

Kazuhiko Nishijima

Masud Chaichian

Anca Tureanu Editors

# Quantum Field Theory

---

By Academician  
Prof. Kazuhiko Nishijima

*A Classic in  
Theoretical Physics*

 Springer

# Quantum Field Theory

Kazuhiko Nishijima

# Quantum Field Theory

By Academician Prof. Kazuhiko Nishijima -  
A Classic in Theoretical Physics

Masud Chaichian • Anca Tureanu

*Editors*

 Springer

*Author*

Kazuhiko Nishijima  
University of Tokyo  
Tokyo, Japan

*Editors*

Masud Chaichian  
University of Helsinki  
Helsinki, Finland

Anca Tureanu  
University of Helsinki  
Helsinki, Finland

*Translated by*

Yuki Sato  
Tokuyama College  
National Institute of Technology  
Yamaguchi, Japan

ISBN 978-94-024-2189-7      ISBN 978-94-024-2190-3 (eBook)  
<https://doi.org/10.1007/978-94-024-2190-3>

Originally published by Kinokuniya Company Ltd., Tokyo, Japan, 1987

Japanese original edition BA NO RIRON by Kazuhiko Nishijima published in 1987 by Kinokuniya Company Ltd., Tokyo

© Springer Nature B.V. 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Back Cover: Photo of Prof. Nishijima, 2008 (courtesy of late Mrs. Hideko Nishijima)

This Springer imprint is published by the registered company Springer Nature B.V.  
The registered company address is: Van Godewijkstraat 30, 3311 GX Dordrecht, The Netherlands

# Foreword

In the Preface, Professors Masud Chaichian and Anca Tureanu describe in vivid detail how the English version of this book on quantum field theory by the great Japanese physicist Kazuhiko Nishijima evolved as a labor of love on their parts. Nishijima enjoyed visiting Helsinki and Finland (Who doesn't? I certainly do.) regularly and was often seen consulting a physics book in Japanese, of which he was actually the author. (I can tell you from personal experience that authors like to read their own books; I certainly find my books ridiculously clear and easy to understand.) Chaichian and Tureanu urged Nishijima to have the book translated into English. After Nishijima's death in 2009, they found a translator, had it proofread, improved the English, arranged for a publisher, and invited one of the most eminent theoretical physicists of our times, Professor Yoichiro Nambu, to write a Foreword.

After Nambu's untimely death in 2015, they asked me to step in, perhaps because I have also written a textbook on quantum field theory. Any reader qualified to read this book knows full well that I am far from the level of achievements attained by Yoichiro Nambu, for whom I have the greatest admiration as a person and as a physicist. Nevertheless, out of respect for Nishijima, Nambu, Chaichian, and Tureanu, I agreed to do so.

I believe that Dick Feynman said that learning physics is something like painting a house. (In case you have no experience painting houses, one coat is never enough. Only the most shoddy house painters would do that.) John Wheeler told me something similar: one should not imbibe physics like sipping water from a cup, but more like drinking from a fire hydrant. Anybody who tells you that he or she could learn quantum field theory by reading a single textbook is not only kidding you, but also himself or herself. During my four years in graduate school, I attended Julian Schwinger's course on quantum field theory four times, not because I was especially stupid, but because quantum field theory is a profound subject. In this spirit, I strongly urge students, and also physics professors who claim that they understand quantum field theory, to read this book in addition to whatever other books they have read or possibly in parallel with other books they are reading. I

also find that this book treats some topics that are not discussed in detail, or omitted entirely, in more recent books on quantum field theory.

Santa Barbara, CA, USA  
July 2022

Anthony Zee

# Preface to the English Edition

During several years of our collaboration with Prof. Kazuhiko Nishijima, we had noticed that always he was carrying a book in Japanese, which he consulted from time to time on different subjects. Prof. Nishijima used to visit us at the University of Helsinki regularly in the years 1985–2008. During one of his visits, he finally left a copy of that book with us in Helsinki to keep it for him for his future use. Not knowing which book it was, we asked about it and he explained that it was a book on Quantum Field Theory he had written. Upon our question whether it will be translated into English, Prof. Nishijima said that it was his wish to translate the book himself.

After Prof. Nishijima passed away in Tokyo on 15 February 2009, the book and Prof. Nishijima's wish sadly came to our mind and we reached a wishful thought of getting the book by all means translated into English as a tribute to Prof. Nishijima. Prof. Yoichiro Nambu, the 2008 laureate of the Nobel Prize in Physics, knew Prof. Nishijima very well and in the early 1960's was nominating him for the Nobel Prize for the discovery of the strangeness quantum number. Knowing this, we discussed the plan with Prof. Nambu, and he supported the idea with great enthusiasm, mentioning that he knew the book well and that he would like to write a foreword for its English edition. However, he also mentioned about the well-known difficulty of translating from Japanese to any other language of any group of English kind, and that it was for this reason that Nishijima wanted to translate his book himself.

We contacted Springer-Verlag about publishing the translated book and they consulted Prof. Nambu for a review. Prof. Nambu wrote a very strong evaluation, mentioning the transparency of the book and, in several parts, the originality of the presentations of the subject. The next task was to find a person who could translate the book, being a physicist with a good grasp of English, especially as regards the terminology used in the book. It was only thanks to the efforts of Prof. Nambu, who advertised the project in Japan, asking for volunteers, that a translator was found. Eventually, a young researcher in particle physics at Nagoya University, Dr. Yuki Sato, wrote to us mentioning that he would like to take the task merely as a sign of

respect to Prof. Nishijima, one of Japan's greatest physicists who made significant contributions to theoretical physics.

The translation of the book by Dr. Yuki Sato, together with us the undersigned as editors going through it and checking the text and all the formulas again and again, as well as adding the reference list, took a few years and meanwhile Prof. Nambu passed away. Instead, Prof. Anthony Zee from University of California, Santa Barbara, most kindly accepted to write the foreword for the book.

After the translation of the book, another professional native English translator and a physicist by education, Mr. Stephen Lyle, went through the whole translated text and performed a more complete polishing of the text into a fine English language. We express our utmost thanks to Mr. Stephen Lyle for his devotion to perfection. We would like to thank Dr. Ramon Khanna, executive editor, and Ms. Christina Fehling, editorial assistant, at Springer Nature for their kind support and advice during the whole publication process, and also for their unlimited patience with several delays in our preparing the English edition of the book to send for print, with the aim that it will become a genuine tribute to Prof. Nishijima.

During the preparation of the English edition of the book we have received the encouragement of several physicists in Japan, to whom we are most grateful. Our special gratitude goes to Prof. Misao Sasaki from Kyoto University, and Prof. Kei-Ichi Kondo from Chiba University, for their constant support and much help during several years.

Helsinki, Finland  
July 2022

Moshe M. Chaichian  
Anca Tureanu



# Preface of the Author

One of the important concepts in quantum mechanics is the dual particle-wave nature of matter. The same object is treated as a particle in one case and as a wave in another case, which is not the case in classical physics. This was one of the launch pads of quantum mechanics. Its typical example is light: in classical physics it was mainly described by the wave theory, although there existed also Newton's particle theory. However, once Einstein's photon hypothesis was proposed at the beginning of this century,<sup>1</sup> one could no longer claim that light is a wave or a particle. In this way quantum mechanics made its appearance. In its framework, the electron, which is thoroughly treated as a particle in the classical theory, is always described as a wave. With that the following problem comes up: How can we introduce the particle-wave duality in the same theory? Quantum field theory answers this question.

Quantum field theory was developed no later than the birth of quantum mechanics. It has experienced many changes up to the present. In the early stage, studies were entirely restricted to the area of electrodynamics, and turned out to be quite successful. Nevertheless, early quantum field theory reached a deadlock called the *problem of divergences*. However, with the advent of the renormalisation theory after the war, quantum electrodynamics based on perturbation theory reached completion. On the other hand, after pions were created by accelerators, the struggles of field theory started again. The reasons was that the perturbative approach can not be used for strong interactions, and one encountered the difficulty of extending the renormalisation theory to an approximation method beyond the perturbation theory. As a consequence, field theory was stagnant for a while; instead, the *S*-matrix theory based on dispersion relations became pre-eminent. In addition, after hundreds of "elementary" particles were created by accelerators, it became necessary to answer the question: What on earth are fundamental particles? And this is how the quark

---

This is the only Preface written by Prof. Nishijima in Japanese to the 1st edition published in 1987 in Tokyo; the 8th edition with extended chapters was published in 1997 in Tokyo.

<sup>1</sup> "This century" means the twentieth century.

model appeared. On the other hand, from the studies of weak interactions, gauge theories were introduced. This book covers a survey of quantum field theory related to the developments above.

After the emergence of the gauge theory, there have been quite many developments toward new directions. For example, the study of monopoles, solitons, and instantons related to the topological nature of classical gauge theories; lattice gauge theory, in which gauge invariance is given the leading role to the detriment of Lorentz invariance; applications of grand unified theories to cosmology; more recently, supersymmetric theories, supergravity theories and superstring theories, and so on. These theories are developing; each of them is sufficiently vast already as to require an independent book. I myself do not yet understand them very well, so that I have omitted them in this book, considering that more adequate people will write about them.

Although I published “Fundamental Particles” (1963) [1] and “Fields and Particles” (1969) [2] with the W.A. Benjamin Corporation as text books for elementary particle theory and quantum field theory, the contents of those books are by now outdated, and the topics to be emphasized have also changed in time. However, as far as the fundamentally invariable parts are concerned, this book is based on the two books above. Additionally, since I have had a reluctance to write a book in the Mathematical Library in ignorance of mathematics, I have decided to write the book based not on the mathematical logic but on the physical logic, on the advice of Professor Seizo Ito. It has been long time since the request to write this book, but Professor Ito has patiently waited, and I would like to express my appreciation for him.

In the 61 year of the Showa era (1986), at Rakuhoku,

Tokyo, Japan

Kazuhiko Nishijima

# Contents

<b>1</b>	<b>Elementary Particle Theory and Field Theory</b> .....	1
1.1	Classification of Interactions and Yukawa's Theory .....	2
1.2	The Muon as the First Member of the Second Generation .....	4
1.3	Quantum Electrodynamics .....	4
1.4	The Road from Pions to Hadrons .....	5
1.5	Strange Particles as Members of the Second Generation .....	7
1.6	Non-conservation of Parity .....	8
1.7	Second Generation Neutrinos .....	10
1.8	Democratic and Aristocratic Hadrons—The Quark Model .....	11
<b>2</b>	<b>Canonical Formalism and Quantum Mechanics</b> .....	15
2.1	Schrödinger's Picture and Heisenberg's Picture .....	15
2.2	Hamilton's Principle .....	17
2.3	Equivalence Between the Canonical Equations and Lagrange's Equations .....	17
2.4	Equal-Time Canonical Commutation Relations .....	18
<b>3</b>	<b>Quantization of Free Fields</b> .....	21
3.1	Field Theory Based on Canonical Formalism .....	21
3.1.1	Canonical Commutation Relations .....	22
3.1.2	Euler–Lagrange Equations .....	24
3.1.3	Hamiltonian .....	25
3.2	Relativistic Generalization of the Canonical Equations .....	26
3.3	Quantization of the Real Scalar Field .....	27
3.4	Quantization of the Complex Scalar Field .....	32
3.5	Dirac Equation .....	34
3.6	Relativistic Transformations of Dirac's Wave Function .....	38
3.7	Solutions of the Free Dirac Equation .....	45
3.8	Quantization of the Dirac Field .....	50
3.9	Charge Conjugation .....	55
3.10	Quantization of the Complex Vector Field .....	59

