frac{T(\Xi^- \rightarrow \Lambda\pi^-)}{T(\Xi^0 \rightarrow \Lambda\pi^0)} \right|^2 = 2 \end{aligned}$$

- 3.

$$\begin{aligned} |s\Sigma^+ \rangle &= \sqrt{\frac{1}{3}} \left| \frac{3}{2}, \frac{1}{2} \right\rangle + \sqrt{\frac{2}{3}} \left| \frac{1}{2}, \frac{1}{2} \right\rangle \\ |s\Sigma^- \rangle &= \left| \frac{3}{2}, -\frac{3}{2} \right\rangle \\ |n\pi^+ \rangle &= \sqrt{\frac{1}{3}} \left| \frac{3}{2}, \frac{1}{2} \right\rangle + \sqrt{\frac{2}{3}} \left| \frac{1}{2}, \frac{1}{2} \right\rangle \\ |n\pi^- \rangle &= \left| \frac{3}{2}, -\frac{3}{2} \right\rangle \\ |p\pi^0 \rangle &= \sqrt{\frac{2}{3}} \left| \frac{3}{2}, \frac{1}{2} \right\rangle - \sqrt{\frac{1}{3}} \left| \frac{1}{2}, \frac{1}{2} \right\rangle \end{aligned}$$

From above relations

$$\begin{aligned}
 a^+ &= \langle \Sigma^+ | T | n\pi^+ \rangle = \frac{1}{3} T_{\frac{3}{2}} + \frac{2}{3} T_{\frac{1}{2}} \\
 a^- &= \langle \Sigma^- | T | n\pi^- \rangle = T_{\frac{3}{2}} \\
 a^0 &= \langle \Sigma^+ | T | p\pi^0 \rangle = \frac{\sqrt{2}}{3} (T_{\frac{3}{2}} - T_{\frac{1}{2}}) \\
 \therefore a^+ + \sqrt{2}a^0 &= a^-
 \end{aligned}$$

Problem 15.10: From Eqs. (15.132), we can calculate the transition amplitude and decay rate as

$$\begin{aligned}
 \mathcal{M} &= \frac{G_s}{\sqrt{2}} \{ f_+(k_{K\mu} + k_{\pi\mu}) + f_-(k_{K\mu} - k_{\pi\mu}) \} \bar{u}(p_\nu) \gamma^\mu (1 - \gamma^5) v(p_e) \\
 d\Gamma &= \sum_{\text{spin}} \frac{|\mathcal{M}|^2}{2m_K} dLIPS = (2\pi)^4 \delta^4(k_K - k_\pi - p_e - p_\nu) \\
 &\quad \sum_{\text{spin}} \frac{|\mathcal{M}|^2}{2m_K} \frac{d^3 k_\pi d^3 p_e d^3 p_\nu}{(2\pi)^6 2\omega_\pi 2E_e 2E_\nu} \\
 k_K^\mu + k_\pi^\mu &= 2k_K^\mu - (k_K^\mu - k_\pi^\mu) = 2k_K^\mu - (p_e^\mu + p_\nu^\mu)
 \end{aligned}$$

The second term becomes m_ℓ by operating on leptonic current. We set $m_\ell = 0$ from now on.

$$\begin{aligned}
 \sum_{\text{spin}} |\mathcal{M}|^2 &= 2G_s^2 |f_+|^2 k_{K\mu} k_{K\nu} \sum |\bar{u}(p_\nu) \gamma^\mu (1 + \gamma^5) v(p_e)|^2 \\
 &\stackrel{\text{(C.9)}}{=} 16G_s^2 |f_+|^2 k_{K\mu} k_{K\nu} \{ p_e^\mu p_\nu^\nu + p_e^\nu p_\nu^\mu - g^{\mu\nu} (p_e \cdot p_\nu) \\
 &\quad + i\epsilon^{\mu\nu\rho\sigma} p_{e\rho} p_{\nu\sigma} \} \\
 &= 16G_s^2 |f_+|^2 \{ 2(k_K \cdot p_e)(k_K \cdot p_\nu) - m_K^2 (p_e \cdot p_\nu) \} \quad (1.17)
 \end{aligned}$$

$$\begin{aligned}
 \therefore d\Gamma &= (2\pi)^4 \delta^4(k_K - k_\pi - p_e - p_\nu) \frac{G_s^2 |f_+|^2}{m_K} \\
 &\quad \times 8 \{ 2(k_K \cdot p_e)(k_K \cdot p_\nu) - m_K^2 (p_e \cdot p_\nu) \} \frac{d^3 k_\pi d^3 p_e d^3 p_\nu}{(2\pi)^9 2\omega 2E_e 2E_\nu} \quad (1.18)
 \end{aligned}$$

To calculate energy spectrum of the electron, one decomposes $d\Gamma$ as follows.

$$\begin{aligned}
 d\Gamma &= \frac{G_s^2 |f_+|^2}{m_K} \frac{d^3 p_e}{(2\pi)^3 2E_e} 8 \{ (2k_K \cdot p_e) k_{K\mu} - m_K^2 p_{e\mu} \} \\
 &\quad \times \int (2\pi)^4 \delta^4(k_K - k_\pi - p_e - p_\nu) p_\nu^\mu \frac{d^3 k_\pi d^3 p_\nu}{(2\pi)^6 2\omega 2E_\nu} \\
 &\quad \int \dots \stackrel{\text{(C.24)}}{=} \frac{(1 - \gamma)^2}{16\pi} (k_K - p_e)^\mu = \frac{1}{4\pi} \left(\frac{E_{\text{max}} - E_e}{m_K - 2E_e} \right)^2 (k_K - p_e)^\mu \\
 \gamma &= \frac{m_\pi^2}{(k_K - p_e)^2} = \frac{m_\pi^2}{m_K(m_K - 2E_e)}, \quad E_{\text{max}} = \frac{m_K^2 - m_\pi^2}{2m_K}
 \end{aligned}$$

Substituting

$$\begin{aligned} \{2(k_K \cdot p_e)k_{K\mu} - m_K^2 p_{e\mu}\}(k_K - p_e)^\mu &= k_K^2(k_K \cdot p_e) - 2(k_K \cdot p_e) \\ &= m_K^2(m_K - 2E_e) \end{aligned}$$

in Eq. (I.16), we have

$$\frac{d\Gamma}{dE_e} = \frac{G_s^2 |f_+|^2}{2\pi^3} m_K E_e^2 \frac{(E_{\max} - E_e)^2}{m_K - 2E_e}$$

To obtain π spectrum, we rewrite Eq. (I.15) as follows.

$$\begin{aligned} d\Gamma &= \frac{G_s^2 |f_+|^2}{4\pi^2 m_K} \frac{d^3 k_\pi}{(2\pi)^3 2\omega} 8\{2k_{K\mu} k_{K\nu} - g_{\mu\nu} m_K^2\} \\ &\quad \int \delta^4(k_K - k_\pi - p_e - p_\nu) p_e^\mu p_\nu^\nu \frac{d^3 p_e d^3 p_\nu}{2E_e 2E_\nu} \\ &\quad \int \dots \stackrel{(10.53)}{=} \frac{\pi}{24} (g^{\mu\nu} + 2q^\mu q^\nu) = 4\{(k_K \cdot q)^2 - m_K^2 q^2\} \\ &\quad = 4m_K^2(\omega^2 - m_\pi^2) = 4m_K^2 k_\pi^2 \end{aligned}$$

where $q = k_K - k_\pi$.

$$\therefore \frac{d\Gamma}{d\omega} = \frac{G_s^2 |f_+|^2}{12\pi^3} m_K k_\pi^3$$

In order to calculate the total decay rate, we integrate the above formula over energy.

Noting

$$\begin{aligned} \omega_{\min} &= m_\pi = \frac{m_K}{2} 2\gamma, \quad \gamma = \frac{m_\pi}{m_K}, \\ \omega_{\max} &= \frac{m_K^2 + m_\pi^2}{2m_K} = \frac{m_K}{2} (1 + \gamma^2) \\ \Gamma &= \frac{G_s^2 |f_+|^2}{192\pi^3} m_K^5 \int_{2\gamma}^{1+\gamma^2} \left(\sqrt{x^2 - 4\gamma^2} \right)^3 dx = \frac{G_s^2 |f_+|^2}{768\pi^3} f \left(\frac{m_\pi^2}{m_K^2} \right) \\ f(\gamma) &= 1 - 8\gamma + 8\gamma^3 - \gamma^4 - 12\gamma^2 \log \gamma \end{aligned}$$

In deriving the result, we have used

$$\begin{aligned} \int (x^2 + C)^{3/2} dx &= \frac{1}{8} [2x(x^2 + C)^{3/2} + 3Cx(x^2 + C)^{1/2} \\ &\quad + 3C^2 \log |x + \sqrt{x^2 + C}|] \end{aligned}$$

Chapter 16: Neutral Kaons

Problem 16.1: Substituting $p|K^0\rangle + q|\bar{K}^0\rangle = \begin{bmatrix} p \\ q \end{bmatrix}$ in Eq. (16.10),

$$\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \lambda \begin{bmatrix} p \\ q \end{bmatrix}$$

Solutions can be obtained by solving the secular equations.

$$\begin{cases} (A_{11} - \lambda)p + A_{12}q = 0 \\ A_{21}p + (A_{22} - \lambda)q = 0 \end{cases}$$

From the above equation, we get

$$\begin{aligned} \frac{q}{p} &= \frac{\lambda - A_{11}}{A_{12}} = \frac{A_{21}}{\lambda - A_{22}} \\ \lambda^2 - (A_{11} + A_{22})\lambda + A_{11}A_{22} - A_{12}A_{21} &= 0 \\ \therefore \lambda_{S,L} &= \frac{A_{11} + A_{22}}{2} \pm \sqrt{\left(\frac{A_{11} - A_{22}}{2}\right)^2 + A_{12}A_{21}} \\ \lambda_S \lambda_L &= A_{11}A_{22} - A_{12}A_{21} \end{aligned}$$

Problem 16.2: Solutions to $x \begin{bmatrix} M & \frac{\Delta}{2} \\ \frac{\Delta}{2} & M \end{bmatrix}$ are give by

$$|K_1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |K_2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Then to first order in δ , the following equality is also true.

$$\begin{aligned} \begin{bmatrix} M + \delta\Delta & \frac{\Delta}{2} \\ \frac{\Delta}{2} & M - \delta\Delta \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 + \delta \\ 1 - \delta \end{bmatrix} &= \left(M + \frac{\Delta}{2}\right) \frac{1}{\sqrt{2}} \begin{bmatrix} 1 + \delta \\ 1 - \delta \end{bmatrix} \\ \therefore |K_{1M}\rangle &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 + \delta \\ 1 - \delta \end{bmatrix} = |K_1\rangle + \delta|K_2\rangle \end{aligned}$$

is a solution in matter with eigenvalue $M + \Delta/2$. Similarly

$$|K_{2M}\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 - \delta \\ -1 - \delta \end{bmatrix} = |K_2\rangle - \delta|K_1\rangle$$

is also a solution with eigenvalue $M - \Delta/2$. Note, as is shown in Eq. (16.35), M represents $(\lambda_S + \lambda_L)/2 - \frac{\pi N}{m}(f + \bar{f})$.

Problem 16.3: Use Eq. (16.13) in the text,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = U \begin{bmatrix} A_S & 0 \\ 0 & A_L \end{bmatrix} U^{-1}, \quad \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = U \begin{bmatrix} \alpha' \\ \beta' \end{bmatrix}$$

U can be obtained as follows. Using Eqs. (16.65), $\psi = \alpha|K^0\rangle + \beta|\bar{K}^0\rangle$ can be reexpressed in the $|K_S\rangle, |K_L\rangle$ base as

$$\begin{aligned}
 \psi &= \alpha'|K_S\rangle + \beta'|K_L\rangle \\
 &= \alpha'(p\sqrt{1-z}|K^0\rangle + q\sqrt{1+z}|\bar{K}^0\rangle) \\
 &\quad + \beta'(p\sqrt{1+z}|K^0\rangle - q\sqrt{1-z}|\bar{K}^0\rangle) \\
 &= (\alpha'p\sqrt{1-z} + \beta'p\sqrt{1+z})|K^0\rangle + (\alpha'q\sqrt{1+z} - \beta'q\sqrt{1-z})|\bar{K}^0\rangle \\
 &= \alpha|K^0\rangle + \beta|\bar{K}^0\rangle \\
 \therefore U &= \begin{bmatrix} p\sqrt{1-z} & p\sqrt{1+z} \\ q\sqrt{1+z} & -q\sqrt{1-z} \end{bmatrix} \quad U^{-1} = \frac{1}{2} \begin{bmatrix} \frac{\sqrt{1-z}}{p} & \frac{\sqrt{1+z}}{q} \\ \frac{\sqrt{1+z}}{p} & -\frac{\sqrt{1-z}}{q} \end{bmatrix}
 \end{aligned}$$

The rest of the calculations is straightforward.

Problem 16.5: From Eq. (16.52c), CP violation means there is a small imaginary part in $M_{12}^* \Gamma_{12}$. $\therefore M_{12}^* \Gamma_{12} = \bar{M}_{12} \bar{\Gamma}_{12} e^{i\Delta\xi}$ $\Delta\xi \ll 1$ where $\bar{M}_{12}, \bar{\Gamma}_{12}$ are signed real numbers. Writing $M_{12} = \bar{M}_{12} e^{i\xi_M}$, $\Gamma_{12} = \bar{\Gamma}_{12} e^{i\xi_M} e^{i\Delta\xi}$,

$$\begin{aligned}
 \sqrt{A_{12}A_{21}} &= [(M_{12} - i\Gamma_{12}/2)(M_{12}^* - i\Gamma_{12}^*/2)]^{1/2} \\
 &= \bar{M}_{12} - i\bar{\Gamma}_{12}/2 + O(\Delta\xi^2) \\
 \therefore A_\lambda &= -2\sqrt{A_{12}A_{21}} \simeq -2(\bar{M}_{12} - i\bar{\Gamma}_{12}/2) \rightarrow (16.81d)
 \end{aligned}$$

Problem 16.6: Use Eq. (16.79d) and Eq. (16.80d).