<b>4</b>	<b>Invariant Functions and Quantization of Free Fields</b>	65
4.1	Unequal-Time Commutation Relations for Real Scalar Fields	65
4.2	Various Invariant Functions	68
4.3	Unequal-Time Commutation Relations of Free Fields	71
4.4	Generalities of the Quantization of Free Fields	73
<b>5</b>	<b>Indefinite Metric and the Electromagnetic Field</b>	81
5.1	Indefinite Metric	81
5.2	Generalized Eigenstates	84
5.3	Free Electromagnetic Field in the Fermi Gauge	91
5.4	Lorentz Condition and Physical State Space	96
5.5	Free Electromagnetic Field: Generalization of Gauge Choices	99
<b>6</b>	<b>Quantization of Interacting Systems</b>	105
6.1	Tomonaga–Schwinger Equation	105
6.2	Retarded Product Expansion of the Heisenberg Operators	111
6.3	Yang–Feldman Expansion of the Heisenberg Operators	113
6.4	Examples of Interactions	114
<b>7</b>	<b>Symmetries and Conservation Laws</b>	127
7.1	Noether’s Theorem for Point-Particle Systems	127
7.2	Noether’s Theorem in Field Theory	130
7.3	Applications of Noether’s Theorem	132
7.4	Poincaré Invariance	134
7.5	Representations of the Lorentz Group	139
7.6	Spin of a Massless Particle	141
7.7	Pauli–Gürsey Group	144
<b>8</b>	<b><i>S</i>-Matrix</b>	151
8.1	Definition of the <i>S</i> -Matrix	151
8.2	Dyson’s Formula for the <i>S</i> -Matrix	152
8.3	Wick’s Theorem	154
8.4	Feynman Diagrams	161
8.5	Examples of <i>S</i> -Matrix Elements	165
	8.5.1 Compton Scattering	166
	8.5.2 Pion Decay to Muons	167
8.6	Furry’s Theorem	174
8.7	Two-Photon Decays of Neutral Mesons	175
<b>9</b>	<b>Cross-Sections and Decay Widths</b>	181
9.1	Møller’s Formulas	181
9.2	Examples of Cross-Sections and Decay Widths	187
9.3	Inclusive Reactions	192
9.4	Optical Theorem	196
9.5	Three-Body Decays	199

<b>10</b>	<b>Discrete Symmetries</b> .....	209
10.1	Symmetries and Unitary Transformations .....	209
10.2	Parity of Antiparticles .....	212
10.3	Isospin Parity and $G$ -Conjugation .....	215
10.4	Antiunitary Transformations .....	217
10.5	CPT Theorem .....	221
<b>11</b>	<b>Green's Functions</b> .....	227
11.1	Gell-Mann–Low Relation .....	227
11.2	Green's Functions and Their Generating Functionals .....	229
11.3	Different Time-Orderings in the Lagrangian Formalism .....	233
11.4	Matthews' Theorem .....	236
11.5	Example of Matthews' Theorem with Modification .....	241
11.6	Reduction Formula in the Interaction Picture .....	246
11.7	Asymptotic Conditions .....	248
11.8	Unitarity Condition on the Green's Function .....	255
11.9	Retarded Green's Functions .....	258
<b>12</b>	<b>Renormalization Theory</b> .....	263
12.1	Lippmann–Schwinger Equation .....	263
12.2	Renormalized Interaction Picture .....	274
12.3	Mass Renormalization .....	276
12.4	Renormalization of Field Operators .....	279
12.5	Renormalized Propagators .....	282
12.6	Renormalization of Vertex Functions .....	288
12.7	Ward–Takahashi Identity .....	295
12.8	Integral Representation of the Propagator .....	299
12.8.1	Integral Representation .....	300
12.8.2	Self-Energy .....	302
12.8.3	Integral Representation of the Electromagnetic Field Propagator .....	304
12.8.4	Goto–Imamura–Schwinger Term .....	306
<b>13</b>	<b>Classification of Hadrons and Models</b> .....	309
13.1	Unitary Groups .....	309
13.1.1	Representations of a Group .....	310
13.1.2	Direct Product Representation .....	311
13.1.3	Lie Groups .....	311
13.1.4	Orthogonal Group $O(n)$ .....	311
13.1.5	Unitary Group $U(n)$ .....	312
13.1.6	Special Unitary Group $SU(2)$ .....	313
13.2	The Group $SU(3)$ .....	314
13.2.1	Generators of $SU(3)$ .....	317
13.2.2	$I$ -, $U$ -, and $V$ -Spin .....	319
13.2.3	Three-Body Quark Systems .....	320
13.2.4	Mass Formulas .....	324

13.2.5	Baryon Magnetic Moments .....	326
13.2.6	$SU(3)$ -Invariant Interactions .....	328
13.2.7	Casimir Operator .....	328
13.3	Universality of $\pi$ -Meson Decay Interactions .....	330
13.4	Beta-Decay .....	332
13.5	Universality of the Fermi Interaction .....	336
13.6	Quark Model in Weak Interactions .....	341
13.7	Quark Model in Strong Interactions .....	344
13.7.1	Mass Formula .....	345
13.7.2	Magnetic Moments .....	347
13.8	Parton Model .....	348
<b>14</b>	<b>What Is Gauge Theory?</b> .....	<b>353</b>
14.1	Gauge Transformations of the Electromagnetic Field .....	354
14.2	Non-Abelian Gauge Fields .....	356
14.3	Gravitational Field as a Gauge Field .....	358
<b>15</b>	<b>Spontaneous Symmetry Breaking</b> .....	<b>361</b>
15.1	Nambu–Goldstone Particles .....	361
15.2	Sigma Model .....	364
15.3	The Mechanism of Spontaneous Symmetry Breaking .....	368
15.4	Higgs Mechanism .....	371
15.5	Higgs Mechanism with Covariant Gauge Condition .....	378
15.6	Kibble’s Theorem .....	380
15.6.1	Adjoint Representation .....	380
15.6.2	Kibble’s Theorem .....	383
<b>16</b>	<b>Weinberg–Salam Model</b> .....	<b>389</b>
16.1	Weinberg–Salam Model .....	389
16.2	Introducing Fermions .....	393
16.3	GIM Mechanism .....	396
16.4	Anomalous Terms and Generation of Fermions .....	397
16.5	Grand Unified Theory .....	402
<b>17</b>	<b>Path-Integral Quantization Method</b> .....	<b>409</b>
17.1	Quantization of a Point-Particle System .....	409
17.2	Quantization of Fields .....	416
<b>18</b>	<b>Quantization of Gauge Fields Using the Path-Integral Method</b> .....	<b>421</b>
18.1	Quantization of Gauge Fields .....	421
18.1.1	A Method to Specify the Gauge Condition .....	422
18.1.2	The Additional Term Method .....	423
18.2	Quantization of the Electromagnetic Field .....	423
18.2.1	Specifying the Gauge Condition .....	424
18.2.2	The Additional Term Method .....	426
18.2.3	Ward–Takahashi Identity .....	426
18.2.4	Gauge Transformations for Green’s Functions .....	427

18.3	Quantization of Non-Abelian Gauge Fields .....	428
18.3.1	A Method to Specify the Gauge Condition .....	429
18.3.2	The Additional Term Method .....	431
18.3.3	Hermitization of the Lagrangian Density .....	431
18.3.4	Gauge Transformations of Green's Functions .....	432
18.4	Axial Gauge .....	434
18.5	Feynman Rules in the $\alpha$ -Gauge .....	435
<b>19</b>	<b>Becchi–Rouet–Stora Transformations</b> .....	<b>439</b>
19.1	BRS Transformations .....	439
19.2	BRS Charge .....	442
19.3	Another BRS Transformation .....	444
19.4	BRS Identity and Slavnov–Taylor Identity .....	446
19.5	Representations of the BRS Algebra .....	450
19.6	Unitarity of the $S$ -Matrix .....	455
19.7	Representations of the Extended BRS Algebra .....	457
19.8	Representations of BRS Transformations for Auxiliary Fields .....	459
19.9	Representations of BRSNO Algebras .....	462
<b>20</b>	<b>Renormalization Group</b> .....	<b>469</b>
20.1	Renormalization Group for QED .....	470
20.2	Approximate Equations for the Renormalization Group .....	472
20.2.1	Approximation Neglecting Vacuum Polarization .....	472
20.2.2	Approximation Taking into Account Vacuum Polarization .....	474
20.3	Ovsianikov's Equation .....	479
20.4	Linear Equations for the Renormalization Group .....	483
20.5	Callan–Symanzik Equation .....	485
20.6	Homogeneous Callan–Symanzik Equation .....	491
20.7	Renormalization Group for Non-Abelian Gauge Theories .....	495
20.8	Asymptotic Freedom .....	499
20.8.1	Electron–Positron Collision .....	499
20.8.2	Bjorken Scaling Law .....	502
20.9	Gauge Dependence of Green's Functions .....	504
<b>21</b>	<b>Theory of Confinement</b> .....	<b>509</b>
21.1	Gauge Independence of the Confinement Condition .....	510
21.2	Sufficient Condition for Colour Confinement .....	512
21.3	Colour Confinement and Asymptotic Freedom .....	516
<b>22</b>	<b>Anomalous Terms and Dispersion Theory</b> .....	<b>525</b>
22.1	Examples of Indefiniteness and Anomalous Terms .....	526
22.1.1	Vacuum Polarization .....	526
22.1.2	Goto–Imamura–Schwinger Term .....	528
22.1.3	Triangle Anomaly Term .....	528
22.1.4	Trace Anomaly Term .....	531

22.2	Dispersion Theory for Green's Functions .....	533
22.3	Subtractions in Dispersion Relations .....	535
22.4	Heisenberg Operators .....	537
22.5	Subtraction Condition .....	538
22.6	Anomalous Trace Identity .....	543
22.7	Triangle Anomaly Terms .....	548
22.7.1	Renormalization Condition .....	549
22.7.2	Ward–Takahashi Identity for $C_\lambda$ .....	555
22.7.3	Proof of the Adler–Bardeen Theorem Using the Callan–Symanzik Equation .....	555
	<b>Postface</b> .....	559
	<b>References</b> .....	561
	<b>Index</b> .....	569



# Chapter 1

## Elementary Particle Theory and Field Theory



In a broad sense, field theory can be defined as quantum mechanics with infinite degrees of freedom. This broad definition includes condensed matter physics, which treats non-relativistic many-body problems. However, in this textbook, we discuss the formalism of field theory in a narrow sense, namely, in the framework of relativistic quantum mechanics. The objects of field theory in this narrow sense are the so-called *elementary particles*, and the academic discipline that studies the properties of elementary particles is called *elementary particle theory*. In a way, we can thus say that field theory provides a “grammar” for describing the properties of elementary particles. Since the two theories were developed together, they cannot be clearly separated. Elementary particle theory is concerned with the types of elementary particles that can exist and the types of interactions that they undergo. In each case, there is a phenomenological level and a fundamental level. For instance, atoms, which had been thought to be indivisible, in fact consist of atomic nuclei and electrons, while the atomic nuclei themselves consist of protons and neutrons. Thus, the solution to the problem of what the truly fundamental particles are has changed over time. Correspondingly, the solution to the problem of what the fundamental interactions are has also changed. We do not know if there will ever be an end to the changing viewpoints. However, what has been established so far is that particles called *quarks* and *leptons* seem to be the fundamental particles, while the fundamental interactions are assumed to be gauge interactions. We will discuss later how concepts such as quarks and gauge interactions arised. For now, we shall just give an overview of the developments of elementary particle theory at the phenomenological level.

The concept of field is a generic one used to refer to dynamical variables defined at each point in space. Examples of fields in macroscopic physics are the temperature and the product of the density and the stream velocity of a fluid. In microphysics, only very few such dynamical variables actually survive at the classical level, and perhaps the only candidate is the electromagnetic field. After the birth of quantum mechanics a crucial question was how to apply quantum

mechanics to the electromagnetic field. As mentioned in the foreword, although light is described classically as an electromagnetic wave, at the microscopic level, it also exhibits particle-like features, as expressed in Einstein's photon hypothesis. Therefore, it is essential that the quantum mechanical description of the electromagnetic field should contain such particle-like characteristics. This line of research was pursued by Dirac in [3]. In 1928, Dirac [4] also rewrote the non-relativistic Schrödinger equation as a relativistic equation. This led to the development of quantum electrodynamics. In 1928, Heisenberg and Pauli [5, 6] formulated quantum electrodynamics in a systematic way, and this provided the starting point for field theory.

Dirac's relativistic theory of electrons predicted the existence of a positron, the antiparticle of the electron. Its existence was confirmed by Anderson in 1932 [7]. Similarly, the existence of the anti-proton, the antiparticle of the proton, was also expected. However, we had to wait until 1955 for confirmation of its existence [8]. On the other hand, again in nuclear physics, the discovery of the neutron by Chadwick in [9, 10] was of fundamental importance, leading to the picture of atomic nuclei consisting of protons and neutrons. Yet, the *birth of elementary particle theory* originated in Yukawa's meson theory.