Problem 16.7: Using

$$|K^0\rangle = \frac{1}{2p}(|K_S\rangle + |K_L\rangle), \quad |\bar{K}^0\rangle = \frac{1}{2q}(|K_S\rangle - |K_L\rangle), \quad |p|^2 + |q|^2 = 1$$

We calculate the decay asymmetry δ_f . Denoting the final state phase space as ρ ,

$$\begin{aligned}
 \delta_f &\equiv \frac{\Gamma(\bar{K}^0 \rightarrow f) - \Gamma(K^0 \rightarrow f)}{\Gamma(\bar{K}^0 \rightarrow f) + \Gamma(K^0 \rightarrow f)} = \frac{\Gamma(\bar{K}^0 \rightarrow f) - \Gamma(K^0 \rightarrow f)}{\Gamma(K_S \rightarrow f)} \\
 &= \frac{1}{\Gamma(K_S \rightarrow f)} \left\{ \left| \frac{1}{2q} (\langle f|T|K_S\rangle - \langle f|T|K_L\rangle) \right|^2 \rho \right. \\
 &\quad \left. - \left| \frac{1}{2p} (\langle f|T|K_S\rangle + \langle f|T|K_L\rangle) \right|^2 \rho \right\} \\
 &= \frac{1}{4} \left\{ \left| \frac{1}{q} \left(1 - \frac{\langle f|T|K_L\rangle}{\langle f|T|K_S\rangle} \right) \right|^2 - \left| \frac{1}{p} \left(1 + \frac{\langle f|T|K_L\rangle}{\langle f|T|K_S\rangle} \right) \right|^2 \right\} \\
 &= \frac{|1 - \eta_f|^2}{4|q|^2} - \frac{|1 + \eta_f|^2}{4|p|^2} \simeq \frac{|p|^2 - |q|^2}{4|p|^2|q|^2} - \frac{2\text{Re}\eta_f}{4|p|^2|q|^2} \\
 &\simeq 2(\text{Re}\epsilon - \text{Re}\eta_f)
 \end{aligned}$$

Now consider $|f\rangle = |\pi^+\pi^-\rangle$ or $|\pi^0\pi^0\rangle$. Referring to Eq. (16.120), we have

$$\begin{aligned}\eta_{+-} &= \epsilon_0 + \epsilon', \quad \epsilon_0 - 2\epsilon' \\ \therefore \delta_{\pi^0\pi^0} &= 2\text{Re}(\epsilon - \eta_{00}) = 4\text{Re}\epsilon', \quad \delta_{+-} = 2\text{Re}(\epsilon - \eta_{+-}) = -2\epsilon'\end{aligned}$$

Problem 16.8:

$$\begin{aligned}\left|\frac{\bar{A}_f}{A_f}\right| &= \left|\frac{(A_{Sf} - A_{Lf})/2q}{(A_{Sf} + A_{Lf})/2p}\right| = \left|\frac{p}{q} \frac{1 - \eta_f}{1 + \eta_f}\right| \simeq 1 + 2\text{Re}(\epsilon - \eta_f) \\ \left|\frac{\bar{A}_{\pi^0\pi^0}}{A_{\pi^0\pi^0}}\right| - \left|\frac{\bar{A}_{\pi^+\pi^-}}{A_{\pi^+\pi^-}}\right| &= 2\text{Re}(\eta_{+-} - \eta_{00}) = 6\text{Re}\epsilon'\end{aligned}$$

Chapter 17: Hadron Structure

Problem 17.1:

$$\begin{aligned}(17.8) : \quad F(Q^2) &= \frac{\int d\mathbf{r} \rho(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}}}{\int d\mathbf{r} \rho(\mathbf{r})} : \\ \rho(\mathbf{r}) &= \rho_0 \quad r \leq R, \quad \rho = 0 \quad r > R \\ \int d\mathbf{r} \rho(\mathbf{r}) &= 4\pi\rho_0 \int_0^R r^2 dr = \frac{4\pi\rho_0}{3} R^3 \\ \int d\mathbf{r} \rho e^{i\mathbf{Q}\cdot\mathbf{r}} &= 2\pi\rho_0 \int_0^R r^2 dr \int_{-1}^{+1} e^{iQrz} dz = \frac{4\pi\rho_0}{Q} \int_0^R r \sin Qr dr \\ &= \frac{4\pi\rho_0}{Q^3} \int_0^{RQ} \xi \sin \xi d\xi = \frac{4\pi\rho_0}{Q^3} \{\sin \xi - \xi \cos \xi\} \\ \therefore F(Q^2) &= \frac{3(\sin \xi - \xi \cos \xi)}{\xi^3}, \quad \xi = QR \\ (17.9) : \quad \rho(\mathbf{r}) &= \rho_0 e^{-r/a} \\ \int d\mathbf{r} \rho(\mathbf{r}) &= 4\pi\rho_0 \int_0^\infty e^{-r/a} r^2 dr = 8\pi\rho_0 a^3 \\ \int d\mathbf{r} \rho_0 e^{-r/a} e^{i\mathbf{Q}\cdot\mathbf{r}} &= 2\pi\rho_0 \int_0^\infty e^{-r/a} r^2 dr \int_{-1}^{+1} e^{iQrz} dz \\ &= \frac{2\pi\rho_0}{iQ} \int_0^\infty r (e^{-(r/a-iQr)} - e^{-(r/a+iQr)}) dr \\ &= \frac{2\pi\rho_0}{iQ} \left(\frac{1}{(1/a-iQ)^2} - \frac{1}{(1/a+iQ)^2} \right) = 8\pi\rho_0 \frac{a^3}{(1+a^2Q^2)^2} \\ \therefore F(Q^2) &= \frac{1}{(1+a^2Q^2)^2}\end{aligned}$$

Problem 17.2:

$$eJ^\mu(x) = e\bar{u}(p') \left[F_1(q^2) \gamma^\mu - \frac{\chi}{2M} F_2(q^2) i\sigma^{\mu\nu} q_\nu \right] u(p) e^{iqx}$$

We first prove $\partial_\mu J^\mu = 0$. As $\partial_\mu \rightarrow i q_\mu$, $q_\mu \gamma^\mu = \not{p}' - \not{p}$, inserting the Dirac equation $\bar{u}(p') \not{p}' = M \bar{u}(p')$, $\not{p} u(p) = M u(p)$, the first term vanishes. The second term $q_\mu \sigma^{\mu\nu} q_\nu$ vanishes also from the different symmetry of $q_\mu q_\nu$ and $\sigma^{\mu\nu}$. If there exists a term proportional to q^μ , it enters in J^μ as $F_3 q^\mu$. Since, $q_\mu q^\mu = q^2$ does not vanish generally, the current conservation requires $F_3 = 0$. Next, we show that a term like $(p' + p)^\mu$ can be expressed in terms of γ^μ and $i \sigma^{\mu\nu} q_\nu$.

$$\begin{aligned} i \sigma^{\mu\nu} q_\nu &= \frac{1}{2} [-\gamma^\mu (\not{p}' - \not{p}) + (\not{p}' - \not{p}) \gamma^\mu] \\ &= \frac{1}{2} [-2(p' + p)^\mu + 2 \not{p}' \gamma^\mu + 2 \gamma^\mu \not{p}] \\ &= -(p' + p)^\mu + 2M \gamma^\mu \end{aligned}$$

Next, using Eq. (17.15), we calculate $J^\mu A_\mu$. In the nonrelativistic limit,

$$\begin{aligned} \frac{(p' + p)^\mu}{2M} &\longrightarrow (1; \mathbf{v}), \quad q_0 \rightarrow 0 \quad \therefore \quad \frac{(p' + p)^\mu A_\mu}{2M} \rightarrow (\phi - \mathbf{v} \cdot \mathbf{A}) \\ i \sigma^{\mu\nu} q_\mu A_\nu &\longrightarrow \frac{A_0}{2} [\boldsymbol{\gamma} \cdot \mathbf{q} \gamma^0 - \gamma^0 \boldsymbol{\gamma} \cdot \mathbf{q}] + i \varepsilon^{ijk} \sigma_k q_i A_j \\ &= -A_0 \boldsymbol{\alpha} \cdot \boldsymbol{\alpha} + \boldsymbol{\sigma} \cdot (i \mathbf{q} \times \mathbf{A}) = \boldsymbol{\sigma} \cdot \mathbf{B} \end{aligned}$$

where we have used the fact that $\boldsymbol{\alpha}$ goes to zero in the nonrelativistic limit as it has no diagonal elements and that $i \mathbf{q} \times \mathbf{A} = \nabla \times \mathbf{A} = \mathbf{B}$.

Problem 17.3: Energy-momentum conservation gives

$$\begin{aligned} M_X^2 &= P_F^2 = (k + p - k')^2 = (k - k')^2 + 2p \cdot (k - k') + p^2 \\ &= q^2 + 2M(\omega - \omega') + M^2 \end{aligned}$$

Using $Q^2 = -q^2$, $\omega - \omega' = \frac{Q^2}{2M} + \frac{M_X^2 - M^2}{2M}$. For the elastic scattering,

$$\begin{aligned} M_X &= M, \\ Q^2 &= -(k - k')^2 = -k^2 - k'^2 + 2k \cdot k' = 2\omega\omega'(1 - \cos \theta) \\ \omega - \omega' &= \frac{\omega\omega'}{M} (1 - \cos \theta) \end{aligned}$$

Problem 17.4: Since, $q^\mu L_{\mu\nu} = q^\nu L_{\mu\nu} = 0$, only terms proportional to $g^{\mu\nu}$, $p^\mu p^\nu$ in $W^{\mu\nu}$ do not vanish.

$$\begin{aligned} g^{\mu\nu} L_{\mu\nu} &\stackrel{(17.34b)}{=} 2[2k' \cdot k + 2q^2] = 2q^2 = -8\omega\omega' \sin^2 \frac{\theta}{2} \\ p^\mu p^\nu L_{\mu\nu} &= 2 \left[2(p \cdot k)(p \cdot k') + \frac{q^2 p^2}{2} \right] \\ &= 2 \left[2M^2 \omega\omega' - 2M^2 \omega\omega' \sin^2 \frac{\theta}{2} \right] \\ &= 4M^2 \omega\omega' \cos^2 \frac{\theta}{2} \\ \therefore L_{\mu\nu} W^{\mu\nu} &= 4\omega\omega' \left[W_2 \cos^2 \frac{\theta}{2} + 2W_1 \sin^2 \frac{\theta}{2} \right] \end{aligned}$$

Substituting $L_{\mu\nu} W^{\mu\nu}$ in Eq. (17.34a), and using $k \cdot p = M\omega$, $d^3k' = \omega'^2 d\omega' d\Omega$

$$\begin{aligned} d\sigma &= \frac{(4\pi\alpha)^2}{4(k \cdot p)} \frac{4\pi M L_{\mu\nu} W^{\mu\nu}}{Q^4} \frac{d^3k'}{(2\pi)^3 2\omega'} \\ &= \frac{4\alpha^2}{Q^4} \omega'^2 \left[W_2 \cos^2 \frac{\theta}{2} + 2 W_1 \sin^2 \frac{\theta}{2} \right] d\omega' d\Omega \end{aligned}$$

Problem 17.5: Comparing Eq. (17.12) with Eq. (17.41)

$$\int W_2 d\omega' = \frac{\omega'}{\omega}, \quad \int 2 W_1 d\omega' = \frac{Q^2}{2M^2} \frac{\omega'}{\omega}$$

Then using Eq. (17.43), we can obtain Eq. (17.42).

Alternative Method: Insert $J^\mu(p) = \bar{u}(p')\gamma^\mu u(p)$ in Eq. (17.34c) for $W^{\mu\nu}$.

$$\begin{aligned} &\sum_{\text{spin}} J_p^\mu J_p^{\nu*} \\ &= \sum \bar{u}(p')\gamma^\mu u(p) \bar{u}(p)\gamma^\nu u(p') = \text{Tr}[(\not{p}' - M)\gamma^\mu(\not{p} - M)\gamma^\nu] \\ &\stackrel{\text{(C.10)}}{=} 4 \left[p^\mu p'^\nu + p^\nu p'^\mu + \frac{q^2}{2} g^{\mu\nu} \right] \end{aligned}$$

Although

$$p = (M; 0), \quad p' = p + q = P_F = (E_F; \mathbf{P}_F) = (M + \omega - \omega'; \mathbf{k} - \mathbf{k}')$$

q term in p' does not contribute because of Eq. (17.39). Considering final state density $\sum_F = \frac{d^3 P_F}{(2\pi)^3 2E_F}$, etc.

$$\begin{aligned} W^{\mu\nu} &= \frac{(2\pi)^4}{4\pi M} \sum_F \delta^4(p + q - P_F) \frac{1}{2} \sum_{\text{spin}} J_p^\mu J_p^{\nu*} \\ &= \frac{M}{E_F} \delta(M + \nu - E_F) \left[-g^{\mu\nu} \frac{Q^2}{4M^2} + \frac{p^\mu p^\nu}{M^2} \right] \end{aligned}$$

Considering

$$E_F = \sqrt{(\mathbf{k} - \mathbf{k}')^2 + M^2} = \sqrt{Q^2 + \nu^2 + M^2}, \quad \frac{\partial(\nu - E_F)}{\partial\nu} = \frac{M}{E_F}$$

we have

$$\begin{aligned} W^{\mu\nu} &= \delta \left(\nu - \frac{Q^2}{2M} \right) \left[-g^{\mu\nu} \frac{Q^2}{4M^2} + \frac{p^\mu p^\nu}{M^2} \right] \\ \therefore \quad W_2 &= \delta \left(\nu - \frac{Q^2}{2M} \right), \quad W_1 = \frac{Q^2}{4M} \delta \left(\nu - \frac{Q^2}{2M} \right) \end{aligned}$$

Problem 17.6: Using Eq. (17.30) in the text, $x = \frac{Q^2}{2M\nu} < 1$. For γ it is obvious from Eqs. (17.61b).

Problem 17.7: Using equality

$$\frac{\partial x}{\partial \cos \theta} = -\frac{\omega}{M} \frac{1-\gamma}{\gamma}, \quad \frac{\partial \gamma}{\partial \omega'} = -\frac{1}{\omega}$$

that can be derived from Eqs. (17.61), we have

$$J \equiv \left| \begin{array}{cc} \frac{\partial x}{\partial \cos \theta} & \frac{\partial x}{\partial \omega'} \\ \frac{\partial \gamma}{\partial \cos \theta} & \frac{\partial \gamma}{\partial \omega'} \end{array} \right| = \frac{1}{M} \frac{(1-\gamma)}{\gamma}$$

Considering

$$\begin{aligned} \omega'^2 &= \omega^2(1-\gamma)^2, \quad \cos^2 \frac{\theta}{2} \simeq 1, \quad \sin^2 \frac{\theta}{2} = \frac{M^2 x \gamma}{s(1-\gamma)}, \\ s &= 2\omega M, \quad \nu = \frac{s\gamma}{2M} \\ \frac{d^2 \sigma}{dx d\gamma} &= \frac{2\pi}{J} \frac{d^2 \sigma}{d\omega' d\Omega} \stackrel{(17.41)}{=} 2\pi M \frac{\gamma}{1-\gamma} \frac{4\alpha^2}{Q^4} \left(\frac{s^2}{4M^2} \right) (1-\gamma)^2 \\ &\quad \times \left[W_2 + \frac{2M^2 x \gamma}{s(1-\gamma)} W_1 \right] \\ &= \frac{4\pi \alpha^2}{Q^4} s \left[\nu W_2(1-\gamma) + 2x M \left(\frac{\gamma^2}{2} \right) W_1 \right] \end{aligned}$$

Problem 17.8: $\varepsilon^\mu(0)$ is a Lorentz vector, and satisfies

$$q_\mu \varepsilon^\mu(0) = a \left(q \cdot p - \frac{(p \cdot q)}{q^2} q^2 \right) = 0$$

Namely, it satisfies the condition to be a photon polarization vector. Furthermore, it is normalized as $\varepsilon_\mu(0)\varepsilon^\mu(0) = 1$ and is time-like. In fact, using variables evaluated in the target protons rest frame,

$$\begin{aligned} a &= \frac{Q}{M\sqrt{Q^2 + \nu^2}}, \quad p^\mu = (M; 0, 0, 0), \quad (p \cdot q) = M\nu, \\ q^\mu &= (\nu; 0, 0, \sqrt{\nu^2 + Q^2}), \quad (Q^2 = -q^2) \end{aligned}$$

We can show it agrees with Eq. (17.77b) in the text.

Problem 17.9: Using $\varepsilon(0)$ in the Problem 17.8, $W^{\mu\nu}$ can be expressed as

$$W^{\mu\nu} \stackrel{(17.38)}{=} \left(\left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right) W_1 + \frac{\varepsilon^\mu(0)\varepsilon^\nu(0)}{a^2 M^2} W_2 \right)$$

Using $\varepsilon_\mu(\pm)\varepsilon^\mu(0) = 0$, $\varepsilon_\mu(\pm)g^{\mu\nu}\varepsilon_\nu(\pm)^* = -1$

$$\begin{aligned} \sigma_T &= \frac{1}{2} \frac{4\pi^2 \alpha}{K} \sum_{\lambda=\pm} \varepsilon_\mu(\lambda) \varepsilon_\nu(\lambda)^* W^{\mu\nu} = \frac{4\pi^2 \alpha}{K} W_1 \\ \sigma_L &= \frac{4\pi^2 \alpha}{K} \varepsilon_\mu(0) \varepsilon_\nu(0)^* W^{\mu\nu} = \frac{4\pi^2 \alpha}{K} \left(-W_1 + \frac{1}{a^2 M^2} W_2 \right) \\ &= \frac{4\pi^2 \alpha}{K} \left[\left(1 + \frac{\nu^2}{Q^2} \right) W_2 - W_1 \right] \end{aligned}$$

In deriving the last two equalities, we have used

$$\varepsilon_\mu(0)g^{\mu\nu}\varepsilon_\nu(0)^* = 1, \quad a^{-2} = p^2 - \frac{(p \cdot q)}{q^2} = M^2 \left(1 + \frac{v^2}{Q^2}\right)$$

Problem 17.10: Denoting the parton mass as μ , and the energy and momentum of the proton in the Breit frame as $(E; 0, 0, P)$, the parton's momentum transfer q^2 can be calculated as

$$\begin{aligned} p &= (p^0; p_x, p_y, xP), \quad p' = (p^0; p_x, p_y, -xP); \\ p^0 &= \sqrt{p_x^2 + p_y^2 + x^2 P^2 + \mu^2} \\ q^2 &= (p' - p)^2 = -4x^2 P^2, \quad \epsilon^\mu(0) = (1, 0, 0, 0) \end{aligned}$$

As the electron scatters with the parton elastically, the structure function $W^{\mu\nu}$ on the parton target is proportional to

$$L^{\mu\nu} = 2 \left(p^\mu p'^\nu + p^\nu p'^\mu + g^{\mu\nu} \frac{q^2}{2} \right)$$

(See Eq. (17.34b)). Therefore, the photon-parton cross section in the Breit frame is given by

$$\begin{aligned} \sigma_L \propto \varepsilon_\mu(0)\varepsilon_\nu(0)^* W^{\mu\nu} &\propto L^{00} = 2 \left(2p^0 p'^0 + \frac{q^2}{2} \right) \\ &= 4(p_T^2 + \mu^2 + x^2 P^2 - x^2 P^2) = 4(p_T^2 + \mu^2) \\ \sigma_T &\propto \frac{1}{2} \sum_{\lambda=\pm} \varepsilon_\mu(\lambda)\varepsilon_\nu(\lambda)^* W^{\mu\nu} \propto \frac{1}{2} (L^{xx} + L^{yy}) = 2p_x p'_x + 2p_y p'_y - q^2 \\ &= Q^2 + 2p_T^2 \xrightarrow{Q^2 \rightarrow \infty} Q^2, \quad \therefore \quad \frac{\sigma_L}{\sigma_T} = \frac{4(p_T^2 + \mu^2)}{Q^2} \end{aligned}$$

Problem 17.11:

$$\frac{d^2\sigma}{d\omega' d\Omega} = \frac{4\alpha^2}{Q^4} \omega'^2 \left[W_2 \cos^2 \frac{\theta}{2} + 2W_1 \sin^2 \frac{\theta}{2} \right] \quad (I.19)$$

We re-write the contents in $[\dots]$ as follows

$$\begin{aligned} W_2 + \tan^2 \frac{\theta}{2} W_1 &= 2 \tan^2 \frac{\theta}{2} \\ &\times \frac{\left(1 + \frac{v^2}{Q^2}\right) W_2 - W_1 + \left\{1 + 2 \left(1 + \frac{v^2}{Q^2}\right) \tan^2 \frac{\theta}{2}\right\} W_1}{2 \left(1 + \frac{v^2}{Q^2}\right) \tan^2 \frac{\theta}{2}} \\ &\equiv 2 \tan^2 \frac{\theta}{2} \frac{1}{1 - \varepsilon} (W_1 + \varepsilon W_L) \end{aligned}$$

Here we used $W_L = (1 + v^2/Q^2) W_2 - W_1$. Inserting

$$\frac{\omega'^2 \cos^2(\theta/2)}{Q^2} = \frac{1}{4} \frac{\omega'}{\omega} \frac{1}{\tan^2(\theta/2)} \quad W_{T,L} = \frac{K}{4\pi^2\alpha} \sigma_{T,L}$$

in Eq. (I.19)

$$\frac{d^2\sigma}{d\omega'd\Omega} = \Gamma(\sigma_T + \sigma_L), \quad \Gamma = \frac{\alpha K}{2\pi^2 Q^2} \frac{\omega'}{\omega} \frac{1}{1-\varepsilon}$$

Problem 17.12: First we derive the parton approximation for ε . According to Eqs. (17.61), $\nu \sim O(s/M)$, $Q^2 \sim O(s)$, $\nu^2 \gg Q^2$,

$$\begin{aligned} \varepsilon &= \left[1 + 2 \frac{\nu^2 + Q^2}{Q^2} \tan^2 \frac{\theta}{2} \right]^{-1} \simeq \left[1 + 2 \frac{\left(\frac{s}{2M}\gamma\right)^2}{sxy} \frac{M^2 xy}{s(1-\gamma)} \right]^{-1} \\ &= \frac{2(1-\gamma)}{1+(1-\gamma)^2} \\ \therefore \frac{1}{1-\varepsilon} &= \frac{1+(1-\gamma)^2}{\gamma^2} \end{aligned}$$

Next considering $\frac{\partial\gamma}{\partial\omega'} = -\frac{1}{\omega}$, $\frac{\partial Q^2}{\partial \cos\theta} = -2\omega\omega'$, we have

$$\begin{aligned} dQ^2 d\gamma &= \frac{\partial(\gamma, Q^2)}{\partial(\omega', \cos\theta)} = 2\omega' d\omega' d\cos\theta, \\ \frac{d^2\sigma}{dQ^2 d\gamma} &= \frac{\pi}{\omega'} \frac{d^2}{d\omega' d\Omega} = \frac{\pi}{\omega'} \frac{\alpha^2 K}{2\pi^2 Q^2} \frac{\omega'}{\omega} \frac{1+(1-\gamma)^2}{\gamma^2} \sigma_T \end{aligned}$$

Inserting $K/\omega \simeq \nu/\omega = \gamma$

$$d^2\sigma = \frac{\alpha}{2\pi} \frac{1+(1-\gamma)^2}{\gamma} \sigma_T \frac{dQ^2}{Q^2} d\gamma$$

Problem 17.13: Use (C.8c).

Chapter 18: Gauge Theories

Problem 18.1: Let the field strength inside the solenoid with radius a be B_0 . Since the magnetic field is along the z direction, we expect the vector potential in $\phi = \tan^{-1}(\gamma/x)$ direction. The magnetic field inside the solenoid can be reproduced by the potential $\mathbf{A}_{\text{in}} = \frac{B_0}{2}(-\gamma, x) = \frac{1}{2}\mathbf{B} \times \mathbf{r} = \frac{B_0}{2}\rho\mathbf{e}_\phi$ where $\rho = \sqrt{x^2 + \gamma^2}$.

The vector potential outside the solenoid should take a form $\nabla f(\mathbf{r})$, and we note $\nabla\phi$ takes the form

$$\mathbf{A}_{\text{out}} = k \left(-\frac{\gamma}{x^2 + \gamma^2}, \frac{x}{x^2 + \gamma^2}, 0 \right) = \frac{k}{\rho} \mathbf{e}_\phi.$$

The constant k can be determined by the condition,

$$\oint \mathbf{A}_{\text{out}} \cdot d\mathbf{r} = \int B_n dS = B_0 \pi a^2$$

As $\mathbf{A}_{\text{out}} = (k/\rho)\mathbf{e}_\phi$, $\oint_{|r=\rho|} \mathbf{A} \cdot d\mathbf{r} = \int_0^{2\pi} \frac{k}{\rho} \rho d\phi = 2\pi k = \pi a^2 B_0$, $\therefore k = \frac{1}{2} B_0 a^2$

k can also be determined by the condition

$$A_{\text{in}}(\rho = a) = A_{\text{out}}(\rho = a)$$

Let us consider a function $\phi = \tan^{-1}(\frac{y}{x})$, which has a cut at $y = 0, x < 0$ because of periodicity of the angle. ϕ differs by 2π on both sides of the cut.

$$\begin{aligned}\nabla\phi &\equiv \nabla \tan^{-1}\left(\frac{y}{x}\right) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\right) \\ (x < 0, y > 0), \quad \tan\phi < 0 &\Rightarrow \phi = \pi \quad \text{at } y \rightarrow +0 \\ (x < 0, y < 0), \quad \tan\phi > 0 &\Rightarrow \phi = -\pi \quad \text{at } y \rightarrow -0 \\ \therefore \quad \phi_{x<0, y \rightarrow +0} - \phi_{x<0, y \rightarrow -0} &= 2\pi\end{aligned}$$

By a gauge transformation

$$A \rightarrow A' = A + \nabla\chi, \quad \chi = -\frac{1}{2}B_0 a^2 \phi(x, y), \quad \phi(x, y) = \tan^{-1}(y/x)$$

A' can be made 0 everywhere except at the cut, which is called a Dirac string.

Problem 18.7:

$$\begin{aligned}D^\mu \Phi^\dagger (D_\mu \Phi) &\sim \Phi^\dagger \left(-ig_W \mathbf{W} \cdot \mathbf{t} - i\frac{g_B}{2} B \cdot Y \right) \\ &\times \left(ig_W \mathbf{W} \cdot \mathbf{t} + i\frac{g_B}{2} B \cdot Y \right) \Phi\end{aligned}$$

Use $\mathbf{t} = \boldsymbol{\tau}/2$, $Y\Phi = 1$ and replace Φ with $(0, v/\sqrt{2})$ and

$$\begin{aligned}D^\mu \Phi^\dagger (D_\mu \Phi) &\sim (0, v/\sqrt{2})(1/2)^2 \\ &\times \begin{bmatrix} g_W W^0 + g_B B, & \sqrt{2}g_W W^+ \\ \sqrt{2}g_W W^-, & -g_W W^0 + g_B B \end{bmatrix}^2 \begin{bmatrix} 0 \\ v/\sqrt{2} \end{bmatrix} \\ &= \frac{v^2}{8} \left[2g_W^2 W^- W^+ + (-g_W W^0 + g_B B)^2 \right] \\ &= \frac{g_W^2 v^2}{4} W^- W^+ + \frac{g_Z^2 v^2}{8} Z^0 Z^0\end{aligned}$$

Considering there is a factor 2 difference in the mass term between charged and neutral boson, namely $\langle \text{mass term in the Lagrangian} = \frac{m^2}{2} \sum_{i=1}^3 W_i^2 = m^2(W^- W^+ + W^{02}/2) \rangle$ we have $m_W = \frac{g_W v}{2}$, $m_Z = \frac{g_Z v}{2}$.

Appendix J

Particle Data

Table J.1 Baryons

Name	Quark contents	Spin	Mass (MeV/ c^2)	Life time (sec) /width (MeV)	Principal decays
N p	uud	1/2	938.272	∞	–
n	udd	1/2	939.565	885.7 s	$p e^- \bar{\nu}_e$
Λ	uds	1/2	1115.68	2.63×10^{-10} s	$p \pi^-, n \pi^0$
Σ^+	uus	1/2	1189.37	8.02×10^{-11} s	$p \pi^0, n \pi^+$
Σ^0	uds	1/2	1192.64	7.4×10^{-20} s	$\Lambda \gamma$
Σ^-	dds	1/2	1197.45	1.48×10^{-10} s	$n \pi^-$
Ξ^0	uss	1/2	1314.9	2.90×10^{-10} s	$\Lambda \pi^0$
Ξ^-	dss	1/2	1321.7	1.64×10^{-10} s	$\Lambda \pi^-$
Λ_c^+	udc	1/2	2286.5	2.00×10^{-13} s	$p K \pi, \Lambda \pi \pi, \Sigma \pi \pi$
$\Delta^{(+++, +, 0, -)}$	$uuu, uud, udd,$ ddd	3/2	1232	118 MeV	$N \pi$
$\Sigma^*(+, 0, -)$	uus, uds, dds	3/2	1385	37 MeV	$\Lambda \pi, \Sigma \pi$
$\Xi^*(0, -)$	uss, dss	3/2	1533	9 MeV	$\Xi \pi$
Ω^-	sss	3/2	1672	8.2×10^{-11} s	$\Lambda K^-, \Xi \pi$

Table J.2 Mesons

Name	Quark contents	Spin	Mass (MeV/ c^2)	Life time (sec) /width (MeV)	Principal decays
$\pi^\pm$	$u\bar{d}, d\bar{u}$	0	139.570	2.60×10^{-8} s	$\mu\nu_\mu$
π^0	$u\bar{u}, d\bar{d}$	0	134.977	8.4×10^{-17} s	$\gamma\gamma$
$K^\pm$	$u\bar{s}, s\bar{u}$	0	493.68	1.24×10^{-8} s	$\mu\nu_\mu, \pi\pi, \pi\pi\pi$
$K^0, \bar{K}^0$	$d\bar{s}, s\bar{d}$	0	497.61	K_S 8.95×10^{-11} s K_L 5.12×10^{-8} s	$\pi\pi$ $\pi\mu\nu_\mu, \pi e\nu_e, \pi\pi\pi$
η	$u\bar{u}, d\bar{d}, s\bar{s}$	0	547.85	1.30 keV	$\gamma\gamma, \pi\pi\pi$
η'	$u\bar{u}, d\bar{d}, s\bar{s}$	0	957.66	0.20 MeV	$\eta\pi\pi, \rho\gamma$
$D^\pm$	$c\bar{d}, d\bar{c}$	0	1869.6	1.04×10^{-12} s	$K\pi\pi, K\mu\nu_\mu, Ke\nu_e$
$D^0, \bar{D}^0$	$c\bar{u}, u\bar{c}$	0	1864.8	4.1×10^{-13} s	$K\pi\pi, K\mu\nu_\mu, Ke\nu_e$
$D_s^\pm$	$c\bar{s}, s\bar{c}$	0	1968.5	5.0×10^{-13} s	$\eta\rho, \eta'\rho, \phi\rho$
$\rho^{(\pm,0)}$	$u\bar{d}, u\bar{u}, d\bar{d}, d\bar{u}$	1	775.5	149 MeV	$\pi\pi$
$K^{* (\pm,0)}$	$u\bar{s}, d\bar{s}, s\bar{d}, s\bar{u}$	1	894	50 MeV	$K\pi$
ω	$u\bar{u}, d\bar{d}$	1	782.6	8.5 MeV	$\pi\pi\pi, \pi\gamma$
ϕ	$s\bar{s}$	1	1019.5	4.3 MeV	$K\bar{K}, \pi\pi\pi$
J/ψ	$c\bar{c}$	1	3097	93 keV	$e^-e^+, \mu^-\mu^+, \rho\pi, \omega\pi$
Υ	$b\bar{b}$	1	9460	54 keV	$e^-e^+, \mu^-\mu^+, \tau^-\tau^+$

Appendix K

Constants

Table K.1 Constants

Quantity	Symbol	Value unit
Speed of light in vacuum	c	$2.997925 \times 10^8 \text{ m s}^{-1}$
Planck constant	h	$6.626069 \times 10^{-34} \text{ J s}$
Planck constant, reduced	$\hbar$	$1.054572 \times 10^{-34} \text{ J s}$ $= 6.582119 \times 10^{-22} \text{ MeV s}$
Conversion constant	$\hbar c$	$197.326963 \text{ MeV fm}^{\text{a}}$
Electron charge magnitude	e	$1.602176 \times 10^{-19} \text{ C}$
Electron mass	m_e	$0.510999 \text{ MeV}/c^2$
Proton mass	m_p	$938.272013 \text{ MeV}/c^2$ $= 1.672622 \times 10^{-27} \text{ kg}$
Neutron mass	m_n	$939.56536 \text{ MeV}/c^2$
Avogadro constant	N_A	$6.022142 \text{ mol}^{-1}$
Boltzmann constant	k	$1.380650 \times 10^{-23} \text{ J K}^{-1}$ $= 8.617343 \times 10^{-5} \text{ eV K}^{-1}$
Fine structure constant	$\alpha = e^2/4\pi\epsilon_0\hbar c$	7.297353×10^{-3} $= 1/137.036000^{\text{b}}$
Classical electron radius	$r_e = \alpha\hbar/m_e c$	$2.817940 \times 10^{-15} \text{ m}$
Bohr radius	$a_\infty = (1/\alpha)\hbar/m_e c$	$0.529177 \times 10^{-10} \text{ m}$
Thomson cross section	$\sigma_T = 8\pi r_e^2/3$	$0.665245 \text{ barn}^{\text{c}}$
Gravitational constant	G_N	$6.67428 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$ $= 6.70881 \times 10^{-39} \hbar c (\text{GeV}/c^2)^{-2}$
Planck mass	$\sqrt{\hbar c/G_N}$	$1.22089 \times 10^{19} \text{ GeV}/c^2$
Planck length	$\sqrt{\hbar G_N/c^3}$	$1.61624 \times 10^{-35} \text{ m}$
Planck time	$\sqrt{\hbar G_N/c^5}$	$5.39124 \times 10^{-44} \text{ s}$
Fermi coupling constant	$G_F/(\hbar c)^3$	$1.16637 \times 10^{-5} \text{ GeV}^{-2}$
$W^\pm$ boson mass	m_W	$80.398 \text{ GeV}/c^2$
Z^0 boson mass	m_Z	$91.1876 \text{ GeV}/c^2$
Weak mixing angle	$\sin^2 \theta_W^{\text{d}}$	0.2231
Strong coupling constant	$\alpha_s(m_Z)$	0.1176

a fm = 10^{-15} m.

b At $Q^2 = 0$, $\sim 1/128$ at $Q^2 = m_W^2$.

c barn = 10^{-28} m^2 .

d defined by $1 - m_W^2/m_Z^2$.