## 1.1 Classification of Interactions and Yukawa's Theory

Current theories such as grand unified theories or the theory of supergravity start by recognizing that there exist a variety of interactions with different strengths. Among these interactions, the best known, apart from gravity, is the electromagnetic interaction which, as mentioned before, has been studied since the birth of quantum mechanics. However, from the early days it was known that there were lots of phenomena that could not be explained by the electromagnetic interaction alone. Beta-decay is an example. The lifetime of a radioactive atomic nucleus is longer than that of states excited electromagnetically. This implies that the interaction causing beta-decay is much weaker than the electromagnetic interaction. In fact, in order to explain the continuous energy spectrum of electrons emitted in beta-decay, Pauli introduced the neutrino hypothesis [11, 12]. Pauli had a very convenient excuse for proposing such a new particle, namely, saving the conservation laws of energy, momentum, and angular momentum.

Fermi substantiated Pauli's suggestion by proposing the *four-fermion interaction* [13] to describe beta-decay in field theory. This was the second interaction discovered after the electromagnetic interaction. It meant that field theory now had to deal with electromagnetic and weak interactions.

After the detection of the neutron in 1932 [9, 10], it gradually became clear that an atomic nucleus consists of protons and neutrons. It was now essential to understand the nature of the so-called *nuclear force* binding protons and neutrons to form atomic nuclei. Since the neutron is electrically neutral, the nuclear force cannot

**Table 1.1** Elementary particles and their interactions

Elementary particles	Interactions
$\nu$	Weak
e	Weak + electromagnetic
n, p	Weak + electromagnetic + strong
$\gamma$	Electromagnetic (carrier)

be the electromagnetic force. Just after the proposal of Fermi's theory, Ivanenko and Tamm calculated the nuclear force by assuming that it was produced by exchange of an electron–neutrino pair via Fermi's interaction [14–16]. It turned out that the calculated force was too weak to be at the origin of the nuclear force.

Consequently, the nuclear force had to be of an entirely different kind from the electromagnetic and the weak interactions. It was also understood that various particles have their own roles and take part only in certain interactions. This is shown in Table 1.1. The symbols for the fundamental particles are  $\nu$  (neutrino), e (electron), n (neutron), p (proton), and  $\gamma$  (photon), while “weak”, “electromagnetic”, and “strong” stand for the weak interaction, the electromagnetic interaction, and the strong interaction (the nuclear force), respectively. The proton and the neutron are collectively dubbed *nucleons* in the sense of elements of an atomic nucleus.

In 1935, Yukawa put forward the *meson theory*, reflecting the principle that *all forces are mediated by fields* [17]. Since the electromagnetic interaction is mediated by the electromagnetic field, the strong interaction and the weak interaction must be mediated by fields, too. Hence, Yukawa introduced the *meson field* as the mediator of the strong interaction. Just as quantizing the electromagnetic field yields the light quantum (photon), it turns out that quantizing any field yields corresponding quanta describing their particle-like aspects. The quantum corresponding to the meson field is an elementary particle called a *meson*. One of Yukawa's most important contributions is that he clarified the relationship between the mass of a quantum and the range of action of a force. In general, the range of a force mediated by some field is given by the Compton wavelength of its quantum. For instance, the range of the electromagnetic interaction mediated by a massless photon is infinite. Since the quantum of the meson field mediates the nuclear force with a range of around  $10^{-13}$  cm, the meson mass must be about 200 times heavier than the electron mass. This mass is midway between the nucleon and electron masses, hence the name meson. Moreover, from a phenomenological standpoint, the quantum of the field mediating the weak interaction described by Fermi's model at a single point must be very massive. Yukawa thought that both the strong nuclear force and the weak interaction might be mediated by the meson field. Shortly afterwards, Klein suggested that a different field should be introduced for the weak interaction [18]. In any case, at that time, such ideas were mere hypotheses. In order to prove them, mesons had to be produced. However, since no accelerator had enough energy to create them, the only way to detect them was by studying cosmic rays. Historically, Yukawa's particle was discovered after some bumps and detours.

## 1.2 The Muon as the First Member of the Second Generation

Theorists predict the existence of experimentally unknown particles, while experimentalists detect sometimes particles which theorists have not predicted. The relation between Yukawa's particle and the muon can be described in this way. Using a modern term, we can say that the world surrounding us consists of elementary particles of the first generation. This is because elementary particles of the second generation are in general unstable, except for the neutrino. Thus, the discoveries of elementary particles of the second generation always gave rise to confusion. Elementary particles such as the muon, the strange quark, and the charm quark are in the second generation. The muon was discovered first.

In 1937, a particle with mass about 100 MeV, the muon, was detected in cosmic rays [19]. For almost a decade afterwards, this particle was thought to be Yukawa's particle.

In 1945, Conversi, Pancini, and Piccioni found that the interaction between the muon and an atomic nucleus was almost as strong as the interaction involved in the muon's weak decay [20], whence they were of a similar (weak) nature. One consequence of this discovery was that the muon and Yukawa's particle, the pion, had to be different particles. One year later, Powell's group observed the following two-step decay process [21] on a photographic plate:

$$\pi^+ \rightarrow \mu^+ + \nu, \quad \mu^+ \rightarrow e^+ + \nu + \nu.$$

Here  $\pi$ ,  $\mu$ , and  $\nu$  stand for the pion, the muon, and the neutrino, respectively. Although there are several different neutrinos, we do not distinguish them here.

A two-meson theory which admits the existence of two different mesons, the pion and the muon, was developed by Sakata, Inoue, and Tanigawa in [22, 23]. One year later, it was also proposed in the United States by Marshak and Bethe [24].

## 1.3 Quantum Electrodynamics

After the Second World War, quantum electrodynamics made amazing progress. Quantum electrodynamics, or QED, is the most typical field theory. By introducing a new weapon called *renormalization*, it became possible to calculate, within the framework of QED, the Lamb shift in a hydrogen atom, i.e., the energy shift between the  $2s$  and  $2p$  states, and the anomalous magnetic moment of the electron. The formulation of the covariant perturbative theory by Tomonaga, Schwinger, Feynman, Dyson, and others was instrumental in this success. At this point, QED achieved the position of being the most successful discipline in exact science. Moreover, the notion of renormalizability was closely connected to subsequent developments in field theory, and in particular to the discovery of gauge theories, as the most important guiding principle in field theories.

## 1.4 The Road from Pions to Hadrons

In 1948, in Berkeley, pions were created artificially for the first time [25]. Furthermore, in 1952, Fermi's group found a resonance state of the pion–nucleon system, which they named  $\Delta$  [26]. This resonance appears in the reaction



where  $N$  stands for a nucleon. The existence of resonances is characteristic of strong interactions of the pion–nucleon system at low energies.

It soon became clear that the covariant perturbative theory which was such a useful tool in QED had almost no utility for understanding strong interactions. This made theorists suspicious of the usefulness of field theory for strong interactions. In fact, determinations of the pion's spin or parity were based on a general invariance principle, and the covariant perturbative theory was completely useless for this purpose. Of course, a non-perturbative method was developed in field theory, but it was not successful because the prescriptions proposed by renormalization were intimately connected to the perturbation theory.

This situation meant that many preferred to investigate the pion–nucleon system without considering the details of strong interactions, but considering this system in its own right. One thing that came up was the idea of *isospin*. The concept of isospin invariance was introduced by Kemmer in 1938 to make a connection between the charge independence of the nuclear force and Yukawa's theory [27]. The isospin transformation is a symmetry in a virtual isospin space. It became an active field of study again when pions could be created in the laboratory. This led to the first two examples of a modern term, *flavour*.

Another idea that came up was the dispersion relation, which is deeply connected to causality in field theory. This gives the relation between the absorptive (imaginary) part of a scattering amplitude and its dispersive (real) part. The form of the dispersion relation does not depend on the details of the interactions. The important factors for dispersion relations are just the representations of spacetime and internal symmetries carried by the particle states, together with the type of mass spectrum the dynamical system possesses. However, the dispersion relations leave slight signatures of the original interactions. For instance, although the best known dispersion relation for the forward scattering of the pion–nucleon system includes terms with first-order poles with respect to some Lorentz invariant variables, the poles can be described by the renormalized coupling constant of Yukawa's interaction. In addition, the dispersion relation for electromagnetic form factors is given by the charges of particles adjusted by so-called *subtraction constants*. In each case, these constants are renormalized coupling constants in field theory.

Much can be learned from the dispersion relation method. First, comparing with perturbative calculations in field theory, we can know what type of renormalization is carried out perturbatively. Considering a given  $S$ -matrix element (a scattering amplitude, for example) at some order in perturbation theory, its absorption part

can be expressed in terms of lower order  $S$ -matrix elements because of the unitarity of the  $S$ -matrix. In this evaluation, integrals have to be carried out in a finite phase-space, but divergences do not appear here. Next, we obtain the dispersion part of the  $S$ -matrix element from its absorption part using the dispersion formula. This corresponds to carrying out a *Hilbert transformation*, and this integral sometimes diverges. Such a divergence just corresponds to the one in field theory. If a divergence exists, we can remove it using a dispersion formula in which a subtraction has been made. These subtractions correspond to the renormalization in field theory, and the subtraction constant corresponds to the renormalized coupling constant. Thus in the dispersion theory, the divergence appears only in the dispersion part, not in the absorption part. Furthermore, the validity of the subtraction is not restricted to a perturbative approach.

We may ask ourselves whether the dynamics of elementary particles can be formulated in terms of the  $S$ -matrix, that is, whether or not we can determine the  $S$ -matrix by combining the unitarity and the dispersion relations of the  $S$ -matrix as mentioned above. In this case, the renormalization is automatically processed by introducing the dispersion formula in which subtractions have been made, thereby realizing a renormalization procedure which is independent of perturbation theory. In this formalism, stable particles which cause poles to occur in scattering amplitudes give the same dispersion relations no matter whether they are elementary particles or composite particles. In other words, what is important in the  $S$ -matrix theory is only the stability of particles, so it does not matter whether the particles are elementary or composite. This ended up having a great significance on subsequent developments in the elementary particle theory. The theory proposed by Chew and Low in 1956 is a typical example of the non-relativistic  $S$ -matrix theory [28]. The idea was developed further by the discovery of the Mandelstam representation in 1958 [29]. In this representation, scattering amplitudes can be written as double dispersion integrations, and they have crossing symmetries (we shall not touch on such symmetries here, but they will be explained later).

Note, however, that the dynamical  $S$ -matrix theory encounters two difficulties. The first is the impossibility or extreme difficulty in finding the complete set of dispersion formulas. The second difficulty is that the solution cannot be determined unambiguously, even under the approximation where we neglect many-particle states. This is related to the fact that we cannot tell the difference between elementary particles and composite particles in the  $S$ -matrix theory, as mentioned before. For example, for a given pair of dispersion formulas we can adopt the Lagrangian of Yukawa's theory or the Lagrangian of quantum chromodynamics (QCD).

The  $S$ -matrix theory itself, however, went in a different direction. From the study of the high-energy behaviour of scattering amplitudes with crossing symmetries, the Regge pole theory, or Regge trajectory theory, was formulated in 1960s [30]. In the 1950s, following the discovery of the antiproton, many hadrons were detected at the Bevatron using bubble chambers, and the Regge trajectory was applied to classify those hadrons. "Hadron" is the generic name for elementary particles participating in strong interactions. When we consider a two-body bound state, the

Regge trajectory is obtained by plotting the relation between the angular momentum  $J$  extended to a continuous variable and the mass squared of the state,  $M^2$ . In this method, allowing many hadrons with different quantum numbers except for their spins to follow the same trajectory clarifies the dynamical family relationships among them, and this of course can help to classify them. In particular, Chew and Frautschi succeeded in classifying many hadrons by considering that all the Regge trajectories are continuous lines with universal tangents [31]. As mentioned before, the Regge trajectory corresponds to a bound state, so the success of the Chew–Frautschi plot implied that all hadrons were composite rather than fundamental. The way we interpret this fact is a problem of fundamental importance for elementary particle theory, and we will discuss it in great depth later.

## 1.5 Strange Particles as Members of the Second Generation

Elementary particles which obey Fermi statistics are classified into hadrons, which engage in strong interactions, and leptons, which do not. The muon, which belongs to the second generation, is a lepton. We have already mentioned its detection and the confusion this created. Similarly, when second generation hadrons were detected, they caused the same confusion. For this reason, hadrons belonging to the second generation were first called *strange particles*. In connection with the detection of the muon, we mentioned that experimentalists sometimes detect particles that have not been predicted by theorists, and this is indeed what happened in the case of strange particles.

In 1947, Rochester and Butler detected two V-particles in a cloud chamber [32]. The V-particle was named after the shape of the track in the cloud chamber, indicating a particle decay. Although the observations continued for two more years, no new particle was discovered. In such a case, two things can happen: either we find nothing when we carry out additional tests or we suddenly detect many cases by improving experimental methods. In the case of V-particles or strange particles, the latter happened. The improved method was to detect V-particles from cosmic rays by climbing high mountains. In the early 1950s, experimental groups in Pasadena and Manchester observed dozens of V-particles on the White Mountains [33] and at the Pic du Midi in the Pyrenees [34], respectively. The biggest problem with the V-particle was to reconcile its frequent detections with its relatively long lifetime of about  $10^{-10}$  seconds (long for an elementary particle). This problem was somewhat similar to the muon problem. With hindsight, this was because both particles belonged to the second generation. Many new particles were subsequently detected in cloud chambers and on photographic plates. For example, *hyperons* like  $\Lambda^0$ ,  $\Sigma^\pm$ , and  $\Theta^-$ , which are hadrons satisfying Fermi statistics, heavy K-mesons (also called  $\theta$ - or  $\tau$ -mesons), and so on.

In order to explain the discrepancy between the frequent detections and the long lifetime, it was suggested that V-particles or strange particles might be created in pairs. Such theories were proposed in 1951 by Nambu, Yamaguchi, and the author

[35, 36], by Oneda [37], and by Miyazawa [38], then by Pais in [39]. However, the pair-production theories were hard to demonstrate from observations of cosmic rays. What these models had in common was that strange particles were always produced in pairs through strong interactions, and the isolated strange particles decayed through weak interactions. This explained both the frequent detections and the long lifetime. However, some phenomena could not be explained by this hypothesis alone. For example, it was not understood why the decay processes  $\Sigma^- \rightarrow \Lambda^0 + \pi^-$  and  $\Sigma^- \rightarrow n + \pi^-$  did not occur through strong interactions. Nor indeed why the heavy mesons with positive charges were more common than those with negative charges both on plates and in cloud chambers. These questions could not be answered using the above hypothesis.

In 1953, the pair production of strange particles that had not been confirmed by cosmic rays was verified experimentally using an accelerator, the Cosmotron, at Brookhaven [40]. The concept of *strangeness* was introduced by Nakano and the author [41], and independently by Gell-Mann [42]. This concept is closely related to isospin. All the members of a given isospin multiplet have the same value of the strangeness. Further, strangeness is an additive quantum number. It is conserved in strong and electromagnetic interactions, but not in weak interactions. As we shall see later, parity is a quantum number with the same features. Moreover, the K-meson came to play an extremely important role in elementary particle theory. In the history of modern physics, we may say that the K-meson has played the most important role after the hydrogen atom. The concept of strangeness is closely related to the fact that the neutral K-meson, denoted by  $K^0$ , differs from its antiparticle  $\bar{K}^0$ . Furthermore, the non-conservation of parity in weak interactions was discovered through the diversity of K-meson decay patterns. The fact that the CP transformation, where C stands for charge conjugation and P for parity, is not conserved in weak interactions was discovered through the observation of certain decay patterns of the neutral K-mesons. As can be seen from this, the K-meson continued to provide many new ideas. To introduce a modern term, strangeness is a third example of flavour and it is carried by hadrons of the second generation. We shall see later that the symmetry group  $SU(3)$  was proposed as a way of combining the isospin symmetry group  $SU(2)$  with strangeness.

## 1.6 Non-conservation of Parity

The strange K-meson was given different names according to its decay patterns:

$$\theta \rightarrow 2\pi, \quad \tau \rightarrow 3\pi. \quad (1.1)$$

The question was whether  $\theta$  and  $\tau$  were the same particle or not. The improvement in the accuracy of measurements revealed that the masses and lifetimes of  $\theta$  and  $\tau$  coincided. In addition, it became clear that the two production rates were constant, and quite independent of the production process. This suggested that  $\theta$  and  $\tau$  were



the same type of particle. However, from the analyses by Dalitz [43] and Fabri [44], this implied the non-conservation of parity in decay processes. The discrepancy between the identification of the particles and the conservation of parity was called the  $\theta$ - $\tau$  puzzle. Under the circumstances, given that the non-conservation of parity might occur, not only in K-meson decays, but in all weak interactions, Lee and Yang proposed various methods of experimental validation [45]. The  $^{60}\text{Co}$  experiment by Wu's group was the first to confirm the non-conservation of parity in beta decays [46]. In particular, it was shown that parity was maximally broken in decay processes which involved neutrinos. This led to the two-component theory of the neutrino put forward by Lee and Yang [47]. Since parity is a space-time symmetry, in contrast to strangeness, many physicists were shocked by the non-conservation of parity.

After this discovery, theoretical research was very active. Here we mention the most important development. Assuming that all forces are mediated by fields, we can determine the characteristics of the fields that mediate the weak interactions by identifying the specific types of phenomenological Fermi interaction. An important first step in this direction was the  $V - A$  theory proposed in 1958 [48–50]. This theory treated the fields mediating weak interactions as vector fields. The quanta of these vector fields were called W bosons, where W here stands for “weak”.

In this connection, in the late 1940s, the conserved vector current (CVC) hypothesis was put forward [49], based on the universality of the Fermi interaction. The idea was proposed by Marshak and Sudarshan [48], by Sakurai [50], and by Gerstein and Zeldovich [51], and further promoted by Gell-Mann and Feynman [49]. According to experiments, the coupling constants for vector couplings of Fermi interactions for various processes are the same, irrespective of the processes, in spite of corrections from strong interactions. This can be explained by assuming that a four-vector density expressing hadronic parts of the Fermi interaction is proportional to an isospin four-current density. This is the *CVC hypothesis*, which will be explained in more detail in the text. When describing phenomenological Fermi interactions as products of two Yukawa interactions mediated by W bosons, each Yukawa interaction is described as the product of the conserved isospin current density and the W field. It is suggestive to compare with the fact that the electromagnetic interaction is the product of the conserved current density and the four-potential of the electromagnetic field.

The idea of including the axial vector current, corresponding to  $A$  in  $V - A$ , developed into the *current algebra method* [52, 53], inspired by the fact that in quantum mechanics a set of conserved quantities forms an algebra. In Sect. 1.4, we presented the methods based on the dispersion formulas and on symmetries for dealing with the strong interactions, although without going into the details. The current algebra method is an extension of the latter. The three components of isospin satisfy the same commutation relations as the components of angular momentum, and they generate the  $SU(2)$  algebra. Although this is the algebra formed by just the spatial integral of the time component of the  $V$  part of  $V - A$ , i.e., the algebra satisfied by isospin, if we take  $V - A$  itself, we also get the  $SU(2)$  algebra. Since the  $V - A$  current corresponds to the left-handed current, which will be explained later,

we write this algebra as  $SU(2)_L$ . Introducing the strangeness-changing current, this algebra is extended to  $SU(3)_L$ . Furthermore, introducing  $SU(3)_R$  in the context of the  $V - A$  theory, the current algebra is extended to  $SU(3)_L \times SU(3)_R$ . The first successful application of this algebra was the Adler–Weisberger formula derived in 1965 [54–56]. Then, in 1963, Cabbibo extended the concept of the universal Fermi interaction to the strangeness-changing interaction [57], and this was an important step on the path to the quark model.

## 1.7 Second Generation Neutrinos

The neutrino was first introduced by Pauli in 1933 [11, 12], and immediately used by Fermi to formulate the field theory of beta-decay [13]. The existence of this neutrino was demonstrated by Cowan, Reines, et al. in [58]. This was the first generation neutrino.

Research on several processes involving weak interactions suggested that there might be a lepton number conservation law and a corresponding selection rule. However, two different definitions were given for the lepton number. One taking  $e^-$ ,  $\mu^+$ , and  $\nu$  as leptons was given by Konopinski and Mahmoud in [59], and another taking  $e^-$ ,  $\mu^-$ , and the left-handed (anti)neutrino as leptons was given by Lee and Yang in [47]. It seemed that neither definition was in conflict with experiment. Then, in 1957, assuming a four-component neutrino and putting the case that each conservation law holds, the author set out to investigate which theories were consistent [60]. Combining the above two types of conservation law in the right way, it turned out that the electron family and the muon family could be attributed separate conservation laws. In the electron family, the difference between the numbers of  $e^-$  and  $\nu_L$  and the numbers of  $e^+$  and  $\bar{\nu}_R$  was conserved, and likewise in the muon family for the difference between the numbers of  $\mu^-$  and  $\bar{\nu}_L$  and the numbers of  $\mu^+$  and  $\nu_R$ .

In short, there were two types of lepton number conservation law: one concerned the lepton number of the first generation and the other concerned the lepton number of the second generation. The subscripts L and R stand for left-handed and the right-handed, respectively, and both are two-component neutrinos. There are thus two types of two-component neutrino: one belongs to the electron family in the first generation and the other to the muon family in the second generation. A similar idea was also introduced by Schwinger [61].

The second generation neutrino hypothesis was confirmed by a group at Columbia University in 1962 using the AGS accelerator in Brookhaven [62].

## 1.8 Democratic and Aristocratic Hadrons—The Quark Model

In 1948, Fermi and Yang constructed a theory treating the pion as a bound state of a nucleon and an anti-nucleon [63]. They were motivated by the idea that, since there were so many so-called elementary particles, it would be better to consider some of them as composite particles. In 1955, including the strange  $\Lambda$  particle with a pair of fundamental particles, the proton and the neutron, Sakata extended the Fermi–Yang model for hadrons [64]. He put forward the idea that all hadrons were bound states of protons, neutrons, the  $\Lambda$  particles, and their antiparticles. Given that there might exist a symmetry between p, n, and  $\Lambda$ , Ikeda, Ogawa, and Ohnuki introduced the symmetry group  $SU(3)$ , and studied the types of representations the hadrons belonged to [65]. For mesons and the hadrons satisfying Bose statistics, Sakata’s model achieved many successful results. The  $\eta$  meson found in 1961 [66] was one successful example of Sakata’s predictions. However, his model could not give a good explanation for baryons. In any case, this was just one attempt to interpret hundreds of hadrons as bound states of the most fundamental particles.

In 1961, Gell-Mann [52, 67] and Ne’eman [68] independently extracted the concept of  $SU(3)$  symmetry from Sakata’s model. In Sakata’s model, p, n, and  $\Lambda$  were assigned to the three-dimensional representation of  $SU(3)$ . On the other hand, including  $\Sigma$  and  $\Theta$  with the particles in Sakata’s model, they were able to assign these particles to the eight-dimensional representation. This was subsequently called the *eighfold way*. For mesons, there is no qualitative difference between Sakata’s model and the eighfold way, because mesons constitute the eight-dimensional representation in both models.

The  $SU(3)$  symmetry is not exact, but explicitly broken. This led to the relation called the *Gell-Mann–Okubo formula* [69, 70] among the masses of particles in one irreducible representation. The highlight of this development was the observation of  $\Omega^-$  with a mass of 1972 MeV, while the mass predicted by the mass formula was 1983 MeV [71].

Let us return to the question of how to interpret the large number of hadrons discussed in Sect. 1.4. From the phenomenological point of view, hadrons can be categorized by the Chew–Frautschi plot [72] and the  $SU(3)$  group. The success of this classification method implied that all the observed hadrons were composite particles. There were two possible ways to interpret this outcome: one was that no fundamental particle existed other than the already observed hadrons, and the other was that more fundamental particles existed, but that they had not yet been observed. These were referred to as *democratic hadrons* and *aristocratic hadrons*, respectively.