References

- 1 B. Fowler *Nucl. Phys. B (Pro. Supple)*, 36:37, 1994, *ibid.* C. Peyrou, p. 59 and other articles p. 3–568 in “30 years of bubble chamber physics”.
- 2 S.A. Abel and O. Lebedev. *J. High Energy Phys.*, 01:133, 2006.
- 3 H. Abramowicz *et al.* *Z. Phys. C*, 17:283, 1983.
- 4 M.R. Adams *et al.* *Phys. Rev. D*, 54:3006, 1996.
- 5 S.L. Adler. *Phys. Rev.*, 140:B736, 1965.
- 6 S.L. Adler. *Phys. Rev.*, 143:1144, 1966.
- 7 S.L. Adler. *Phys. Rev.*, 177:2426, 1969.
- 8 Y. Aharonov and D. Bohm. *Phys. Rev.*, 115:485, 1959.
- 9 L.A. Ahrens *et al.* *Phys. Lett. B*, 202:284, 1988.
- 10 ALEPH Collaboration. *Nucl. Instrum. Methods A*, 294:121, 1990.
- 11 ALEPH Collaboration. *Eur. Phys. J. C*, 2:197, 1998.
- 12 C. Alff *et al.* *Phys. Rev. Lett.*, 9:322, 1962.
- 13 C. Alff *et al.* *Phys. Rev. Lett.*, 9:325, 1962.
- 14 W.W.M. Allison. The Interaction of Charged Particles and Photons in Matter. In C.W. Fabjan and J.E. Pilcher, editors, *Proceedings of the ICFA School on Instrumentation in Elementary Particle Physics at Trieste, Italy, 1987*, pages 3–42, World Scientific, 1987.
- 15 W.W.M. Allison and J.H. Cobb. *Annu. Rev. Nucl. Part. Sci.*, 30:253, 1980.
- 16 G. Altarelli. *Annu. Rev. Nucl. Part. Sci.*, 39:357, 1989.
- 17 G. Altarelli and G. Parisi. *Nuovo Cimento*, 4:335, 1974.
- 18 G. Altarelli and G. Parisi. *Nucl. Phys. B*, 126:298, 1977.
- 19 L. Alvarez-Gaumé *et al.* *Phys. Lett. B*, 458:347, 1999.
- 20 U. Amaldi *et al.* *Phys. Lett. B*, 260:447, 1991.
- 21 C.D. Anderson. *Phys. Rev.*, 41:405, 1932.
- 22 C.D. Anderson. *Phys. Rev.*, 43:491, 1933.
- 23 H.L. Anderson. *Phys. Rev.*, 100:279, 1955.
- 24 D. Andrews *et al.* *Phys. Rev. Lett.*, 44:1108, 1980.
- 25 P.L. Anthony *et al.* *Phys. Rev. Lett.*, 75:1949, 1995.
- 26 P.L. Anthony *et al.* *Phys. Rev. D*, 56:1373, 1997.
- 27 R. Arai *et al.* *Nucl. Instrum. Methods Phys. Res.*, 217:181, 1983.
- 28 ARGUS Collaboration: H. Albrecht *et al.* *Phys. Lett. B*, 246:278, 1990.
- 29 N. Arkani-Hamed *et al.* *Phys. Lett. B*, 429:263, 1998.
- 30 X. Artru *et al.* *Phys. Rev. D*, 12:1289, 1975.
- 31 J. Ashkin *et al.* *Phys. Rev.*, 101:1149, 1956.
- 32 ATLAS Collaboration. ATLAS-TDR-002; CERN-LHCC-96-041. Technical report, CERN LHCC, 1996.
- 33 J.J. Aubert *et al.* *Phys. Rev. Lett.*, 33:1404, 1974.
- 34 J.-E. Augustin *et al.* *Phys. Rev. Lett.*, 33:1406, 1974.

- 35 N. Bacino. *Phys. Rev. Lett.*, 41:13, 1978.
- 36 J. Bailey *et al.* *Nucl. Phys. B*, 150:1, 1979.
- 37 C.A. Baker *et al.* *Phys. Rev. Lett.*, 97:131801, 2006.
- 38 G. Bakker *et al.* *Nucl. Phys. B*, 16:53, 1970.
- 39 G.S. Bali. *Phys. Rep.*, 343:1, 2001. arXiv:hep-ph/0001312v2.
- 40 B. Balke *et al.* *Phys. Rev. D*, 37:587, 1988.
- 41 C. Baltay. In S. Homma *et al.*, editors, *Proc. 19th Int. Conf. on High Energy Physics, Tokyo, 1978*. Physical Society of Japan, Tokyo, 1979.
- 42 A. Bandyopadhyay *et al.* *Rep. Prog. Phys.*, 72:106201, 2009. arXiv:0710.4947v2 [hep-ph].
- 43 D.P. Barber *et al.* *Phys. Rev. Lett.*, 43:830, 1979.
- 44 M. Bardon *et al.* *Phys. Rev. Lett.*, 14:449, 1965.
- 45 V. Barger *et al.* *Nucl. Phys. B*, 5:411, 1968.
- 46 V. Barger and M. Olsson. *Phys. Rev.*, 146:1080, 1966.
- 47 V. Bargmann *et al.* *Phys. Rev. Lett.*, 2:435, 1959.
- 48 V.V. Barmin *et al.* *Nucl. Phys. B*, 247:293, 1984.
- 49 V.E. Barnes *et al.* *Phys. Rev. Lett.*, 12:204, 1964.
- 50 G. Barr. Presented at Lepton-Photon 99 (LP99), 1999. <http://www.sldnt.slac.stanford.edu/lp99/> (accessed March 2010).
- 51 J.R. Batley *et al.* *Phys. Lett. B*, 544:97, 2002.
- 52 BCDMS Collaboration: A.C. Benvenuti *et al.* *Phys. Lett. B*, 223:485, 1989.
- 53 M.A.B. Bég and A. Sirlin. *Phys. Rep.*, 88:1, 1982.
- 54 A.I. Belesev *et al.* *Phys. Lett. B*, 350:263, 1995.
- 55 J.S. Bell and R. Jackiw. *Nuovo Cimento A*, 60:47, 1969.
- 56 Belle Collaboration: A. Go *et al.* *Phys. Rev. Lett.*, 99:131802, 2007.
- 57 J.B. Bellicard *et al.* *Phys. Rev. Lett.*, 19:527, 1967.
- 58 V.B. Berestetskii, E.M. Lifshitz, and L.P. Pitaevskii. *Relativistic Quantum Theory*. Pergamon, New York, 1971.
- 59 P. Berge *et al.* *Z. Phys. C*, 49:187, 1991.
- 60 J. Bernabeu *et al.* CPT and quantum mechanics tests with kaons. In A. Di Domenico, editor, *Handbook on Neutral Kaon Interferometry at a Phi-Factory*, page 39. Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali di Frascati, 2006. arXiv:hep-ph/0607322v1.
- 61 H. Bethe and W. Heitler. *Proc. R. Soc. London A*, 146:83, 1934.
- 62 J.D. Bjorken. *Phys. Rev.*, 179:1547, 1969.
- 63 J.D. Bjorken and S.D. Drell. *Relativistic Quantum Mechanics*. McGraw-Hill, New York, 1964.
- 64 J.D. Bjorken and S.D. Drell. *Relativistic Quantum Fields*. McGraw-Hill, New York, 1965.
- 65 J.D. Bjorken and E. A. Paschos. *Phys. Rev.*, 185:1975, 1969.
- 66 E. Blanke *et al.* *Phys. Rev. Lett.*, 51:355, 1983.
- 67 F. Bloch and A. Nordsieck. *Phys. Rev.*, 52:54, 1937.
- 68 P. Bloch and L. Tauscher. *Annu. Rev. Nucl. Part. Sci.*, 53:123, 2003.
- 69 E. Bloom *et al.* *Phys. Rev. Lett.*, 23:930, 1969.
- 70 E. Blucher. Presented at Lepton-Photon 99 (LP99), 1999. <http://www.sldnt.slac.stanford.edu/lp99/> (accessed March 2010).
- 71 T. Böhringer *et al.* *Phys. Rev. Lett.*, 44:1111, 1980.
- 72 M. Bourquin *et al.* *Z. Phys. C*, 21:27, 1983.
- 73 A.M. Boyarski *et al.* *Phys. Rev. Lett.*, 34:1357, 1975.
- 74 T.H. Boyer. *Phys. Rev.*, 174:1764, 1968.
- 75 H. Braun *et al.* *Nucl. Phys. B*, 89:210, 1975.
- 76 M. Breidenbach *et al.* *Phys. Rev. Lett.*, 23:935, 1969.
- 77 A. Breskin *et al.* *Nucl. Instrum. Methods*, 124:189, 1975.
- 78 S. J. Brodsky *et al.* *Phys. Rev. D*, 4:1532, 1971.
- 79 C.N. Brown *et al.* *Phys. Rev. Lett.*, 63:2637, 1989.
- 80 R. Brown *et al.* *Nature*, 163:47, 1949.
- 81 F. Bumiller. *Phys. Rev.*, 124:1623, 1961.
- 82 M.T. Burgy *et al.* *Phys. Rev. Lett.*, 1:324, 1958.
- 83 M.T. Burgy *et al.* *Phys. Rev.*, 120:1829, 1960.
- 84 N. Cabibbo. *Phys. Rev. Lett.*, 10:531, 1963.

- 85 C.G. Callan, Jr. and D.J. Gross. *Phys. Rev. Lett.*, 22:156, 1969.
- 86 A. Camacho and A. Camacho-Galván. *Rep. Prog. Phys.*, 70:1937, 2007.
- 87 L. Camilleri *et al.* *Annu. Rev. Part. Nucl. Sci.*, 58:343, 2008.
- 88 D.D. Carmony and R.T. van de Walle. *Phys. Rev. Lett.*, 8:73, 1962.
- 89 P.A. Carruthers. *Introduction to Unitary Symmetry*. John Wiley & Sons, New York, 1966.
- 90 A.A. Carter. *Nucl. Phys. B*, 58:378, 1973.
- 91 W.F. Cartwright *et al.* *Phys. Rev.*, 91:677, 1953.
- 92 H.B.G. Casimir. *Proc., K. Ned. Akad. Wet., B*, 51:793, 1948.
- 93 CERN. <http://public.web.cern.ch/public/en/Research/AccelComplex-en.html> (accessed March 2010).
- 94 J. Chadwick. *Proc. R. Soc. London A*, 136:692, 1932.
- 95 G. Charpak *et al.* *Nucl. Instrum. Methods*, 62:262, 1968.
- 96 M.-S. Chen and P. Zerwas. *Phys. Rev. D*, 12:187, 1975.
- 97 T.-P. Cheng and L.-F. Li. *Gauge Theory of Elementary Particle Physics*. Clarendon Press, Oxford, 1984.
- 98 C.B. Chiu. *Annu. Rev. Nucl. Sci.*, 22:255, 1972.
- 99 J. Christenson *et al.* *Phys. Rev. Lett.*, 13:138, 1964.
- 100 D.L. Clark *et al.* *Phys. Rev.*, 85:523, 1952.
- 101 F.E. Close. *An Introduction to Quarks and Partons*. Academic Press, New York, 1979.
- 102 S. Coleman. *Science*, 206:1290, 1979.
- 103 R. Collela *et al.* *Phys. Rev. Lett.*, 34:1472, 1975.
- 104 E.D. Commins and P.H. Bucksbaum. *Weak Interactions of Leptons and Quarks*. Cambridge University Press, Cambridge, 1983.
- 105 C.L. Cowan, Jr. *et al.* *Science*, 124:103, 1956.
- 106 P. Coyle and O. Schneider. *C.R. Physique*, 3:1143, 2002.
- 107 CPLEAR Collaboration: A. Angelopoulos *et al.* *Phys. Rep.*, 374:165, 2003.
- 108 CPLEAR Collaboration: A. Angelopoulos *et al.* *Phys. Lett. B*, 444:43, 1998.
- 109 CPLEAR Collaboration: R. Adler *et al.* *Nucl. Instrum. Methods Phys. Res. A*, 379:76, 1996.
- 110 J.W. Cronin *et al.* *Phys. Rev. Lett.*, 18:25, 1967.
- 111 G. Culligan *et al.* *Nature*, 180:751, 1957.
- 112 G. Danby *et al.* *Phys. Rev. Lett.*, 9:36, 1962.
- 113 A. Das. *Field Theory: A Path Integral Approach*. World Scientific, Singapore, 2nd edition, 2006.
- 114 DASP Collaboration: R. Brandelik *et al.* *Phys. Lett. B*, 70:132, 1977.
- 115 M. Davier. *Proc. SLAC Summer Institute on Particle Physics*, SLAC-R-528, 1997.
- 116 E.H. de Groot and J. Engels. *Z. Phys. C*, 1:141, 1979.
- 117 A. De Rújula, H. Georgi and S.L. Glashow. *Phys. Rev. D*, 12:147, 1975.
- 118 B.S. Deaver, Jr. and W. Fairbank. *Phys. Rev. Lett.*, 7:43, 1961.
- 119 DELPHI Collaboration. *Nucl. Instrum. Methods Phys. Res. A*, 303:233, 1991.
- 120 T.J. Devlin and J.O. Dickey. *Rev. Mod. Phys.*, 51:237, 1979.
- 121 C.O. Dib and R.D. Peccei. *Phys. Rev. D*, 46:2265, 1992.
- 122 N. Doble *et al.* *Nucl. Instrum. Methods Phys. Res. B*, 119:181, 1996.
- 123 Y.L. Dokshitzer. *Sov. Phys. - JETP*, 46:641, 1977.
- 124 J. Drees and H.E. Montgomery. *Annu. Rev. Part. Sci.*, 33:383, 1983.
- 125 J. Dunkley *et al.* *Astrophys. J. Suppl.*, 180:306, 2009.
- 126 R. Durbin *et al.* *Phys. Rev.*, 83:646, 1951.
- 127 W. Ehrenberg and R.E. Siday. *Proc. Phys. Soc. B*, 62:8, 1949.
- 128 A. Einstein *et al.* *Phys. Rev.*, 47:777, 1935.
- 129 J. Ellis. *Nucl. Instrum. Methods Phys. Res. A*, 284:33, 1989.
- 130 J. Ellis *et al.* *Nucl. Phys. B*, 241:381, 1984.
- 131 F. Englert and R. Brout. *Phys. Rev. Lett.*, 13:321, 1964.
- 132 P.T. Eschstruth *et al.* *Phys. Rev.*, 165:1487, 1968.
- 133 L. Evans *et al.* In G. Altarelli and L. DiLella, editors, *Proton-Antiproton Collider Physics*, page 33. World Scientific, Singapore, 1989.
- 134 C.W. Fabjan *et al.* *Nucl. Instrum. Methods*, 141:61, 1977.