The former idea was put forward in Berkeley around 1960, especially by Chew’s group [73]. In modern physics, there is a tendency always to seek more fundamental levels, such as molecules  $\rightarrow$  atoms  $\rightarrow$  nucleons  $\rightarrow$  elementary particles, but Chew’s group suggested cutting off this sequence. According to this idea, all hadrons should be treated on an equal footing. In other words, since no hadron is more fundamental

than the others, each hadron from the lightest pion to the heaviest nucleon should be thought of as a bound state of other hadrons. In this case, each hadron's mass and spin should be determined by a self-consistent method. In this approach, there is no idea of a most fundamental field, so field theory is not suitable to describe it. In this sense, it is a pure  $S$ -matrix theory. In particular, the dispersion formula is suitable for this purpose because it holds whether the relevant particles are elementary or composite. However, a framework describing the complete dynamics in a mathematically rigorous way was never provided. This was the idea of a democracy of hadrons.

The other idea is that detectable hadrons are composite particles of more fundamental particles. We can say that this is the aristocracy of the hadron world, treating the more fundamental particles as aristocrats. In 1964, Gell-Mann and Zweig proposed the quark model for hadrons based on the  $SU(3)$  group [74, 75]. In this model, quarks became the most fundamental particles. This can be considered as a refinement of Sakata's model. Rather than  $p$ ,  $n$ , and  $\Lambda$  used in Sakata's model, the quark model introduces quarks with three different flavours, namely  $u$ ,  $d$ , and  $s$ . Mesons are considered to be bound states of a quark and an antiquark, which is similar to the construction in Sakata's model, while baryons are considered to be bound states of three quarks. The possible representations are:

$$\text{mesons } \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} ,$$

$$\text{baryons } \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} .$$

One merit of this model is that all the representations appearing on the right-hand side are actually realized in nature. Furthermore, this model succeeded in explaining intrinsic features of hadrons such as the mass and the magnetic moment. However, Gell-Mann [67, 68] and Han and Nambu [76, 77] showed that, when this model is applied to the interpretation of real hadrons, in order to avoid a contradiction with the exclusion principle, an until then unknown degree of freedom had to be introduced. This degree of freedom was called *colour* to distinguish it from flavour.

Unlike the democratic theory of hadrons, the quark model can be easily included in field theory. Quarks and leptons are described by fundamental fields, and hadrons are considered to be composite particles. Further, the left-handed or right-handed current densities in the current algebras can be expressed in terms of very simple bilinear forms using the quark model. In other words, when writing weak interactions in Yukawa form, the hadron parts can be written as products of bilinear forms and  $W$  fields. This was a great advantage, recalling from above that expressing the left-handed and right-handed current densities as hadron fields required highly non-linear representations. Thus the quark model was highly promising from the standpoint of including weak interactions.

The charges of the three types of quark can be determined through the fact that the baryon consists of three quarks. The charge of  $u$  is  $2e/3$ , and the charges of  $d$  and  $s$  are  $-e/3$ , where  $e$  is the positron charge. It is a feature of the quark model that quarks must have fractional charges. However, although experimentalists have

tried to detect such fractional charges, they have not been found yet. Therefore, the hypothesis was put forward that independent quarks could not be observed in principle, for some reason. This is the hypothesis known as *quark confinement*. If we accept this hypothesis, there is a sense in which Chew's idea of a democracy of hadrons can still be realized. This is because the  $S$ -matrix theory treats only observable hadrons. The dispersion relation is given for the  $S$ -matrix elements relevant to the processes involving only hadrons, while the unitarity condition is also expressed by states consisting only of hadrons, which means that the quark states do not appear.

However, the dynamics of hadrons is not determined within the framework of the  $S$ -matrix theory. To do this, we need to know the dynamics of gluons mediating strong interactions and quarks. We can say that such an application of the quark model to both strong and weak interactions has paved the way for theories of gauge fields mediating strong, electromagnetic, and weak interactions. We will discuss the theories of gauge fields in the latter half of this book. In this chapter, we went through the story of trial and error that eventually led to the theories of gauge fields. We have used many terms without giving strict definitions, but these will be explained step by step in the following chapters.

# Chapter 2

## Canonical Formalism and Quantum Mechanics



Quantum field theory is the quantum mechanics of systems with an infinite number of degrees of freedom, and it is constructed as the limit of systems with a finite number of degrees of freedom. After reviewing general features of quantum mechanics for systems with a finite number of degrees of freedom, we will attempt to extend the formalism.

### 2.1 Schrödinger's Picture and Heisenberg's Picture

In general, in quantum mechanics, the equations of motion are either Schrödinger's or Heisenberg's. Given the Hamiltonian of a system as a function of the set of canonically conjugate variables,  $q_1, q_2, \dots, q_f$  and  $p_1, p_2, \dots, p_f$ , Schrödinger's equation is

$$i \frac{\partial}{\partial t} \Psi(t) = H \left( q_r, \frac{1}{i} \frac{\partial}{\partial q_r} \right) \Psi(t) . \tag{2.1}$$

$\Psi(t)$  is called a *probability amplitude* or *state vector*. This equation holds in the so-called *Schrödinger picture*, in which an operator  $\mathcal{O}$  representing a physical quantity is explicitly independent of time. The expectation value of the dynamical quantity  $\mathcal{O}$  at time  $t$  is

$$\langle \mathcal{O} \rangle = \int \Psi^*(t) \mathcal{O} \Psi(t) dq , \quad dq = \rho dq_1 dq_2 \dots dq_f , \tag{2.2}$$

where  $\rho$  stands for an appropriate density.

In this book, we use natural units, such that  $\hbar = c = 1$ , where  $h$  and  $c$  stand for Planck's constant and the speed of light, respectively. From the point of view of relativity, it is natural to take the same unit for time and space; the unit of time is

considered to be the time light takes to travel a unit length, for instance 1 cm. In addition, although the reduced Planck constant  $\hbar = h/2\pi$ , frequently appearing in quantum mechanics, is a very small unit from the perspective of the human scale, it is useful to choose this as unit because it always appears when physical quantities are expressed in the microscopic world. For example, angular momentum always appears in units of  $\hbar$ , and in natural units this is merely an integer or a half-integer. In the natural system, all the units of physical quantities are given as certain powers of length.

In Schrödinger's picture, all the time dependence is assigned to state vectors, while in Heisenberg's picture the time dependence is assigned to the operators corresponding to physical quantities. Since all the physical quantities are functions of the canonical variables, they can be determined using the equations of motion, which are Heisenberg's equations in our case:

$$\frac{dq_r(t)}{dt} = i[H, q_r(t)] , \quad \frac{dp_r(t)}{dt} = i[H, p_r(t)] , \quad (2.3)$$

where the canonical variables must satisfy the canonical commutation relations

$$[q_r(t), p_s(t)] = i\delta_{rs} , \quad [q_r(t), q_s(t)] = 0 , \quad [p_r(t), p_s(t)] = 0 . \quad (2.4)$$

In this case, the expectation value of  $\mathcal{O}$  at time  $t$ , corresponding to (2.2), is given by

$$\langle \mathcal{O} \rangle = \int \Psi^* \mathcal{O}(t) \Psi dq . \quad (2.5)$$

The two representations are related by the following unitary transformations:

$$\Psi(t) = e^{-itH} \Psi , \quad \mathcal{O}(t) = e^{itH} \mathcal{O} e^{-itH} , \quad (2.6)$$

for a state vector and an operator, respectively. Taking appropriate care of the operator orderings, as a consequence of (2.4), we obtain for a function  $F$  of  $p$  and  $q$  the equations

$$\frac{\partial F}{\partial p_r} = i[F, q_r] , \quad \frac{\partial F}{\partial q_r} = -i[F, p_r] . \quad (2.7)$$

Hence, (2.3) can be written in the canonical form

$$\frac{dq_r(t)}{dt} = \frac{\partial H}{\partial p_r(t)} , \quad \frac{dp_r(t)}{dt} = -\frac{\partial H}{\partial q_r(t)} , \quad (2.8)$$

where the  $t$ -dependence has been made explicit in  $p_r(t)$  and  $q_r(t)$ , because these operators are defined in the Heisenberg picture.

## 2.2 Hamilton's Principle

In classical dynamics, the equation of motion, called the *Euler–Lagrange equation*, is derived from the variation principle

$$\delta \int_{t_0}^{t_1} dt L(q_r, \dot{q}_r) = 0 . \quad (2.9)$$

This is *Hamilton's principle*. Considering  $q_r(t_0)$  and  $q_r(t_1)$  as given and imposing the conditions

$$\delta q_r(t_0) = \delta q_r(t_1) = 0 , \quad (2.10)$$

the Euler–Lagrange equations are derived as

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_r} \right) - \frac{\partial L}{\partial q_r} = 0 , \quad (2.11)$$

where  $\dot{q}_r$  is the time derivative of  $q_r$ . To pass from here to the canonical form, it is enough to define the canonically conjugate momentum  $p_r$  as

$$p_r = \frac{\partial L}{\partial \dot{q}_r} . \quad (2.12)$$

When  $\dot{q}_r$  can be written in terms of  $q$  and  $p$ , from this equation, we call  $L$  a *non-singular Lagrangian*. In this case, the Hamiltonian is defined by

$$H = \sum_{r=1}^f p_r \dot{q}_r - L . \quad (2.13)$$

## 2.3 Equivalence Between the Canonical Equations and Lagrange's Equations

In this section, we derive Lagrange's equations by combining the canonical equations and the canonical commutation relations. A combination of the canonical equations (2.3), the canonical commutation relations (2.4), and the definition of the



Hamiltonian (2.13) yield

$$\begin{aligned}\frac{dq_r}{dt} &= i[H, q_r] \\ &= i \left[ \sum_{j=1}^f p_j \dot{q}_j - L, q_r \right] \\ &= \dot{q}_r + i \sum p_j [\dot{q}_j, q_r] - i[L, q_r].\end{aligned}$$

Taking into account the fact that the  $q$ 's commute, we use

$$[L, q_r] = \sum_j \frac{\partial L}{\partial \dot{q}_j} [\dot{q}_j, q_r] + \sum_j \frac{\partial L}{\partial q_j} [q_j, q_r] = \sum_j \frac{\partial L}{\partial \dot{q}_j} [\dot{q}_j, q_r].$$

Inserting this in the previous equation, we find

$$\frac{dq_r}{dt} = \dot{q}_r + i \sum_j \left( p_j - \frac{\partial L}{\partial \dot{q}_j} \right) [\dot{q}_j, q_r].$$

This equation always holds as long as the identity (2.12) does. Starting from the other canonical equation, we find similarly

$$\begin{aligned}\frac{dp_r}{dt} &= i[H, p_r] \\ &= \frac{\partial L}{\partial q_r} + i \sum_j \left( p_j - \frac{\partial L}{\partial \dot{q}_j} \right) [\dot{q}_j, p_r].\end{aligned}$$

Thus, if (2.13) holds,

$$\frac{dp_r}{dt} - \frac{\partial L}{\partial q_r} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_r} \right) - \frac{\partial L}{\partial q_r} = 0.$$

As a consequence, we can understand that the canonical equations are equivalent to Lagrange's equations as long as the canonical commutation relations hold. This fact will be used to quantize fields.

## 2.4 Equal-Time Canonical Commutation Relations

As we have seen in the previous section, the equal-time commutation relations (2.4) are the basic relations of quantum mechanics. We prove below that these relations hold at any time if they hold at some time  $t = t_0$ . First, from the canonical

equations (2.3) and the Jacobi identity, we get the differential equation

$$\frac{d}{dt}[q_r(t), p_s(t)] = i[H, [q_r(t), p_s(t)]] . \quad (2.14)$$

Solving this equation with the initial condition

$$[q_r(t_0), p_s(t_0)] = i\delta_{rs} , \quad (2.15)$$

we obtain the following relation for arbitrary time  $t$  :

$$[q_r(t), p_s(t)] = i\delta_{rs} . \quad (2.16)$$

# Chapter 3

## Quantization of Free Fields



The dynamical systems described in the previous chapter had a finite number of degrees of freedom  $f$ . How can we extend the theory to the limit  $f \rightarrow \infty$ ? To investigate this problem, we start by implementing an easy extension of the theory and then generalizing it.