- 135 C.W. Fabjan and T. Ludlum. *Annu. Rev. Nucl. Part. Sci.*, 32:335, 1982.
- 136 L.D. Faddeev and V.N. Popov. *Phys. Lett. B*, 25:29, 1967.
- 137 U. Fano. *Annu. Rev. Nucl. Sci.*, 13:1, 1963.
- 138 T. Ferbel. *Experimental Techniques in High Energy Physics*. Addison-Wesley, Reading, MA, 1987.
- 139 E. Fermi and C.N. Yang. *Phys. Rev.*, 76:1739, 1949.
- 140 W. Fetscher. *Phys. Lett. B*, 140:117, 1984.
- 141 W. Fetscher *et al.* *Phys. Lett. B*, 173:102, 1986.
- 142 R.P. Feynman. In *Proc. 3rd Topical Conf. High Energy Collision of Hadrons*. C.N. Yang, J.A. Cole, M. Good, R. Hwa, and J. Lee-Franzini, editors, Gordon Breach, New York, p. 237, 1969.
- 143 R.P. Feynman. *Phys. Rev. Lett.*, 23:1415, 1969.
- 144 R.P. Feynman and M. Gell-Mann. *Phys. Rev.*, 109:193, 1958.
- 145 H.E. Fisk and F. Sciulli. *Annu. Rev. Nucl. Part. Sci.*, 32:499, 1982.
- 146 L. Fonda *et al.* *Rep. Prog. Phys.*, 41:587, 1978.
- 147 H. Frauenfelder and R.M. Steffan. In K. Siegban, editor, α , β , γ *Ray Spectroscopy*, page 1431. North-Holland, Amsterdam, 1966.
- 148 J.I. Friedman. *Rev. Mod. Phys.*, 63:615, 1991.
- 149 J.I. Friedman and H.W. Kendall. *Annu. Rev. Nucl. Sci.*, 22:203, 1972.
- 150 M. Froissart. *Phys. Rev.*, 123:1053, 1961.
- 151 K. Fujikawa and R.E. Shrock. *Phys. Rev. Lett.*, 45:963, 1980.
- 152 T. Gabriel *et al.* *Nucl. Instrum. Methods*, 171:237, 1980.
- 153 J.-M. Gaillard *et al.* *Phys. Rev. Lett.*, 18:20, 1967.
- 154 M.K. Gaillard and B.W. Lee. *Phys. Rev. D*, 10:897, 1974.
- 155 G.M. Garibian *et al.* *Nucl. Instrum. Methods*, 125:133, 1975.
- 156 S. Gasiorowicz. *Elementary Particle Physics*. John Wiley & Sons, New York, 1966.
- 157 J. Gasser and U.-G. Meißner. *Phys. Lett. B*, 258:219, 1991.
- 158 H. Geiger and E. Marsden. *Philos. Mag.*, 25:604, 1913.
- 159 M. Gell-Mann. *Phys. Rev.*, 92:833, 1953.
- 160 M. Gell-Mann. *Phys. Rev.*, 125:1067, 1962.
- 161 M. Gell-Mann. *Phys. Lett.*, 8:214, 1964.
- 162 M. Gell-Mann and M. Levy. *Nuovo Cim.*, 16:705, 1960.
- 163 M. Gell-Mann, P. Ramond, and R. Slansky. *Supergravity*. North-Holland, Amsterdam, 1979.
- 164 S. Gentile and M. Pohl. *Phys. Rep.*, 274:287, 1996.
- 165 H. Georgi. *Lie Algebras in Particle Physics*. Perseus Books, New York, 2nd edition, 1999.
- 166 H. Georgi and S.L. Glashow. *Phys. Rev. Lett.*, 32:438, 1974.
- 167 S.S. Gershtein and Ya. B. Zeldovich. *Sov. Phys. – JETP*, 2:596, 1956.
- 168 F.J. Gilman. *Phys. Rev.*, 167:1365, 1968.
- 169 F.J. Gilman. *Phys. Rev. D*, 35:3541, 1987.
- 170 V.L. Ginzburg and L.D. Landau. *Zh. Eksp. Teor. Fiz.*, 20:1064, 1950.
- 171 S. Gjesdal *et al.* *Phys. Lett. B*, 52:113, 1974.
- 172 D.A. Glaser. *Phys. Rev.*, 87:665, 1952.
- 173 S.L. Glashow. *Nucl. Phys.*, 22:579, 1961.
- 174 S.L. Glashow *et al.* *Phys. Rev. D*, 2:1285, 1970.
- 175 R.L. Gluckstern. *Nucl. Instrum. Methods*, 24:381, 1963.
- 176 M.L. Goldberger and K.M. Watson. *Collision Theory*. John Wiley & Sons, New York, 1964.
- 177 G. Goldhaber. In *Proc. of the Summer Institute on Particle Physics*, volume 198 of *SLAC Publication*, page 379. 1979.
- 178 G. Goldhaber *et al.* *Phys. Rev. Lett.*, 37:255, 1976.
- 179 M. Goldhaber *et al.* *Phys. Rev.*, 109:1015, 1958.
- 180 J. Goldstone. *Nuovo Cimento*, 19:154, 1961.
- 181 M.L. Good. *Phys. Rev.*, 110:550, 1958.
- 182 K. Gottfried. *Phys. Rev. Lett.*, 18:1174, 1967.
- 183 K. Goulianos. *Phys. Rep. C*, 101:169, 1983.
- 184 M.G. Green, S.L. Lloyd, P.N. Ratoff, and D.R. Ward. *Electron-Positron Physics at the Z*. IOP Publishing, Bristol, 1998.
- 185 O.W. Greenberg. *Phys. Rev. Lett.*, 13:598, 1964.

- 186 B.R. Greene. *The Elegant Universe*. W.W. Norton, New York, 1999.
- 187 G.N. Gribov and L.N. Lipatov. *Sov. J. Nucl. Phys.*, 15:438, 675, 1972.
- 188 D.J. Gross and Llewellyn Smith C.H. *Nucl. Phys. B*, 14:337, 1969.
- 189 D.J. Gross and F. Wilczek. *Phys. Rev. Lett.*, 30:1343, 1973.
- 190 D.J. Gross and F. Wilczek. *Phys. Rev. D*, 8:3633, 1973.
- 191 G.S. Guralnik *et al.* *Phys. Rev. Lett.*, 13:585, 1964.
- 192 F. Gürsey and L.A. Radicati. *Phys. Rev. Lett.*, 13:173, 1964.
- 193 M.Y. Han and Y. Nambu. *Phys. Rev.*, 139:B1006, 1965.
- 194 L.N. Hand. *Phys. Rev.*, 129:1834, 1963.
- 195 G. Hanson *et al.* *Phys. Rev. Lett.*, 35:1609, 1975.
- 196 Y. Hara. *Phys. Rev.*, 134:B701, 1964.
- 197 B.W. Harris *et al.* *Phys. Rev. A*, 62:052109, 2000.
- 198 P.G. Harris *et al.* *Phys. Rev. Lett.*, 82:904, 1999.
- 199 F.J. Hasert *et al.* *Phys. Lett. B*, 46:121, 1973.
- 200 F.J. Hasert *et al.* *Phys. Lett. B*, 46:138, 1973.
- 201 S.W. Hawking. *Commun. Math. Phys.*, 87:395, 1982.
- 202 K Hayashi and T. Nakano. *Prog. Theor. Phys.*, 38:491, 1967.
- 203 W. Heitler. *The Quantum Theory of Radiation*. Oxford University Press, Oxford, 3rd edition, 1954. Republished by Dover, Mineola, NY, 1984.
- 204 A.J.G. Hey and R.L. Kelly. *Phys. Rep.*, 96:71, 1983.
- 205 P.W. Higgs. *Phys. Rev. Lett.*, 13:508, 1964.
- 206 R. Hofstadter. *Rev. Mod. Phys.*, 28:214, 1956.
- 207 G.E. Hogan *et al.* *Phys. Rev. Lett.*, 42:948, 1979.
- 208 K. Huang. *Quarks, Leptons and Gauge Fields*. World Scientific, Singapore, 2nd edition, 1992.
- 209 F. Hubaut. *Nucl. Instrum. Methods A*, 518:31, 2004.
- 210 P. Huet and M.E. Peskin. *Nucl. Phys. B*, 434:3, 1995.
- 211 J. Iizuka. *Prog. Theor. Phys. Suppl.*, 37:21, 1966.
- 212 M. Ikeda *et al.* *Prog. Theor. Phys.*, 22:715, 1959.
- 213 W.R. Innes *et al.* *Phys. Rev. Lett.*, 39:1240, 1977.
- 214 N. Isgur and G. Karl. *Phys. Rev. D*, 18:4187, 1978.
- 215 C. Itzykson and J.-B. Zuber. *Quantum Field Theory*. McGraw-Hill, New York, 1980. Republished by Dover, Mineola, NY, 2006.
- 216 J.D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, New York, 3rd edition, 1998.
- 217 M. Jacob and G.C. Wick. *Annals of Phys.*, 7:404, 1959.
- 218 R.C. Jaklevic *et al.* *Phys. Rev.*, 140:A1628, 1965.
- 219 F. Jegerlehner and A. Nyffeler. *Phys. Rep.*, 477:1, 2009. F. Jegerlehner. arXiv:hep-ph/0703125v3.
- 220 P.K. Kabir. *The CP Puzzle: Strange Decays of the Neutral Kaon*. Academic Press, Boston, MA, 1968.
- 221 P.K. Kabir. *Phys. Rev. D*, 2:540, 1970.
- 222 G. Kallen. *Elementary Particle Physics*. Addison-Wesley, Reading, MA, 1964.
- 223 T. Kamae *et al.* *Nucl. Instrum. Methods Phys. Res. A*, 252:423, 1986.
- 224 K. Kasahara. *Phys. Rev. D*, 31:2737, 1985.
- 225 KEK-E246 Collaboration: M. Abe *et al.* *Phys. Rev. Lett.*, 93:131601, 2004.
- 226 KEK-E246 Collaboration: M. Abe *et al.* *Phys. Rev. D*, 73:072005, 2006.
- 227 K.W. Kendall. *Rev. Mod. Phys.*, 63:597, 1991.
- 228 L.A. Khalfin. *Soviet Phys. Dokl.*, 2:340, 1957. [Original: *Dokl. Akad. Nauk. USSR* 115:277, 1957].
- 229 L.A. Khalfin. *Pis'ma Zh. Eksp. Teor. Fiz.*, 8:106, 1968.
- 230 T.W.B. Kibble. *J. Math. Phys.*, 2:212, 1961.
- 231 T. Kinoshita. *Phys. Rev. Lett.*, 2:477, 1959.
- 232 T. Kinoshita *et al.* *Phys. Rev. Lett.*, 52:717, 1984.
- 233 O.C. Kistner and B.M. Rustad. *Phys. Rev.*, 114:1329, 1959.
- 234 S. Klein. *Rev. Mod. Phys.*, 71:1501, 1999.
- 235 KLOE Collaboration: F. Ambrosino *et al.* *Phys. Lett. B*, 636:173, 2006.
- 236 KLOE Collaboration: F. Ambrosino *et al.* *Phys. Lett. B*, 642:315, 2006.

- 237 KLOE Collaboration: F. Bossi *et al.* *Riv. Nuovo Cim.*, 031(10):531, 2008. arXiv:0811.1929v2 [hep-ex].
- 238 KLOE Collaboration: G. D'Ambrosio and G. Isidori. *J. High Energy Phys.*, 12:011, 2006. arXiv:hep-ex/0610034v1.
- 239 M. Kobayashi and K. Maskawa. *Prog. Theor. Phys.*, 49:652, 1973.
- 240 R.D. Kohaupt and G. Voss. *Annu. Rev. Nucl. Part. Sci.*, 33:67, 1983.
- 241 F.W.J. Koks and J. Van Klinken. *Nucl. Phys. A*, 272:61, 1976.
- 242 E. Komatsu *et al.* *Astrophys. J. Suppl.*, 180:330, 2009.
- 243 C. Kraus *et al.* *Eur. Phys. J. C*, 40:447, 2005.
- 244 KTeV Collaboration: A. Alavi-Harati *et al.* *Phys. Rev. D*, 67:012005, 2003.
- 245 KTeV Collaboration: A. Alavi-Harati *et al.* *Phys. Rev. D*, 70:079904, 2004.
- 246 H. Kühnelt and O. Nachtmann. *Nucl. Phys. B*, 88:41, 1975.
- 247 J. Lach and L. Pondrom. *Annu. Rev. Nucl. Sci.*, 29:203, 1979.
- 248 W.E. Lamb and R.C. Retherford, Jr. *Phys. Rev.*, 72:241, 1947.
- 249 S.K. Lamoreaux. *Phys. Rev. Lett.*, 78:5, 1997.
- 250 D.L. Landau and I.J. Pomeranchuk. *Dokl. Akad. Nauk. SSSR*, 92:535, 735, 1953. English transl.: *The Collected Papers of L.D. Landau*, D. ter Haar, Pergamon Press, Oxford 1965.
- 251 L.R. Landau. In *Proc. LP-HEP Conf. Geneva, July 1991*, S. Hegarty, K. Potter and E. Quercigh, editors, World Scientific, 1992.
- 252 P. Langacker. *Phys. Rep.*, 72:185, 1981.
- 253 P. Langacker and M. Luo. *Phys. Rev. D*, 44:817, 1991.
- 254 Lawrence Radiation Laboratory. *Introduction to the Detection of Nuclear Particles in a Bubble Chamber*. The Ealing Press, Cambridge, MA, 1964.
- 255 T.D. Lee. *Particle Physics and Introduction to Field Theory*. Harwood Academic Publishers, Chur, Switzerland, revised and updated 1st edition, 1988.
- 256 T.D. Lee *et al.* *Phys. Rev.*, 75:905, 1949.
- 257 T.D. Lee and C.S. Wu. *Annu. Rev. Nucl. Sci.*, 16:511, 1966.
- 258 T.D. Lee and C.N. Yang. *Phys. Rev.*, 104:254, 1956.
- 259 Y.K. Lee *et al.* *Phys. Rev. Lett.*, 10:253, 1963.
- 260 H. Lehmann *et al.* *Nuovo Cimento*, 6:319, 1957.
- 261 W.R. Leo. *Techniques for Nuclear and Particle Physics Experiments*. Springer, Berlin, 1987.
- 262 LEP Electroweak working group. *Phys. Rep.*, 427:257–454, 2006.
- 263 H. Leutwyler and M. Roos. *Z. Phys. C*, 25:91, 1984.
- 264 S.J. Lindenbaum and L.C.L. Yuan. *Phys. Rev.*, 111:1380, 1958.
- 265 J. Litt and R. Meunier. *Annu. Rev. Nucl. Part. Sci.*, 23:1, 1973.
- 266 M.S. Livingston and J.P. Blewett. *Particle Accelerators*. McGraw-Hill, New York, 1962.
- 267 P.C. Macq *et al.* *Phys. Rev.*, 112:2061, 1958.
- 268 B.C. Maglić *et al.* *Phys. Rev. Lett.*, 7:178, 1961.
- 269 L. Maiani, editor. *The Second Daphne Physics Handbook: CP and CPT Violation in Neutral Kaon Decays*. INFN, Frascati, 1995.
- 270 Z. Maki. *Prog. Theor. Phys.*, 31:331, 1964.
- 271 Z. Maki *et al.* *Prog. Theor. Phys.*, 23:1174, 1960.
- 272 W.A.W. Mehlhop *et al.* *Phys. Rev.*, 172:1613, 1968.
- 273 W. Meissner and R. Ochsenfeld. *Naturwissenschaften*, 21:787, 1933.
- 274 L. Michel. *Proc. Phys. Soc. A*, 63:514, 1950.
- 275 A.B. Migdal. *Phys. Rev.*, 103:1811, 1956.
- 276 J.P. Miller *et al.* *Rep. Prog. Phys.*, 70:795, 2007.
- 277 V.A. Miransky *et al.* *Phys. Lett. B*, 221:177, 1989.
- 278 A. Misaki. *Nucl. Phys. B – Proc. Suppl.*, 33:192, 1993.
- 279 P.J. Mohr and B.N. Taylor. *Rev. Mod. Phys.*, 77:1, 2005.
- 280 M. Morita. *Beta Decay and Muon Capture*. W.A. Benjamin, New York, 1973.
- 281 Muon ($g - 2$) Collaboration: G.W. Bennett *et al.* *Phys. Rev. D*, 73:072003, 2006.
- 282 NA48: A. Lai *et al.* *Eur. Phys. J. C*, 22:231, 2001.
- 283 O. Nachtmann. *Elementary Particle Physics*. Springer, Berlin, 1989.
- 284 H. Nagaoka. *Philos. Mag.*, 7:445, 1904.