### 3.1 Field Theory Based on Canonical Formalism

First, we consider  $N$  systems with no interaction among them. The total Lagrangian is given by the sum of Lagrangians corresponding to each system, namely,

$$L = \sum_{n=1}^N L_n(q_n, \dot{q}_n) , \quad (3.1)$$

and Hamilton's principle, or the principle of least action, is given by

$$\delta \int_{t_0}^{t_1} dt \sum_{n=1}^N L_n(q_n, \dot{q}_n) = 0 . \quad (3.2)$$

Then, treating  $N$  as a continuous infinity, we replace the discrete number  $n$  by the continuous parameter  $x$ . At the same time, we change the symbol of the dynamical variable from  $q_n$  to  $\varphi(x)$ . We call the dynamical variable, which is a function of continuous variables, a *field*. At the same time, designating a domain of integration, the principle of least action assumes the form

$$\delta \int dt L = \delta \int dt \left( \int dx \mathcal{L}[\varphi(x), \dot{\varphi}(x)] \right) = 0 . \quad (3.3)$$

$\mathcal{L}$  is called a *Lagrangian density*. In the case when  $x$  stands for the spatial coordinates in three dimensions,  $dx$  denotes the volume element  $d^3x = dx dy dz$ . At the same time,  $dt d^3x = d^4x$  stands for the four-dimensional volume element in natural units. In this system,  $t$  is sometimes written as  $x_0$ . Thus, the principle of least action in three-dimensional space can be written as

$$\delta \int_{\Omega} d^4x \mathcal{L}[\varphi(x), \dot{\varphi}(x)] = 0, \quad (3.4)$$

where  $\Omega$  stands for a certain four-dimensional domain and  $x$  denotes the four-dimensional coordinate. This is the principle of least action for the case when there are no interactions between fields at different points, but in this case we cannot describe the waves propagating through space. In relativistic field theory based on the principle of actions through a medium, a field at some point interacts only with its infinitesimal neighbours. Put another way, the Lagrangian density is a function of  $\varphi(x)$ ,  $\dot{\varphi}(x)$ , and  $\varphi(x + dx)$ . Then, using

$$\varphi(x + dx) = \varphi(x) + \frac{\partial \varphi(x)}{\partial x_k} dx_k, \quad k = 1, 2, 3,$$

the Lagrangian density becomes a function of  $\varphi$  and its time-like and space-like derivatives. Putting these together, we can write

$$\mathcal{L} = \mathcal{L} \left[ \varphi(x), \frac{\partial \varphi(x)}{\partial x_{\mu}} \right], \quad \mu = 1, 2, 3, 4, \quad (3.5)$$

where  $x_4 = ix_0 = it$ . In general, there exist several fields, and we distinguish them by an additional label  $\alpha$  as  $\varphi_{\alpha}(x)$ . Thus, the principle of least action can be written as

$$\delta \int d^4x \mathcal{L} \left[ \varphi_{\alpha}(x), \frac{\partial \varphi_{\alpha}(x)}{\partial x_{\mu}} \right] = 0. \quad (3.6)$$

### 3.1.1 Canonical Commutation Relations

Let us consider how we can extend the canonical commutation relations for a system with a finite number of degrees of freedom to those for a system with an infinite number of degrees of freedom. First, writing the equal-time canonical commutation relations in a system with finite degrees of freedom as

$$\sum_s [q_r, p_s] A_s = i A_r, \quad (3.7)$$

where  $A_r$  is an arbitrary quantity, we can make a dictionary to translate from discrete variables to continuous variables:

$$q_s \rightarrow \varphi_\alpha(x) ,$$

$$p_s = \frac{\partial L}{\partial \dot{q}_s} \rightarrow \pi_\alpha(x) = \frac{\partial}{\partial \dot{\varphi}(x)} \mathcal{L} \left[ \varphi_\alpha(x), \frac{\partial \varphi_\alpha(x)}{\partial x_\mu} \right] ,$$

$$\sum_s \rightarrow \sum_\alpha \int d^3x .$$

Using these rules, we translate (3.7) to continuous variables. Since the condition is given at equal times, we assume  $x_0 = x'_0$  in the following, and we have

$$\sum_\beta \int d^3x' [\varphi_\alpha(x), \pi_\beta(x')] f_\beta(x') = i f_\alpha(x) . \quad (3.8)$$

Assuming that this holds for arbitrary  $f_\beta(x')$ , we get the canonical commutation relation in field theory:

$$[\varphi_\alpha(x), \pi_\beta(x')] = i \delta_{\alpha\beta} \delta^3(x - x') , \quad (3.9)$$

where  $x_0 = x'_0$ , and

$$\delta^3(x - x') = \delta(x - x') \delta(y - y') \delta(z - z') .$$

As we shall see later, commutation relations like this hold only for fields satisfying Bose statistics. For fields satisfying Fermi statistics, we need to replace the commutation relations by anti-commutation relations. We shall discuss this issue later. In addition, the relations corresponding to

$$[q_r, q_s] = [p_r, p_s] = 0 \quad (3.10)$$

are

$$[\varphi_\alpha(x), \varphi_\beta(x')] = [\pi_\alpha(x), \pi_\beta(x')] = 0 , \quad x_0 = x'_0 . \quad (3.11)$$

The canonical quantization of fields as presented above was introduced by Heisenberg and Pauli.

### 3.1.2 Euler–Lagrange Equations

The equations of motion for fields are derived through the principle of least action

$$\delta \int_{\Omega} d^4x \mathcal{L} = 0, \quad (3.12)$$

where  $\varphi_{\alpha}$  is given in the four-dimensional domain  $\Omega$ , such that, on the boundary surface  $\partial\Omega$  of the domain, the variations of the field vanish, i.e.,  $\delta\varphi_{\alpha} = 0$ . The variation considered here is  $\delta\varphi_{\alpha}(x) = \varphi'_{\alpha}(x) - \varphi_{\alpha}(x)$ , and a variation commutes with a derivative. In addition, from now on we shall write derivatives as follows:

$$\frac{\partial\varphi_{\alpha}(x)}{\partial x_{\mu}} = \partial_{\mu}\varphi_{\alpha}(x) = \varphi_{\alpha,\mu}(x). \quad (3.13)$$

We start with the following equation:

$$\delta \int_{\Omega} d^4x \mathcal{L} = \int_{\Omega} d^4x \left[ \frac{\partial\mathcal{L}(x)}{\partial\varphi_{\alpha}(x)} \delta\varphi_{\alpha}(x) + \frac{\partial\mathcal{L}(x)}{\partial\varphi_{\alpha,\mu}(x)} \delta\varphi_{\alpha,\mu}(x) \right]. \quad (3.14)$$

Using the commutativity of a variation and a derivative, viz.,

$$\delta\varphi_{\alpha,\mu}(x) = \partial_{\mu}\delta\varphi_{\alpha}(x), \quad (3.15)$$

we perform a partial integral, using Gauss' theorem:

$$\int_{\Omega} d^4x \partial_{\mu} A_{\mu} = \int_{\partial\Omega} d\sigma_{\mu} A_{\mu}, \quad (3.16)$$

where  $d\sigma_{\mu} = d^4x/dx_{\mu}$ . Thus,

$$\begin{aligned} \delta \int_{\Omega} d^4x \mathcal{L}(x) &= \int_{\Omega} d^4x \left[ \frac{\partial\mathcal{L}(x)}{\partial\varphi_{\alpha}(x)} - \frac{\partial}{\partial x_{\mu}} \left( \frac{\partial\mathcal{L}(x)}{\partial\varphi_{\alpha,\mu}(x)} \right) \right] \delta\varphi_{\alpha}(x) \\ &\quad + \int_{\partial\Omega} d\sigma_{\mu} \frac{\partial\mathcal{L}(x)}{\partial\varphi_{\alpha,\mu}(x)} \delta\varphi_{\alpha}(x). \end{aligned} \quad (3.17)$$

From the assumption that the variations of the field vanish on the boundary surface, the second term vanishes. Using the fact that the variations  $\delta\varphi_{\alpha}$  are arbitrary, the first term leads to

$$[\mathcal{L}]_{\varphi_{\alpha}} = \frac{\partial\mathcal{L}}{\partial\varphi_{\alpha}} - \frac{\partial}{\partial x_{\mu}} \left( \frac{\partial\mathcal{L}}{\partial\varphi_{\alpha,\mu}} \right) = 0. \quad (3.18)$$

These are the Euler–Lagrange equations. We call the left-hand side an *Euler derivative*.

### Example: Klein–Gordon Equation

The following equation is the *Klein–Gordon equation*:

$$(\square - m^2)\varphi = 0 , \quad (3.19)$$

where  $\square = \partial_\mu^2 = \Delta - \partial_0^2$ . We choose the Lagrangian density

$$\mathcal{L} = -\frac{1}{2}[(\partial_\mu\varphi)^2 + m^2\varphi^2] . \quad (3.20)$$

In the following we shall use the Pauli metric. That is to say, we define an inner product of vectors as

$$x \cdot y = \mathbf{x} \cdot \mathbf{y} - x_0 y_0 . \quad (3.21)$$

Thus, the field equation derived from (3.20) becomes

$$[\mathcal{L}]_\varphi = -m^2\varphi + \partial_\mu^2\varphi = (\square - m^2)\varphi = 0 ,$$

and (3.19) is reproduced. The kinetic term in the Lagrangian density is  $\dot{\varphi}^2/2$  due to the overall minus sign, and since this term is positive, the energy in the whole system is positive-definite, which will be explained below.

### 3.1.3 Hamiltonian

To obtain the Hamiltonian for fields we again apply the translation from point-particle systems. We rewrite the defining identity of the Hamiltonian (2.13) using the rule given before:

$$H = \sum_\alpha \int d^3x \pi_\alpha(x) \dot{\varphi}_\alpha(x) - \int d^3x \mathcal{L}(x) \equiv \int d^3x \mathcal{H}(x) , \quad (3.22)$$

$$\mathcal{H}(x) = \sum_\alpha \pi_\alpha(x) \dot{\varphi}_\alpha(x) - \mathcal{L}(x) , \quad (3.23)$$

where  $H$  is the Hamiltonian and  $\mathcal{H}$  is the Hamiltonian density.

### Example: Hamiltonian for Real Scalar Field

Rewriting the previous example (3.20),

$$\mathcal{L} = \frac{1}{2}[\dot{\varphi}^2 - (\nabla\varphi)^2 - m^2\varphi^2]. \quad (3.24)$$

From this, the canonical variable conjugate to  $\varphi$  is

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\varphi}} = \dot{\varphi}. \quad (3.25)$$

Therefore,

$$\mathcal{H} = \pi\dot{\varphi} - \mathcal{L} = \frac{1}{2}[\pi^2 + (\nabla\varphi)^2 + m^2\varphi^2]. \quad (3.26)$$

This is clearly positive-definite.

As we have seen above, the canonical formalism for point-particle systems can be extended to field theory by using the simple translation rule.

## 3.2 Relativistic Generalization of the Canonical Equations

The canonical equation in quantum mechanics, viz.,

$$\frac{d\mathcal{O}}{dt} = i[H, \mathcal{O}], \quad (3.27)$$

holds directly for field theory. Since  $P_0$  is the time component of the energy-momentum vector, writing  $P_4 = iP_0 = iH$  and considering  $\mathcal{O}(x)$  as a polynomial in the field operators, we can write (3.27) in the form

$$[P_4, \mathcal{O}(x)] = i\frac{\partial\mathcal{O}(x)}{\partial x_4}. \quad (3.28)$$

If a theory is relativistically invariant, we can generalize this equation to

$$[\mathcal{O}(x), P_\mu] = \frac{1}{i}\frac{\partial}{\partial x_\mu}\mathcal{O}(x), \quad (3.29)$$

so there must be spatial components  $P_k$  ( $k = 1, 2, 3$ ) satisfying this equation.