- 285 Y. Nagashima. *Nucl. Phys. A*, 478:265c, 1988.
- 286 T. Nakano and K. Nishijima. *Prog. Theor. Phys.*, 10:581, 1953.
- 287 Y. Nambu. A 'superconductor' model of elementary particles and its consequences, Talk at Purdue (1960). In T. Eguchi and K. Nishijima, editors, *Broken Symmetries, Selected Papers of Y. Nambu*. World Scientific, Singapore, 1995.
- 288 Y. Nambu. *Proc. Int. Conf. on Symmetries and Quark Models*, R. Chand, editor, Gordon and Breach, 1970, p. 269
- 289 Y. Nambu and G. Jona-Lasinio. *Phys. Rev.*, 122:345, 1961.
- 290 Y. Nambu and G. Jona-Lasinio. *Phys. Rev.*, 124:246, 1961.
- 291 B. Naroska. *Phys. Rep.*, 148:67, 1987.
- 292 NASA. <http://map.gsfc.nasa.gov/media/080998/index.html>. (accessed March 2010).
- 293 S.H. Neddermeyer and C.D. Anderson. *Phys. Rev.*, 51:884, 1937.
- 294 Y. Ne'eman. *Nucl. Phys.*, 26:222, 1961.
- 295 New Muon Collaboration (NMC): M. Arneodo *et al.* *Nucl. Phys. B*, 483:3, 1997.
- 296 K. Nishijima. *The Field Theory (in Japanese)* Kinokuniya Shoten, Shinjuku, Tokyo, 1987.
- 297 K. Niu *et al.* *Prog. Theor. Phys.*, 46:1644, 1971.
- 298 M. Oda. *Cosmic Rays*. Shokabo, Tokyo, 1972. In Japanese.
- 299 T. Ohshima. *Prog. Theor. Phys. Suppl.*, 119:5, 93, 1995.
- 300 S. Okubo. *Prog. Theor. Phys.*, 27:949, 1962.
- 301 S. Okubo. *Phys. Lett.*, 5:165, 1963.
- 302 L.B. Okun and I.Ya. Pomeranchuk. *Soviet Phys. - JETP*, 3:307, 1956.
- 303 S. Olariu and I.I. Popescu. *Rev. Mod. Phys.*, 57:339, 1985.
- 304 L. O'RaiFeartaigh. *The Dawning of Gauge Theory*. Princeton University Press, Princeton, NJ, 1997.
- 305 J. Orear *et al.* *Phys. Rev.*, 102:1676, 1956.
- 306 N. Osakabe *et al.* *Phys. Rev A*, 34:815, 1986.
- 307 J.P. Ostriker. *Proc. Natl. Acad. Sci. USA*, 74:1767, 1977.
- 308 E.W. Otten and C. Weinheimer. *Rep. Prog. Phys.*, 71:086201, 2008.
- 309 Ö. Özdemir *et al.* *J. Geophys. Res.*, 100:2193, 1995.
- 310 Pais, A. *Phys. Rev.*, 86:663, 1952.
- 311 Particle Data Group: C. Amsler *et al.* Review of particle physics. *Phys. Lett. B*, 667:1, 2008.
- 312 Particle Data Group: D.E. Groom *et al.* Review of particle physics. *Eur. Phys. J. C*, 15:1, 2000.
- 313 Particle Data Group: R.M. Barnett *et al.* Review of particle physics. *Phys. Rev. D*, 54:1, 1996.
- 314 R. Peccei. *The Strong CP Problem.*, in: CP Violation, Advanced Series on Directions in High Energy Physics, Vol. 3, C. Jarlskog, editor, World Scientific, p. 503, 1989.
- 315 R. Peccei and H.R. Quinn. *Phys. Rev. D*, 16:1791, 1977.
- 316 C. Pellegrini. *Annu. Rev. Nucl. Sci.*, 22:1, 1972.
- 317 M.L. Perl. *Annu. Rev. Nucl. Part. Sci.*, 30:299, 1980.
- 318 M.L. Perl *et al.* *Phys. Rev. Lett.*, 35:1489, 1975.
- 319 S. Perlmutter. *Phys. Today*, April:53, 2003.
- 320 S. Perlmutter *et al.* *Astrophys. J.*, 517:565, 1999.
- 321 I. Peruzzi *et al.* *Phys. Rev. Lett.*, 37:569, 1976.
- 322 M.E. Peskin and D.V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, Boulder, CO, 1995.
- 323 PLUTO Collaboration: C. Berger *et al.* *Z. Phys. C*, 1:343, 1979.
- 324 H.D. Politzer. *Phys. Rev. Lett.*, 30:1346, 1973.
- 325 I.Ya. Pomeranchuk. *Soviet Phys. - JETP*, 3:306, 1956.
- 326 C.Y. Prescott *et al.* *Phys. Lett. B*, 77:347, 1978.
- 327 C.Y. Prescott *et al.* *Phys. Lett. B*, 84:524, 1979.
- 328 PSI (Paul Scherrer Institut). <http://abe.web.psi.ch/accelerators/ringcyc.php> (accessed March 2010).
- 329 C. Rebbi. *Phys. Rep. C*, 12:1, 1974.
- 330 F. Reines and C.L. Cowan, Jr. *Phys. Rev.*, 92:830, 1953.
- 331 A.G. Riess *et al.* *Astronom. J.*, 116:1009, 1998.

- 332 G.D. Rochester and C.C. Butler. *Nature*, 160:855, 1947.
- 333 M. Roos and A. Sirlin. *Nucl. Phys. B*, 29:296, 1971.
- 334 M.E. Rose. *Elementary Theory of Angular Momentum*. John Wiley & Sons, New York, 1957. Republished by Dover, Mineola, NY, 1995.
- 335 A. M. Rossi *et al.* *Nucl. Phys. B*, 84:269, 1975.
- 336 B. Rossi. *High Energy Particles*. Prentice-Hall, Englewood Cliffs, NJ, 1952.
- 337 V.C. Rubin and W.K. Ford. *Astrophys. J.*, 159:379, 1970.
- 338 L.H. Ryder. *Quantum Field Theory*. Cambridge University Press, Cambridge, 2nd edition, 1996.
- 339 S. Sakata. *Prog. Theor. Phys.*, 16:686, 1956.
- 340 M. Sakuda. *Butsuri*, 48:87, 1993.
- 341 M. Sakuda *et al.* *Nucl. Instrum. Methods A*, 311:57, 1992.
- 342 J.J. Sakurai. *Nuovo Cimento*, 7:649, 1958.
- 343 J.R. Sanford. *Annu. Rev. Nucl. Part. Sci.*, 26:151, 1976.
- 344 A. Schenk. *Nucl. Phys. B*, 363:97, 1991.
- 345 F. Schmidt. CORSIKA Shower Images. <http://www.ast.leeds.ac.uk/~fs/showerimages.html> (accessed March 2010), 2008.
- 346 A.L. Sessoms *et al.* *Nucl. Instrum. Methods*, 161:371, 1979.
- 347 P.S. Shawhan. PhD thesis, university of Chicago, December 1999.
- 348 SLAC. <http://www.slac.stanford.edu/spires/topcites/matrix.shtml>. (Accessed March 2010).
- 349 SNO Collaboration: Q.R. Ahmad *et al.* *Phys. Rev. Lett.*, 87:071301, 2001.
- 350 T. Soldner *et al.* *Phys. Lett. B*, 581:49, 2004.
- 351 D.N. Spergel *et al.* *Astrophys. J. Suppl.*, 148:175, 2003.
- 352 J. Steinberger. *Rev. Mod. Phys.*, 61:533, 1989.
- 353 G. Sterman. QCD and jets. Lecture given at TASI 2004. arXiv:hep-ph/0412013v1.
- 354 M.L. Stevenson *et al.* *Phys. Rev.*, 125:687, 1962.
- 355 E.C.G. Sudarshan and R.E. Marshak. In *Proc. Padua Conf. on Mesons and Recently Discovered Particles*, 1957.
- 356 E.C.G. Sudarshan and R.E. Marshak. *Phys. Rev.*, 109:1860, 1958.
- 357 Super-Kamiokande Collaboration: Y. Fukuda *et al.* *Phys. Rev. Lett.*, 81:1562, 1998.
- 358 L. Susskind. *Nuovo Cimento C*, 12:457, 1970.
- 359 G. 't Hooft. *Nucl. Phys. B*, 33:173, 1971.
- 360 G. 't Hooft. *Nucl. Phys. B*, 35:167, 1971.
- 361 G. 't Hooft. *Phys. Rev. Lett.*, 37:8, 1976.
- 362 T2K Collaboration. Letter of Intent: Neutrino Oscillation Experiment at JHF. <http://neutrino.kek.jp/jhfnu/loi/loi.v2.030528.pdf> (accessed March 2010), 2003.
- 363 TASSO Collaboration: R. Brandelik *et al.* *Phys. Lett. B*, 97:453, 1980.
- 364 R.E. Taylor. *Rev. Mod. Phys.*, 63:573, 1991.
- 365 R. Tenchini and C. Verzegnassi. *The Physics of the Z and W Bosons*. World Scientific, Singapore, 2008.
- 366 J.J. Thomson. *Philos. Mag.*, 44:293, 1897.
- 367 A. Tonomura *et al.* *Phys. Rev. Lett.*, 48:1443, 1982.
- 368 A. Tonomura *et al.* *Phys. Rev. Lett.*, 56:792–795, 1986.
- 369 R.D. Tripp. *Annu. Rev. Nucl. Sci.*, 15:325, 1965.
- 370 Y.S. Tsai. *Phys. Rev. D*, 12:3533, 1975.
- 371 Y.S. Tsai. SLAC-pub-2105. <http://www.slac.stanford.edu/cgi-wrap/getdoc/slac-pub-2105.pdf>(accessed March 2010), 1978.
- 372 UA1 Collaboration: G. Arnison *et al.* *Phys. Lett. B*, 122:103, 1983.
- 373 UA1 Collaboration: G. Arnison *et al.* *Phys. Lett. B*, 129:273, 1983.
- 374 UA1 Collaboration: G. Arnison *et al.* *Phys. Lett. B*, 126:398, 1983.
- 375 UA2 Collaboration: P. Bagnaia *et al.* *Phys. Lett. B*, 129:130, 1983.
- 376 R. Utiyama. *Phys. Rev.*, 101:1597, 1956.
- 377 A.M. van Oudenaarden *et al.* *Nature*, 391:768, 1998.
- 378 G. Veneziano. *Nuovo Cimento A*, 57:190, 1968.
- 379 D. Venos *et al.* *Nucl. Instrum. Methods Phys. Res. A*, 560:352, 2006.
- 380 VENUS. *J. Phys. Soc. Japan*, 56:3763, 1987.
- 381 C.F. von Weizsäcker. *Z. Phys. A*, 88:612, 1934.
- 382 S. Weinberg. *Phys. Rev.*, 140:B516, 1965.

- 383 S. Weinberg. *Phys. Rev. Lett.*, 19:1264, 1967.
- 384 W.I. Weisberger. *Phys. Rev.*, 143:1302, 1966.
- 385 V. Weisskopf and E. Wigner. *Z. Phys. A*, 63:54, 1930.
- 386 V. Weisskopf and E. Wigner. *Z. Phys. A*, 65:18, 1930.
- 387 H. Weyl. *Math. Z.*, 2:384, 1918.
- 388 H. Weyl and K. Preuss. *Aksd. Wiss.*, 433, 1918.
- 389 R. Wigmans. *Annu. Rev. Nucl. Part. Sci.*, 41:133, 1991.
- 390 E.J. Williams. *Phys. Rev.*, 45:729, 1934.
- 391 E.T. Worchester. PhD thesis, university of Chicago, December 2007.
- 392 C.S. Wu *et al.* *Phys. Rev.*, 105:1413, 1957.
- 393 T.T. Wu and C.N. Yang. *Phys. Rev. Lett.*, 13:380, 1964.
- 394 T.M. Yan. *Annu. Rev. Nucl. Sci.*, 26:199, 1976.
- 395 T. Yanagida. *Prog. Theor. Phys. B*, 135:66, 1978.
- 396 C.N. Yang and R.N. Mills. *Phys. Rev.*, 96:191, 1954.
- 397 D. Yennie *et al.* *Annals Phys.*, 13:379, 1961.
- 398 C. Zemach. *Phys. Rev.*, 133:B1201, 1964.
- 399 ZEUS Collaboration: S. Chekanov *et al.* *Eur. Phys. J. C*, 21:443, 2001.
- 400 G. Zweig. CERN Reports 8182/TH.401, 8149/TH.412, 1964.
- 401 F. Zwicky. *Helv. Phys. Acta*, 6:110, 1933.

Index

a

absorption coefficient 379
 absorption length 401
 accelerating expansion 798, 801
 action 89
 action principle 89
 active transformation 38
 adiabatic approximation 131, 135
 adiabatic condition 131, 574
 adjoint 69
 adjoint representation 755
 – of the $SU(N)$ 842
 AGS 443
 Aharonov–Bohm effect 748, 751
 allowed transition 557–558
 α_s 528
 Anderson–Neddermeyer 448
 Anderson, C.D. 66
 angular correlation 567
 angular momentum barrier 482, 484
 anomalous magnetic moment, *see also* g-2
 77, 167, 193, 197, 214, 685
 antigravity 126, 800
 antiunitary transformation 242
 Argand diagram 472
 asymmetric collider 366
 asymmetry A_L 643
 asymptotic condition
 – field 131
 – strong 133
 – weak 133
 asymptotic freedom 24, 524, 762, 764
 attenuation length 396
 axial gauge 357, 361
 axion 762, 798

b

B-factory 366
 b quark 539

Baker–Campbell–Hausdorff formula
 225, 252
 bare coupling constant 206
 bare mass 209
 barn 379
 baryon 18
 – decuplet 848
 – multiplets 510
 – number 467, 507
 – octet 848
 – wave function 850
 baryon–lepton symmetry 536
 BCS theory 789
 Bell–Steinberger relation 637, 658
 beta beam 708
 beta decay 555
 beta function (QCD) 763–764
 β function (QCD) 763
 betatron oscillation 375
 Bethe–Bloch formula 384–385
 Bethe–Heitler formula 183, 394
 Bevatron 443, 503–504
 Bjorken scaling 690–692, 699
 Bjorken x 723
 black sphere model 491
 Boltzmann constant 306
 booster 374
 bootstrap model 679, 729
 Born
 – approximation 300
 – series 296
 Bose–Einstein condensate 750, 769
 Bose–Einstein statistics 105
 boson 105
 – weak 258
 bottomonium 528, 539
 box diagram 625
 bra and ket 267
 Breit frame 700, 702

- Breit interaction 552
- Breit–Wigner resonance formula 16
- bremsstrahlung 181, 198, 394
 - inner 451
 - soft 183, 199
- brick wall frame 702
- bubble chamber 403, 466, 503, 710
- c**
- C, *see also* charge conjugation
 - parity
 - of 2-particle system 256
 - of photon 255
 - transformation 252, 635
 - violation 257, 564
- C, P, T transformation 839
- Cabibbo angle 597
- Cabibbo rotation 522, 596–597, 605
- Callan–Gross relation 702, 714
- calorimeter 388, 395, 417
 - e/h compensation 401
 - energy resolution 419
 - sampling- 418
- canonical
 - conjugate 90
 - formalism 267
 - method 282
 - variables 90, 100
- Cartesian coordinates 733, 736, 744
- cascade shower 396
- Casimir effect 122, 149
- Casimir operator 57, 512, 843, 853
 - $SU(3)$ 545
- causality 120, 273
 - condition 273
- CDHS 711
- center of mass frame 136, 228
- central limit theorem 429
- CESR 539
- Chadwick 447
- channel
 - s-, t-, u- 140
- charge
 - independence 261, 454–455
 - symmetry 261, 266, 714
- charge conjugation 79, 252, 265
 - Dirac matrix 818
 - of the Dirac field 254
 - of the vector field 255
 - test 256
- charged current 556
- charm
 - -ed baryon 538
 - -ed meson 537
 - -ed particle 501
 - lifetime of -ed particle 536
 - quantum number 536
 - quark 525
 - spectroscopy 537
- charmonium 527–529
- Cherenkov
 - counter 406
 - differential type 407
 - ring image 407
 - threshold type 407
 - total absorption type 417
 - radiation 390, 392
- Chew–Frautschi plot 496–497, 532
- chi square distribution 438–439
- chiral
 - anomaly 517
 - eigenstate 82, 171–172
 - invariance 585, 772, 781
 - operator 82
 - perturbation theory 762
 - representation 818
 - transformation 586
 - world 773
- chirality 62, 564, 567
 - conservation 84, 172
- Christoffel connection 736, 739, 741, 747
- CKM matrix *see also* KM matrix 600–601
- CKM mixing 554
- Clebsch–Gordan coefficient 458, 512, 835
- cloud chamber 66, 403, 448, 502
- CM frame 228
- Cockroft–Walton 370
- coherence length 770
- coherent scattering 627
- collider detector 422
- collision length 379–380
- color 519, 895
 - charge 6, 20, 24, 509, 539–540
 - counting 541
 - degrees of freedom 18, 519
 - evidence of 520
 - exchange force 541–542, 544, 547
 - factor 545
 - force 4, 6, 540
 - independence 542
 - singlet 533, 544
 - $SU(3)$ 541
- Compton scattering 176
 - path integral method 348, 353
- confinement 6, 18, 24, 521, 529, 533, 761
- conjugate 268

- conjugate momentum 223
- connected diagrams 353
- connection 734–736, 738, 743–744
- conserved charge 96
- conserved isospin current 263
- constituent mass 524
- continuity equation 95, 99
- contravariant vector 42
- cooling 708
- Cooper pair 5, 19, 750, 767, 781
- correlation 429
 - function 300, 304–305
 - n-point 315
 - two-point 300, 318–319
- cosmic microwave background radiation 799
- cosmological constant 799
- Cosmotron 443, 504
- Coulomb gauge 115, 357, 361, 813
- Coulomb potential
 - instantaneous 814
- covariance 48, 429
 - of Dirac equation 71
- covariant
 - formalism 47
 - gauge 115, 119
 - vector 42
- covariant derivative
 - 131, 147, 253, 732, 741–742, 744, 756
- covariant differential operator, *see* covariant derivative 732
- CP
 - constraint 621
 - eigenstate 618
 - invariance 632
 - operator 618
 - $2\pi, 3\pi$ of 619
 - transformation 618
 - transformation property 632
- CP violation 617
 - ϵ_T 655
 - asymmetry A_L 640, 643
 - asymmetry A_S 644
 - decay parameter ϵ_0 (denoted as ϵ), ϵ' 635–636, 640, 644–650, 659
 - decay parameter ϵ_2, ϵ_f 646
 - decay parameter η_{+-}, η_{00} 640
 - value of 642
 - direct 645, 651
 - discovery of 630–631
 - experiments on 664
 - formalism 632
 - in the SM model 675
 - indirect 645, 651–652
 - mixed interference 645
 - mixing parameter $\epsilon, \epsilon_S, \epsilon_L$ 635, 644
 - evaluation of 649
 - phase convention 645
 - models of 673
 - milli-strong 674
 - milliweak 675
 - superweak 675
- CLEAR 622, 664
- CPT
 - constraint 621
 - formalism 632
 - invariance 632
 - foundation of 662
 - theorem 253, 258
 - transformation 258, 840
- CPT violation 656
 - decay parameter B_0, B_2 661
 - decay parameter γ 655, 657
 - mixing parameter δ 635–639, 644, 656–662
- critical density 800
- critical energy 395, 397
- cross section 379
- crossing 140, 174–175
 - symmetry 139, 494
- Curie temperature 774–775
- current–current coupling 556
- current mass 524
- curvature 741, 751, 757–758
 - average 740
 - local 740
 - of space-time 739, 741, 799
 - of the universe 799
 - total 740
- curvilinear coordinate 734, 742–743
- CVC hypothesis 567, 584, 592, 615
 - experimental test of 587
- cyclotron 371
 - frequency 217, 371, 574
 - isochronous 372
 - ring 372
 - sector 372
 - synchro- 372
- d**
- D meson 535–536
- Dalitz decay 451
- Dalitz plot 478
 - for the decay 486
 - of K meson 483
 - of ω 486
 - symmetrical 482

- DAPHNE 622
 - dark energy 19, 122, 798
 - dark matter 19, 797
 - de Broglie wave 25
 - de Broglie wavelength 11
 - decay amplitude 635
 - decomposition
 - of product representations 855
 - decuplet 510
 - Dee 371
 - deep inelastic scattering 679
 - by electron 687
 - by neutrino 712
 - neutrino data 717
 - QCD correction 725–726
 - defect 777
 - deficit angle 740, 750
 - $\Delta I = 1/2$ rule 591, 647
 - violating parameter ω 646
 - Δm 623, 626
 - sign of 623, 629
 - $\Delta S = -\Delta Q$ components 659
 - $\Delta S = \Delta Q$ rule 589, 625
 - Δ function 120, 154
 - delta functional 357
 - $\Delta(1232)$ 457, 469
 - Δ_1 function 120
 - Δ_F , *see* Feynman propagator
 - density effect 386, 400
 - depletion layer 414
 - derivative coupling 487
 - DESY 539
 - determinant
 - of the hermitian operator 291, 293
 - of the Klein–Gordon operator 315
 - DGLAP evolution formula 728
 - diamagnetism
 - perfect 767, 769
 - dichromatic beam 708
 - differential geometry 735–736, 742
 - diffraction pattern 493, 682
 - diffraction scattering 627
 - dimension
 - fifth 747
 - of irreducible representation 851, 856
 - dimensional regularization 195, 825
 - Dirac
 - condition for γ matrices 69
 - equation 62
 - equation, another derivation 811
 - equation, Lorentz-covariant 69
 - equation, plane wave solution 71, 817
 - γ matrix 69, 817
 - matrix 59–60, 817
 - spinor
 - bilinear form 78
 - direct product 47
 - DIS, *see* deep inelastic scattering
 - disconnected diagram 347
 - domain 775
 - double beta decay 87
 - Drell–Yan process 520, 721
 - scaling 723
 - dressed 212
 - drift chamber 412, 415, 417
 - dual tensor 49
 - duality 500

e

 - e^-e^+ to $\mu^-\mu^+$ reaction 174
 - e - μ scattering 167
 - path integral method 351
 - polarized- 171
 - e - μ - τ puzzle, *see* generation puzzle
 - effective action 305
 - effective range 490
 - approximation 490
 - eigen mass in matter 628
 - eightfold way 508
 - Einstein equation 741
 - Einstein–Podolsky–Rosen, *see* EPR paradox 663
 - Einstein’s contraction 41
 - elastic scattering
 - cross section 491
 - electric dipole moment 88
 - of the neutron 249
 - electron
 - magnetic moment 75
 - self-energy 349–350, 830
 - electron holography 751
 - electron volt 364
 - electrostatic spectrometer 574
 - electroweak theory 770
 - emittance 368
 - emulsion 415, 448
 - energy-momentum tensor 96, 225
 - complex field 99
 - Dirac field 100
 - electromagnetic field 99
 - energy vector 426
 - entropy 307, 309
 - ep elastic scattering 683
 - EPR paradox 663
 - equivalent photon 699
 - cross section of 701

- energy of 701
- polarization of 699
- error matrix 435, 438
- η meson 454
- Euclidean 736
 - action 291
 - space 291, 733
 - time 303
- Euler constant 196
- Euler–Lagrange equation 90, 92, 223
- exchange force
 - isospin- 455
 - spin 545–546
- extra dimension 793, 796
- f**
 - Faddeev and Popov’s gauge fixing method 355
 - Faddeev–Popov ghost 354, 359
 - family problem 606
 - Fermi
 - coupling constant 240, 556
 - coupling constant, value of 582
 - golden rule 127
 - interaction 556–557, 559
 - theory 553, 555
 - limitation of 610
 - transition 559–561
 - Fermi–Dirac statistics 105, 113
 - Fermi–Yang model 462
 - fermion 104
 - loop 205, 349
 - mass 781
 - ferromagnetism 774
 - Feynman
 - gauge 116
 - parameter 194
 - path integral 274, 277
 - propagator 147, 151, 318
 - propagator in the Coulomb gauge 815
 - rules 186, 299
 - α 722–723, 725
 - Feynman diagram
 - connected- 353
 - disconnected 347, 354
 - external line 187
 - internal line 187–188
 - irreducible 212
 - loop diagram 188–189
 - reducible 212
 - vertex 187
 - final state interaction 478
 - fine structure 65, 546
 - fine structure constant 23, 35, 215
 - Fitch–Cronin 630
 - flavor 18
 - conservation 554, 579, 598
 - eigenstate 789
 - quark 764
 - $SU(3)$ 221, 509
 - Fock space 112
 - forbidden transition 558
 - force
 - carrier 18–20
 - color exchange 457, 489
 - fundamental 20, 733
 - isospin exchange 457, 487, 489
 - long range 24
 - range 23
 - spin exchange 487, 489
 - tensor 489
 - form factor 681
 - electromagnetic 683
 - examples of 681
 - of KL3 595
 - of the nucleon 683, 686
 - formation zone 391, 393, 399
 - four-fermion interaction 555, 557, 559, 611
 - FP ghost 354
 - Fresnel integral 277
 - Friedman equation 799
 - Froissart bound 493
 - ft value 559
 - functional 279
 - calculus 279
 - derivative 280, 286, 304
 - integral 279
 - Taylor expansion 281
 - vacuum 305
 - fundamental representation 755
 - g**
 - g -2 193, 197, 214
 - g -factor 50, 65, 77
 - G operation 453
 - G parity 453
 - γ matrix
 - trace 173
 - Gamow–Teller transition 240, 559, 561
 - gas electron multiplier, *see* GEM 413
 - gauge
 - condition 93
 - field 732
 - fixing 116, 321, 356
 - invariance 205, 321, 613, 750, 825
 - chiral 587

- principle 731
 - sector 783
 - symmetry 114
 - theory 729
 - non-Abelian 754
 - transformation
 - global 731
 - local 252, 732
 - 2nd kind 732
 - 1st kind 731
 - gauging 732, 745, 747, 755
 - Gaussian distribution 430
 - Gaussian integral 277, 284, 314
 - Gauss's law 796
 - Geiger–Marsden 446
 - Geiger mode 411
 - Gell-Mann matrix 506, 542, 761, 844
 - Gell-Mann–Okubo mass formula 513–514, 551
 - GEM 413
 - general relativity 733, 741, 744–745
 - generating functional 300, 305, 315, 318–319
 - for QED 342, 345
 - of interacting fields 320
 - of the connected Green's function 306
 - of the Dirac field 332
 - of the electromagnetic field 323
 - of the scalar field 312
 - Taylor expansion 319
 - generation 18
 - how many 606
 - puzzle 605, 786–788
 - generator 754
 - of $O(3)$ 40
 - of $SU(N)$ 262
 - of the gauge transformation 252
 - of the group 805, 841–842
 - of the Lorentz group 53–55, 70
 - of the Poincaré group 57
 - of the rotation 54, 264
 - of the transformation 224
 - of the translation 227
 - geodesic line 733–734, 736–737, 743
 - geometrical interpretation of gravity 733
 - GIM mechanism 536, 598, 601
 - Ginzburg–Landau parameter 770
 - Ginzburg–Landau theory 767, 769, 781, 789
 - Glashow–Weinberg–Salam theory 782
 - global symmetry 252
 - gluon 20, 24, 536, 541
 - existence of 720
 - momentum 720
 - Goldstone boson 777–778, 780
 - Goldstone theorem 778
 - goodness of fit 433
 - Gordon equality 193–194, 684, 817
 - grand unified theory 7, 221, 788–789, 791
 - Grassmann number 324
 - complex conjugate 324
 - δ function 328
 - derivative 325
 - Gaussian integral 328, 330
 - integral 326
 - gravitino 795
 - graviton 20, 747, 795
 - Green's function 132, 151–152, 272, 297, 303
 - advanced 132, 153
 - Feynman's 312
 - retarded 132, 153
 - group 803
 - Abelian 803
 - adjoint representation 842, 846
 - conjugate representation 843
 - connectedness 805, 842
 - direct product 843, 847
 - fundamental representation 842, 846
 - generators 805
 - irreducible 512
 - irreducible representation 846
 - Lie group 804
 - nonabelian 803
 - $O(3)$ 805
 - $O(N)$ 804
 - rank of the 843
 - reducible 512, 843
 - representation 804
 - $SL(2, \mathbb{C})$ 809
 - structure constant 805
 - structure constants of $SU(3)$ 844
 - subgroup 804
 - $U(N)$ 804, 841
 - GUT, *see also* grand unified theory 792
 - scale 792
 - GWS theory, *see* Glashow–Weinberg–Salam theory 782
- ## h
- hadron 4, 18, 167
 - hadronic shower 401
 - hadronization 532
 - Hamiltonian formalism 90, 100
 - heavy electron 579
 - Heisenberg
 - equation of motion 100–101, 127
 - picture 101, 127–128
 - uncertainty principle 69

helicity 58, 228
 – conservation 161, 174–175
 – eigenstate 61, 172
 – of electron 569–570
 – of neutrino 576
 – operator 61
 – suppression 566

Helmholtz free energy 306, 309
 – free particle 308
 – harmonic oscillator 308

hierarchy problem 790, 792, 796

Higgs
 – field 779
 – mechanism 6, 778–782, 788
 – particle 6, 19, 782, 788
 – sector 783

Hilbert space 267

hologram 752

horn 706

Hubble 801

Huygens' principle 273

hypercharge 509, 771

– of ν_L , e_L , e_R 772
 – of Higgs 782
 – of leptons 772
 – of quarks 772
 – operator 772

hyperfine structure 64, 546

hypernucleus 483

hyperspace 734, 738–739, 751

i

I spin *see* Isospin

ideal mixing 517

ILC 377

imaginary time 302

impact parameter 445

impulse approximation 693

– conditions for 696

in state 132, 334

incoming state, *see* in state 132

index of refraction 628

inertial frame 743

infinite momentum frame 696

inflation 801

inflation model 792

infrared

– compensation 199
 – divergence 184, 192, 194, 197, 200

instanton 762

interacting fields 131, 147

interaction picture 127

interferogram 752

internal degrees of freedom 17

invariant mass 468

ionization chamber 403, 411

ionization loss 381

ionization mode 410

IOO symmetry 505

isoscalar target 714

isospin 18

– conservation 454
 – current 264, 584–585, 755
 – degree of freedom 457
 – multiplets 260
 – of antiparticles 265
 – operator 754
 – symmetry 260

j

J/ψ 526, 535

– isospin and G-parity 527
 – mass and quantum number 527
 – width 533

Jacobi identity 759

jet 533

– gluon 720

k

K factor 724

K meson 481

Kabir asymmetry 655

Kaluza–Klein theory 747, 795

kaon decay 589

– time development of 638

KK tower 797

KL3 decay 592

– form factor 595

Klein–Gordon equation 25, 32

Klein–Nishina formula 181

KLOE 664

klystron 371

KM matrix 597, 600–601, 617, 676, 791

Kobayashi–Maskawa matrix, *see* KM matrix
 597, 675

KTeV 666

Kurie plot 573

l

LAB frame, *see* laboratory frame

laboratory frame 136

Lagrangian

– complex field 93
 – density 92
 – Dirac field 93
 – electromagnetic field 93
 – formalism 89, 223