In fact, it can be shown that the proper choice for  $P_k$  is

$$P_k = - \int d^3x \sum_{\alpha} \pi_{\alpha}(x) \partial_k \varphi_{\alpha}(x). \quad (3.30)$$



As an example, taking  $\varphi_\alpha$  or  $\pi_\alpha$  as  $\mathcal{O}$ , we find

$$[\varphi_\alpha(x), P_k] = -i \int d^3x' \delta^3(x - x') \partial_k \varphi_\alpha(x') = \frac{1}{i} \frac{\partial}{\partial x_k} \varphi_\alpha(x)$$

and

$$[\pi_\alpha(x), P_k] = i \int d^3x' \pi_\alpha(x') \frac{\partial}{\partial x'_k} \delta^3(x - x') = \frac{1}{i} \frac{\partial}{\partial x_k} \pi_\alpha(x).$$

In this way, we obtain the four-dimensional energy–momentum vector  $P_\mu$ . These four components commute with each other:

$$[P_\mu, P_\nu] = 0. \quad (3.31)$$

Now we sandwich a general operator  $\mathcal{O}(x)$  between the eigenstates  $|a\rangle$  and  $|b\rangle$  of  $P_\mu$ , which satisfy

$$P_\mu |a\rangle = p_\mu |a\rangle, \quad P_\mu |b\rangle = q_\mu |b\rangle. \quad (3.32)$$

This gives

$$\frac{\partial}{\partial x_\mu} \langle a | \mathcal{O}(x) | b \rangle = i \langle a | [\mathcal{O}(x), P_\mu] | b \rangle = -i(p_\mu - q_\mu) \langle a | \mathcal{O}(x) | b \rangle. \quad (3.33)$$

The solution of this differential equation is

$$\langle a | \mathcal{O}(x) | b \rangle = \langle a | \mathcal{O}(0) | b \rangle e^{-i(p-q) \cdot x}, \quad (3.34)$$

and the  $x$ -dependence of  $\mathcal{O}(x)$  can thus be determined. This relation is important and will be used often later.

### 3.3 Quantization of the Real Scalar Field

Since the real scalar field has already appeared as an example, we now proceed to quantize it. First, we consider a non-interacting case. To choose the Lagrangian density  $\mathcal{L}$ , we take into account the following conditions:

1.  $\mathcal{L}$  is invariant under the Lorentz transformations.
2. The Hamiltonian density  $\mathcal{H}$  calculated from  $\mathcal{L}$  is positive-definite.
3. In the non-interacting case, the fields satisfy the Klein–Gordon equation.

We now examine condition 1 above in detail. It is enough that  $\mathcal{L}$  should be invariant only under the proper Lorentz transformations, i.e., the transformations generated

by accumulating infinitesimal Lorentz transformations. Regarding condition 3, if we replace  $p_\mu$  by the operator  $-i\partial/\partial x_\mu$  in the Einstein relation

$$p^2 + m^2 = p_\mu^2 + m^2 = 0 \quad (3.35)$$

and apply the result to a field, we obtain the Klein–Gordon equation. This is the condition associating a given free field to a particle of given mass.

The Lagrangian density (3.20) is a scalar, i.e., Lorentz invariant. In addition, the Hamiltonian density (3.26) calculated from it is positive-definite. The field equation derived using the principle of least action becomes the Klein–Gordon equation itself. Thus we can say that (3.20) is the correct Lagrangian density. The quantity canonically conjugate to  $\varphi$  is the time derivative of  $\varphi$ , as given by (3.25).

The fields are quantized using the equal-time commutation relations:

$$[\varphi(x), \varphi(x')] = [\pi(x), \pi(x')] = 0 \quad (3.36)$$

and

$$[\varphi(x), \pi(x')] = i\delta^3(x - x') , \quad (3.37)$$

where  $x_0 = x'_0$ .

Next, we put the scalar field system inside a cube with volume  $V = L^3$  and impose a periodic condition on it. Then we expand  $\varphi(x)$  in a Fourier series. In this box, the set of functions

$$\{e^{i\mathbf{p}\cdot\mathbf{x}}\} \quad (3.38)$$

forms a complete set, where  $p_j = n_j(2\pi/L)$  and the  $n_j$  are integers, so we get the expansion

$$\varphi(x) = \frac{1}{\sqrt{V}} \sum_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} c(\mathbf{p}, t) . \quad (3.39)$$

Substituting this expansion into the Klein–Gordon equation, we get the equation for the coefficient functions  $c(\mathbf{p}, t)$ :

$$\ddot{c}(\mathbf{p}, t) = -(\mathbf{p}^2 + m^2)c(\mathbf{p}, t) . \quad (3.40)$$

Then, setting  $p_0 = \sqrt{\mathbf{p}^2 + m^2}$ , the general solution for the equation above is

$$c(\mathbf{p}, t) = c_1(\mathbf{p})e^{-ip_0t} + c_2(\mathbf{p})e^{ip_0t} . \quad (3.41)$$

Therefore,

$$\begin{aligned}\varphi(x) &= \frac{1}{\sqrt{V}} \sum_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} [c_1(\mathbf{p})e^{-ip_0t} + c_2(\mathbf{p})e^{ip_0t}] \\ &= \frac{1}{\sqrt{V}} \sum_{\mathbf{p}} [e^{i\mathbf{p}\cdot\mathbf{x}-ip_0t} c_1(\mathbf{p}) + e^{-i\mathbf{p}\cdot\mathbf{x}+ip_0t} c_2(-\mathbf{p})] .\end{aligned}\quad (3.42)$$

In addition,  $\varphi$  is real or Hermitian as an operator, so we have

$$c_2(\mathbf{p}) = c_1^\dagger(-\mathbf{p}) . \quad (3.43)$$

From now on we shall simply write  $c_1$  as  $c$  and use the notation  $\mathbf{p} \cdot \mathbf{x} = \mathbf{p}\mathbf{x} - p_0t$ . Therefore,  $\varphi$  and  $\pi$  are

$$\varphi(x) = \frac{1}{\sqrt{V}} \sum_{\mathbf{p}} [e^{i\mathbf{p}\cdot\mathbf{x}} c(\mathbf{p}) + e^{-i\mathbf{p}\cdot\mathbf{x}} c^\dagger(\mathbf{p})] \quad (3.44)$$

and

$$\pi(x) = \frac{1}{\sqrt{V}} \sum_{\mathbf{p}} [-ip_0 e^{i\mathbf{p}\cdot\mathbf{x}} c(\mathbf{p}) + ip_0 e^{-i\mathbf{p}\cdot\mathbf{x}} c^\dagger(\mathbf{p})] . \quad (3.45)$$

Substituting these expansions into the equal-time commutation relations (3.36), we find

$$[c(\mathbf{p}), c(\mathbf{q})] = [c^\dagger(\mathbf{p}), c^\dagger(\mathbf{q})] = 0 , \quad (3.46)$$

$$[c(\mathbf{p}), c^\dagger(\mathbf{q})] + [c^\dagger(-\mathbf{p}), c(-\mathbf{q})] = 0 . \quad (3.47)$$

Next, we rewrite (3.37) as

$$[\varphi(x), \pi(x')] = \frac{i}{V} \sum_{\mathbf{p}} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} ,$$

where  $x'_0 = x_0$ . Substituting the expansions (3.44) and (3.45) into this equation, we have

$$ip_0 [c(\mathbf{p}), c^\dagger(\mathbf{q})] - ip_0 [c^\dagger(-\mathbf{p}), c(-\mathbf{q})] = i\delta_{\mathbf{p},\mathbf{q}} ,$$

or

$$[c(\mathbf{p}), c^\dagger(\mathbf{q})] - [c^\dagger(-\mathbf{p}), c(-\mathbf{q})] = \frac{1}{p_0} \delta_{\mathbf{p},\mathbf{q}} . \quad (3.48)$$

Combining this equation and (3.47) yields

$$[c(\mathbf{p}), c^\dagger(\mathbf{q})] = \frac{1}{2p_0} \delta_{\mathbf{p}, \mathbf{q}}. \quad (3.49)$$

Then, instead of  $c(\mathbf{p})$ , we define  $a(\mathbf{p})$  by

$$c(\mathbf{p}) = \frac{1}{\sqrt{2p_0}} a(\mathbf{p}), \quad c^\dagger(\mathbf{p}) = \frac{1}{\sqrt{2p_0}} a^\dagger(\mathbf{p}). \quad (3.50)$$

The commutation relations simplify to

$$[a(\mathbf{p}), a(\mathbf{q})] = [a^\dagger(\mathbf{p}), a^\dagger(\mathbf{q})] = 0 \quad (3.51)$$

and

$$[a(\mathbf{p}), a^\dagger(\mathbf{q})] = \delta_{\mathbf{p}, \mathbf{q}}. \quad (3.52)$$

Substituting (3.50) into (3.44), we obtain

$$\varphi(x) = \sum_{\mathbf{p}} \frac{1}{\sqrt{2p_0 V}} \left[ e^{i\mathbf{p} \cdot \mathbf{x}} a(\mathbf{p}) + e^{-i\mathbf{p} \cdot \mathbf{x}} a^\dagger(\mathbf{p}) \right]. \quad (3.53)$$

In order to interpret the operators  $a$  and  $a^\dagger$ , we represent the energy and the momentum in a total system by  $a$  and  $a^\dagger$ :

$$H = \int_V d^3x \mathcal{H}(x) = \frac{1}{2} \sum_{\mathbf{p}} p_0 [a(\mathbf{p}) a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p}) a(\mathbf{p})], \quad (3.54)$$

$$P_k = - \int_V d^3x \pi(x) \partial_k \varphi(x) = \frac{1}{2} \sum_{\mathbf{p}} p_k [a(\mathbf{p}) a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p}) a(\mathbf{p})]. \quad (3.55)$$

Then we define the following so-called number density operator  $n(\mathbf{p})$ :

$$n(\mathbf{p}) = a^\dagger(\mathbf{p}) a(\mathbf{p}). \quad (3.56)$$

Since (3.52) implies  $a(\mathbf{p}) a^\dagger(\mathbf{p}) = n(\mathbf{p}) + 1$ , (3.54) and (3.55) become

$$H = \sum_{\mathbf{p}} p_0 \left[ n(\mathbf{p}) + \frac{1}{2} \right], \quad P_k = \sum_{\mathbf{p}} p_k \left[ n(\mathbf{p}) + \frac{1}{2} \right]. \quad (3.57)$$

What we would like to emphasize here is that the eigenvalues of  $n(\mathbf{p})$  (which can be written as  $n$  by fixing  $\mathbf{p}$ ) become non-negative integers like  $0, 1, 2, \dots$ . To prove this, we start with  $[a, a^\dagger] = 1$ . Therefore,

$$[n, a] = -a, \quad [n, a^\dagger] = a^\dagger. \quad (3.58)$$

Here we describe the eigenstate of  $n$  as  $|n'\rangle$ .  $n'$  is the eigenvalue of the operator  $n$ , satisfying

$$n|n'\rangle = n'|n'\rangle. \quad (3.59)$$

Then, from (3.58),

$$\langle n'|[n, a^\dagger]|n''\rangle = (n' - n'')\langle n'|a^\dagger|n''\rangle = \langle n'|a^\dagger|n''\rangle.$$

Thus, the matrix elements of  $a^\dagger$  are nonzero only if  $n' = n'' + 1$ . In other words,  $a^\dagger$  is the operator raising the eigenvalues of  $n$  by 1. At the same time,  $a$  is the operator lowering the eigenvalues of  $n$  by 1. Therefore, the eigenvalues of  $n$  for the states connected by  $a^\dagger$  or  $a$  become

$$\dots, n_1 - 2, n_1 - 1, n_1, n_1 + 1, n_1 + 2, \dots, \quad (3.60)$$

if we start with a given eigenvalue  $n_1$ . According to the definition (3.56), the eigenvalues of  $n$  cannot be negative, so there must be a minimum value for the series above. Denoting it by  $n_0$ , (3.60) becomes

$$n_0, n_0 + 1, n_0 + 2, \dots. \quad (3.61)$$

We find

$$\begin{aligned} \langle n'|n|n'\rangle &= \sum_{n''} \langle n'|a^\dagger|n''\rangle \langle n''|a|n'\rangle \\ &= \langle n'|a^\dagger|n' - 1\rangle \langle n' - 1|a|n'\rangle \\ &= |\langle n' - 1|a|n'\rangle|^2 = n'. \end{aligned} \quad (3.62)$$

Therefore, we can choose the undetermined phase factor to be

$$\langle n' - 1|a|n'\rangle = \langle n'|a^\dagger|n' - 1\rangle = \sqrt{n'}. \quad (3.63)$$

To find  $n_0$ , we impose the condition

$$a|n_0\rangle = 0. \quad (3.64)$$

Comparing it with (3.63), we can immediately find

$$n_0 = 0 . \quad (3.65)$$

This completely determines the eigenvalues of  $n$ . This  $n$  is called a *particle number operator*.