- real field 93
- $SU(2)$ 760
- $SU(2) \times U(1)$ 772
- Lamb shift 213
- Λ_{QCD} 766
- Landau distribution 385
- Landé g-factor 217
- Larmor frequency 217
- lattice QCD 761
- law of inertia 733, 744
- law of large numbers 428
- leading log approximation 762–763
- least squares method 438
- LEP 369, 374
- lepton 18
 - flavor 579, 605
 - flavor conservation 606
 - number 554
 - number conservation 571
- Levi-Civita tensor 47, 819
- LHC 364, 368, 374
- Lie algebra 804, 841
 - $O(3)$ 40
 - Poincaré group 57
- Lie group 804
 - compact 805
 - noncompact 805
- light guide 405
- limited streamer mode 411
- linac 370
- linear collider 377
- linear operator 290
- Livingston chart 11
- LLA, *see* leading log approximation 762
- long-range force 5
- longitudinal gluon 764
- Lorentz
 - boost 42, 53, 70
 - condition 322
 - contraction 694
 - gauge 115, 357, 359
 - generators for spin 1/2 particle 70
 - invariance 37, 193, 221, 227
 - operator 53, 70
 - operator for spin 1/2 particle 809
 - scalars 45
 - tensor 47
 - vector 41, 44
- Lorentz invariant
 - flux 144
 - matrix element 134, 169
 - phase space 145
- Lorentz transformation 41, 138
 - homogeneous 41
 - inhomogeneous 56
 - orthochronous 46, 120
 - proper 46, 119–120
- LPM effect 399–400
- LS force 546–547
- luminosity 367
 - measurement 427
- m**
- MAC-E 574–575
- magnetic flux 750, 767
 - line 753
 - quantization 749–750, 753
 - tube 767
- magnetic moment
 - of the baryon 525
 - of the proton 524
- magnetic spectrometer 574
- magnon 778
- Majorana
 - phases 791
- Majorana particle 84
- Maki–Nakagawa–Sakata matrix 791
- Mandelstam variable 138, 140, 171, 366
- mass
 - formula 547, 549, 551
 - of the gauge boson 779
 - of W, Z 783
- mass absorption coefficient 396
- mass hierarchy 786–788
- mass matrix
 - 619–620, 625–626, 628, 633, 636, 859
 - eigenvalue 621
 - symmetry constraint 634
- mass operator 619
- matter field 27
- matter particle 18
- maximum likelihood method 433
- Meissner effect 5, 750, 753, 767–770, 781
- meson 18
 - mixing 516
 - octet 846
- meson exchange
 - scalar 164
- metric 734, 736
- metric tensor 41–42, 734, 736, 747
- Michel parameter 583, 604
- minimum interaction 732
- minimum ionization loss 385
- Minkowski
 - metric 41

- space 41, 291, 303, 734, 744
- time 303
- mirror operation 236
- missing energy 427
- Molière unit 399
- momentum transfer 680
- monopole 792
- Mott scattering 157, 184, 192, 213
 - path integral method 345
- multiple scattering 389
- multiwire proportional chamber 411, 415
- muon 448
 - decay 579, 584
- MWPC, *see* multiwire proportional chamber

n

- n photon emission 204
- n -point function 318
- N product, *see* normal ordered
- NA48 666
- Nambu 775
- Nambu–Goldstone boson 778
- narrow band neutrino beam 707
- natural unit 33
- naturalness 792–793
- NBB, *see* narrow band neutrino beam
- negative energy 26, 64
- negative pressure 126, 801
- negative probability 26
- neutral current 598, 773
 - flavor-changing 601
- neutralino 795, 798
- neutrino 444, 448, 571
 - atmospheric 790
 - beams 705
 - detector 709
 - Dirac neutrino 87
 - experiments how to 705
 - factory 708
 - helicity 566
 - helicity measurement 576
 - Majorana neutrino 86–87, 791
 - mass measurement 572, 576
 - oscillation 18, 35, 62, 554, 606, 789
 - reactor neutrino 571
 - right-handed 86, 789
 - second neutrino 578
 - solar 790
 - Weyl neutrino 87
- neutron 447
 - bottle 250
- Nishijima–Gell-Mann’s law 502, 504, 509, 590, 771, 782

- Noether’s theorem 95, 222, 225, 731
- non-Abelian 754
 - field strength 758
 - gauge theory 100, 195, 753–760
- nonet 507, 510, 522
- normal distribution 430
- normal ordered 149
- November Revolution 525
- ν – q scattering 715
- nucleon
 - parton structure of 680
- number operator 102

o

- $O(3)$ 37, 40
- octet 507, 509–510
 - baryon 509
 - meson 507
- off shell 118, 150
- Ω 514, 519
- ω meson 485
- on shell 150, 207
- order parameter 769, 775
- out state 132, 334
- outgoing state, *see* out state 132
- OZI rule 518, 535–536

p

- P , *see also* parity 559
 - violation 559, 562
- P -type interaction 560, 564, 567
- pair creation 396
- parallel transport 734–736, 738, 740, 742, 756–757
- parity 77
 - intrinsic 233
 - of π^+ 449
 - of π^0 450
 - tensor method 451
 - of the photon 234
 - operation 77
 - transformation 42, 46, 62, 233, 839
 - two-body system 235
- parity violation 239, 559–560, 562
 - in the weak interaction 238
- partial wave
 - analysis 462, 472
 - cross section 466
 - expansion 229
 - parity 464, 470
 - S -matrix unitarity 464
 - scattering matrix 468
- particle democracy 679

- particle identification 419
 - partition function 306
 - parton 680
 - distribution function 697, 725
 - kinematics 694
 - luminosity 723
 - model 693
 - in hadron collisions 721
 - variable 695–696
 - passive transformation 38
 - path integral 267, 311
 - determinant 315
 - expression 276
 - harmonic oscillator 285
 - matrix method 287
 - normalization 315
 - operator method 290
 - scattering matrix 294
 - calculation of 297
 - Pauli 444
 - matrix 61
 - Pauli–Dirac
 - representation 62, 818
 - Pauli–Lubanski operator 57
 - Pauli–Villars regularization 195
 - penetration length 770
 - penguin diagram 677
 - PEP 529
 - phase convention 622, 640
 - phase rotation 708
 - phase shift 464
 - phase stability 376
 - phase transformation 252, 738, 749, 754
 - phase transition 775
 - ϕ factory 622
 - photomultiplier 405
 - photon
 - longitudinal 117, 699
 - scalar 117, 700
 - self-energy 349, 825
 - soft 199
 - transverse 699
 - unphysical 117
 - virtual 118
 - π meson 449
 - π^0 to 2γ 520
 - pion cloud 584
 - pion decay 564
 - constant 520, 565
 - Planck
 - energy 23
 - scale 4, 795
 - plasma frequency 123
 - plasmon 779
 - plastic scintillator 405
 - Poincaré transformation 746
 - Poincaré group 56
 - Poisson distribution 204, 431
 - polarization
 - in scattering 465
 - of nucleon 465
 - sum 117–119, 178–179, 822
 - Pomeranchuk
 - theorem 492, 498
 - trajectory 498
 - Pomeron 498
 - positron 448
 - positronium 527, 529
 - principle of detailed balance 247
 - principle of equivalence 743
 - Proca equation 93
 - propagator, *see also* Feynman propagator 127, 272
 - free particle 282
 - proper time 43
 - proportional mode 411
 - proton decay 792
 - proton synchrotron, *see* PS 370
 - PS 443
- q**
- Q-value 526
 - QCD 7, 20, 760–764
 - background 368–369
 - coupling constant 529, 763–764
 - Lagrangian 761
 - lattice 769
 - potential 528
 - QED 4, 167
 - path integral method 340
 - test 213
 - quantization
 - of fermion 104
 - of fields 89, 105–106, 111
 - of harmonic oscillator 102
 - of the Dirac field 112
 - of the electromagnetic field 114
 - in the Coulomb gauge 814
 - quantum anomaly 791, 795
 - quantum chromodynamics, *see* QCD
 - quantum electrodynamics, *see* QED
 - quark
 - counting 605
 - electric charge 18
 - electric charge by DIS data 718
 - mass 19, 550

- model 501, 509, 787, 841
- potentials of 24, 527
- sea- 716
- valence 716
- quintessence 802

r

- R-parity 795
- R-value 520–521
- radiation length 390, 395, 401
- radiative correction 191, 527
- range 388
- rapidity 51, 72, 725
 - pseudo- 53
- reduction formula
 - Dirac fields 337
 - electromagnetic fields 337
 - LSZ 320, 333, 340
 - scalar fields 333
- regeneration 626
- Regge
 - pole 495
 - trajectory 494, 527, 529
- Reggeon 499
- relativistic
 - cross section 141
 - kinematics 136
 - normalization 142
 - wave equation 25
- renormalizability 784
- renormalizable theory 611
- renormalization 191, 208
 - constant 133
 - group equation 765
 - mass- 208
 - scattering amplitude 211
- renormalized coupling constant 206
- resonance 466
 - amplitude 472
 - formula 469
 - structure 458
- RF 370
- ρ meson 475
- Ricci's curvature tensor 99, 741
- RICH 407
- Riemann's tensor 741
- Rosenbluth formula 685
- rotation
 - group 37
 - invariance
 - 2-body- 227
 - π N scattering 465
 - matrix 174, 229, 833, 835

- orthogonal condition 229
- operator 229
- rotational symmetry 737
- running coupling constant
 - 204, 208, 762–765, 792
- Rutherford 446
 - formula 446
 - model 3
 - scattering 135, 389, 680
 - scattering formula 136

s

- S-matrix, *see also* scattering matrix 134
 - in the path integral 297
- S-matrix
- sagitta 416
- Sakata model 462, 505
- scalar product 44
- scaling 690, 723
 - Bjorken 691
 - limit 691, 698, 713
 - variables 696
- scattering length 490
- scattering matrix 127
 - path integral method 351
 - T_{if} 231
- Schrödinger
 - equation 271, 278
 - picture 101, 128
- scintillation counter 404
- screening 764
 - anti 764
- sea of negative-energy 66
- sea quark 716
- second quantization 112
- see-saw mechanism 791
- short-range force 5
- singlet 507
- SLC 369, 371, 377
- solid state detector 413
- space-time translation 223
- spectator 368
- spectroscopy 443
 - hadron 457
- spherical harmonic function 235, 244, 833
- spin
 - of π^+ 449
 - of π^0 450
 - tensor method 451
 - operator 63
- spin-orbit coupling 551
- spin-statistics 119
- spinor 70
 - composite of 807

- iso 263
- representation 803
- spontaneous symmetry breaking
 - 5, 750, 774–778, 792
- SPS 374
- spurion 592
- SSD 413, 415
- standard deviation 429
- standard model 7, 731
 - beyond the 789–799
- static approximation 488
- statistical mechanics 306
- strange particle 501
 - decay of 589
- strangeness 501–502
 - eigenstate 618
 - oscillation 622
- string model 500
 - of hadron 532–533
- strong force 4
- strong CP problem 762, 798
- structure constant 754, 842
- structure function 40, 688, 699–700
 - definition of 689
 - isoscalar target 714
 - ν 712
 - $\bar{\nu}$ 712
 - of neutrino 712
 - of point particle 689
 - physical interpretation of 699
 - scaling limit 698
- $SU(2)$ 505, 615, 805
 - symmetry 501
 - tensor 510
- $SU(2) \times U(1)$ 783
- $SU(2) \times U(1)$ gauge theory 770, 774
- $SU(3)$ 501, 505, 841, 844, 857
 - color 510
 - flavor 510
 - rank of 506
 - singlet 511, 513
 - structure constant 506
 - symmetry 516
 - tensor 510
- $SU(3)_{\text{color}} \times SU(2)_L \times U(1)_Y$ 731
- $SU(4)$ 537, 857
- $SU(5)$ 792, 857
- $SU(6)$ 522, 857
 - wave function 523
- $SU(6) \times O(3)$ 525
- $SU(N)$ 841
- Sudakov
 - double logarithm 198
 - form factor 204
- sum over history 274
- sum rule 719
- superbeam 708
- superconductivity 781
- superconductor
 - type I 767, 770
 - type II 767, 770
- supercurrent 750, 769
- supergravity 747, 793
- superstring 747
- superstring model 795
- supersymmetry 747, 792
- superweak model 675
- superweak phase 637, 647, 650
 - parameter ϵ 649
- SUSY, *see* supersymmetry 792
- SusyGUT 793
- symmetry
 - continuous 222
 - discrete 233
 - family 606
 - horizontal 606, 791
 - internal 221, 251–252
 - list of 222
 - spacetime symmetry 221
- synchrotron 373–374
 - oscillation 376
 - radiation 369, 376
- t
- T, *see also* time reversal 839
 - constraint 621
 - invariance 560, 595, 632
 - transformation 9, 839
 - violation
 - asymmetry A_T 655
 - evidence of 653
 - parameter ϵ_T 655
- T product 134, 151
- τ – θ puzzle 483, 562
- τ lepton
 - coupling type of 604
 - decay 601
 - mass and spin of 602
 - production of 601
- technicolor 793
- temporal gauge 357, 361
- tensor analysis 851
- tensor method 485
- test of T and CPT invariance 653

- Tevatron 366
- Thomas precession 217–218
- Thomson model 445
- Thomson scattering 381
- Thomson, J.J. 444–445
- three-jet 721
- time evolution operator 130, 133, 272
- time projection chamber 420
- time reversal 46, 240
 - of S-matrix 242
 - of scalar and vector fields 244
 - of the Dirac field 246
 - operation 241
 - quantum mechanics 240
 - test 247
 - Wigner’s 241
- TOF 420
- Tomonaga–Schwinger–Feynman 193, 213
- top condensate 19
- total reflection 46
- TPC *see* time projection chamber
- traces of the γ matrices 818
- transition operator 134
- transition radiation 387, 390, 392
 - counter 408
- translational operator 268
- translational symmetry 737
- tree approximation 191
- triangular diagram 520
- Troitzk–Mainz experiment 573
- two point function 152
- type I supernova 798
- u**
- $U(1)$ 251
- U, V spin 844, 846
- ultra cold neutron 249
- ultraviolet divergence 192–193
- unification 792
- unification of forces 8
- unified theory 611, 770, 794–795
 - a step toward 608
- unitarity 764
 - for kaon decay 620, 637
 - of scattering matrix 231
 - of the CKM matrix 601
- unitary gauge 780
- unitary group
 - special unitary group 262
 - $SU(N)$ 262
- unitary limit 469–470, 610
- unitary spin [see flavor $SU(3)$] 505
- universality 582, 596, 601, 604, 608
- Upsilon 539
- v**
- V–A interaction
 - 240, 564, 566–567, 579, 583, 595, 608
- V particles 502
- vacuum
 - energy 121–122, 348, 776, 801
 - fluctuation 69, 121, 123
 - has negative pressure 126
 - polarization 213, 762, 764
 - properties of 126
- vacuum expectation value 777, 781, 784
- vacuum to vacuum transition 349
- valence quark 716
- variance 429
- vector
 - axial 233
 - polar 233
- Veneziano model 500, 795
- vertex 148, 151, 187
 - correction 191, 213
 - diagram 198
 - function 127, 147
- virtual particle 150
- w**
- W**
 - longitudinal polarization 613
- Ward–Takahashi identity 210
- Watson–Sommerfeld transformation 494
- wavelength shifter 404
- WBB *see* wide band neutrino beam
- weak 18
 - charge 20, 553
 - current 553, 556
 - doublet 554
 - eigenstate 596, 789
 - force 4–5
 - interaction 553
 - magnetic force 553
 - magnetism 588
- weak isospin
 - doublet 554
 - independence 615
 - triplet 554
- weight diagram 508, 846
- Weinberg angle 773
- Weinberg–Salam model 609
- Weisskopf–Wigner approximation 620, 859
- Weizsäcker–Williams
 - approximation 699
 - formula 704–705
- Weyl 729
 - equation 61
 - ordering 275

- particle 58, 62, 85, 172
- representation 62, 82, 818
- spinor 61

Wick

- rotation 291, 302–303, 307
- theorem 319

wide band neutrino beam 707

Wigner–Eckart theorem 458, 513

Wu–Yang’s phase 647

y

Yang and Mills 730

Young diagram 512, 854

Yukawa

- interaction 131, 162, 783
- potential 24, 488

z

zero-point energy 103, 121–122, 126

ZEUS 727