Going back to (3.57),

$$H = \sum_{\mathbf{p}} p_0 n(\mathbf{p}) + \text{c-number} , \quad P_k = \sum_{\mathbf{p}} p_k n(\mathbf{p}) + \text{c-number} . \quad (3.66)$$

We drop the c-numbers for two reasons:

1. The choice of operator orderings when going from classical mechanics to quantum mechanics is not unique. For instance, in (3.54), if we go to quantum mechanics after rewriting  $aa^\dagger$  as  $a^\dagger a$  within the framework of classical mechanics, then these c-numbers do not appear.
2. Only the differences between the eigenvalues of these operators can actually be observed.

If we drop the c-numbers, unifying the equations (3.66), we obtain

$$P_\mu = \sum_{\mathbf{p}} p_\mu n(\mathbf{p}) . \quad (3.67)$$

The interpretation of this is that  $n(\mathbf{p})$  stands for the number of particles with momentum  $\mathbf{p}$  and energy  $p_0$ . This interpretation is possible because, as mentioned before, the eigenvalues of  $n(\mathbf{p})$  are non-negative integers. Thus, although we have considered the scalar field as a wave satisfying the wave equation, its particle-like features have shown up as a consequence of imposing the canonical commutation relations to quantize it. This is nothing but an extension of Einstein's photon hypothesis to the scalar field. In addition, as is clear from (3.35),  $m$  stands for the mass of the quantum of the scalar field. As we call  $n$  the particle number operator,  $a$  and  $a^\dagger$  are called the *annihilation operator* and *creation operator*, respectively.

### 3.4 Quantization of the Complex Scalar Field

Complex scalar fields are quantized in a similar manner to real scalar fields. The first method is to quantize a complex scalar field by decomposing it into two real scalar fields. We choose the Lagrangian density

$$\mathcal{L} = - \left( \partial_\mu \varphi^\dagger \cdot \partial_\mu \varphi + m^2 \varphi^\dagger \varphi \right) , \quad (3.68)$$

where  $\varphi^\dagger$  is the operator Hermitian conjugate to  $\varphi$ . In classical theory,  $\varphi^*$  is used as the complex conjugate of  $\varphi$ . However, we use the notation  $\varphi^\dagger$  because  $\varphi$  becomes an operator through quantization. Then, making the change of variables

$$\varphi = \frac{1}{\sqrt{2}}(\varphi_1 - i\varphi_2), \quad \varphi^\dagger = \frac{1}{\sqrt{2}}(\varphi_1 + i\varphi_2), \quad (3.69)$$

the Lagrangian density can be rewritten in the form

$$\mathcal{L} = -\frac{1}{2} \left[ (\partial_\mu \varphi_1)^2 + m^2 \varphi_1^2 \right] - \frac{1}{2} \left[ (\partial_\mu \varphi_2)^2 + m^2 \varphi_2^2 \right]. \quad (3.70)$$

Since these are two real scalar fields, the corresponding quantization method is the same as in the previous section.

The second method is to quantize complex scalar fields without decomposing them into real scalar fields. To begin with, we define the canonical conjugate quantities of  $\varphi$  and  $\varphi^\dagger$  as follows:

$$\pi^\dagger = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}}, \quad \pi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}^\dagger}. \quad (3.71)$$

As a result, we have

$$\pi^\dagger = \dot{\varphi}^\dagger, \quad \pi = \dot{\varphi}. \quad (3.72)$$

The equal-time commutation relations that do not vanish and the Fourier expansions are given by

$$[\varphi(x), \pi^\dagger(x')] = [\varphi^\dagger(x), \pi(x')] = i\delta^3(x - x'), \quad \text{where } x_0 = x'_0, \quad (3.73)$$

$$\varphi(x) = \sum_{\mathbf{p}} \frac{1}{\sqrt{2p_0V}} [a(\mathbf{p})e^{ip \cdot x} + b^\dagger(\mathbf{p})e^{-ip \cdot x}], \quad (3.74)$$

and

$$\varphi^\dagger(x) = \sum_{\mathbf{p}} \frac{1}{\sqrt{2p_0V}} [b(\mathbf{p})e^{ip \cdot x} + a^\dagger(\mathbf{p})e^{-ip \cdot x}]. \quad (3.75)$$

In this way, when quantizing the complex scalar fields, we need two types of operators,  $a$  and  $b$ . Expressing the equal-time commutation relations in terms of  $a$  and  $b$  and repeating the calculation in the previous section, we obtain the following equations:

$$[a(\mathbf{p}), a^\dagger(\mathbf{q})] = [b(\mathbf{p}), b^\dagger(\mathbf{q})] = \delta_{\mathbf{p}, \mathbf{q}}, \quad (3.76)$$

where other commutators such as  $[a(\mathbf{p}), a(\mathbf{q})]$  all vanish. Therefore, introducing

$$n_+(\mathbf{p}) = a^\dagger(\mathbf{p})a(\mathbf{p}), \quad n_-(\mathbf{p}) = b^\dagger(\mathbf{p})b(\mathbf{p}), \quad (3.77)$$

both operators become particle number operators. They both correspond to quanta with mass  $m$ , but there is a quantum number that distinguishes between the two types of quanta, like the electric charge, for example.  $\pi^+$  and  $\pi^-$ , or  $K^0$  and  $\bar{K}^0$ , are examples of particle–antiparticle pairs. As for the latter, it is convenient to use a real scalar representation, and the corresponding particles are  $K_1^0$  and  $K_2^0$ . The two types of quanta, being complex conjugates, are said to be *antiparticles* of each other. This non-interacting theory is invariant under the interchange of the two fields, and the interchange can be expressed as a unitary transformation. Writing it as  $\mathcal{C}$ , we have

$$\mathcal{C}^{-1}\varphi(x)\mathcal{C} = e^{i\alpha}\varphi^\dagger(x), \quad \mathcal{C}^{-1}\varphi^\dagger(x)\mathcal{C} = e^{-i\alpha}\varphi(x), \quad (3.78)$$

where  $\alpha$  is a real number. We call this transformation *charge conjugation*. Since it commutes with  $P_\mu$ , a particle and its antiparticle have the same mass. By changing the phase of  $\varphi$  in a suitable way in the above equation, we can choose the phase factor  $e^{i\alpha}$  to be  $\pm 1$ .

### 3.5 Dirac Equation

Many particles with spin 1/2 obey Dirac's equation. Here we will discuss only the properties relevant to quantization.<sup>1</sup>

To translate the non-relativistic Schrödinger equation into the relativistic Dirac equation, the following conditions are used:

1. The wave function satisfies the Klein–Gordon equation.
2. A probability density  $\rho$  and a current density of probability  $\mathbf{j}$  are defined as quadratic forms of the wave function, satisfying the following equation of continuity:

$$\frac{\partial \rho}{\partial t} + \text{div } \mathbf{j} = 0, \quad \rho \geq 0. \quad (3.79)$$

---

<sup>1</sup> The properties of the Dirac equation are discussed in detail in my Japanese book [78].



To obtain Pauli's two-component wave function in a non-relativistic approximation, the wave function must be multicomponent. We write this as follows:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix}. \quad (3.80)$$

Therefore,  $\rho$  should be written as a quadratic form in  $\psi$  or, more precisely, as the following bilinear form in  $\psi$  and  $\psi^\dagger$ :

$$\rho = \sum_{r,s} \psi_r^* a_{rs} \psi_s. \quad (3.81)$$

In order for this to be positive-definite, the coefficients  $a_{rs}$  should form a positive-definite Hermitian matrix that can be diagonalized by a suitable unitary transformation. Furthermore, since we can rewrite it as a unit matrix by a scale transformation, without loss of generality, we can choose the following form from the beginning:

$$\rho = \sum_r \psi_r^* \psi_r = \psi^\dagger \psi. \quad (3.82)$$

Here,  $\psi^\dagger$  is the Hermitian conjugate of  $\psi$  having the form of a  $1 \times n$  matrix.

In terms of the Hamiltonian  $H$ , the equation of motion can be written in the form

$$i \frac{\partial}{\partial t} \psi = H \psi, \quad \text{or} \quad \left( i \frac{\partial}{\partial t} - H \right) \psi = 0. \quad (3.83)$$

From this equation we can obtain the second order differential equation with respect to time:

$$\left( \frac{\partial^2}{\partial t^2} + H^2 \right) \psi = 0. \quad (3.84)$$

This coincides with the Klein–Gordon equation if we set

$$H^2 = m^2 + \mathbf{p}^2 = m^2 - \nabla^2. \quad (3.85)$$

Assuming that  $H$  is linear in  $p$ , the most general form becomes

$$H = \boldsymbol{\alpha} \cdot \mathbf{p} + m\beta = \alpha_x p_x + \alpha_y p_y + \alpha_z p_z + m\beta, \quad (3.86)$$

where  $\alpha$  and  $\beta$  are suitable matrices. Thus, by using the matrix elements of  $\alpha$  and  $\beta$ , the equation of motion becomes

$$i \frac{\partial}{\partial t} \psi_r = \sum_s \left( \alpha_{rs} \frac{1}{i} \cdot \nabla + m \beta_{rs} \right) \psi_s . \quad (3.87)$$

We now determine the properties of  $\alpha$  and  $\beta$ :

1. Since  $H$  is Hermitian,  $\alpha$  and  $\beta$  are Hermitian matrices.
2.  $H$  must satisfy (3.85).

Substituting (3.86) into (3.85), we arrive at the following relations:

$$\begin{aligned} \alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = \mathbf{1} \quad (\text{unit matrix}), \\ \alpha_x \alpha_y + \alpha_y \alpha_x = \dots = 0, \quad \alpha_x \beta + \beta \alpha_x = \dots = 0. \end{aligned}$$

Combining the above relations, we obtain the alternative form

$$\alpha_k \alpha_l + \alpha_l \alpha_k = 2\delta_{kl} \mathbf{1}, \quad \alpha_k \beta + \beta \alpha_k = 0, \quad \beta^2 = \mathbf{1}. \quad (3.88)$$

Next, we derive the equation of continuity. To this end, we set

$$\rho = \psi^\dagger \psi, \quad \mathbf{j} = \psi^\dagger \alpha \psi . \quad (3.89)$$

Taking the complex conjugate of (3.87) and using the hermiticity of  $\alpha$  and  $\beta$ , we find

$$i \frac{\partial}{\partial t} \psi_r^* = \sum_s \left( \frac{1}{i} \nabla \psi_s^* \cdot \alpha_{sr} - m \psi_s^* \beta_{sr} \right) . \quad (3.90)$$

Combining (3.87) and (3.90), the equation of continuity can be derived:

$$\begin{aligned} i \frac{\partial \rho}{\partial t} &= i \sum_r \left( \frac{\partial \psi_r^*}{\partial t} \psi_r + \psi_r^* \frac{\partial \psi_r}{\partial t} \right) \\ &= \sum_{s,r} \left( \frac{1}{i} \nabla \psi_s^* \cdot \alpha_{sr} \psi_r - m \psi_s^* \beta_{sr} \psi_r \right) + \sum_{r,s} \left( \psi_r^* \alpha_{rs} \frac{1}{i} \cdot \nabla \psi_s + m \psi_r^* \beta_{rs} \psi_s \right) \\ &= -i \nabla \psi^\dagger \cdot \alpha \psi - i \psi^\dagger \alpha \cdot \nabla \psi \\ &= -i \operatorname{div}(\psi^\dagger \alpha \psi) \\ &= -i \operatorname{div} \mathbf{j} . \end{aligned}$$

### Pauli Representation

To satisfy the commutation relations (3.88), the dimension of the matrices of  $\alpha$  and  $\beta$  must be at least four. In order to derive explicit representations, Pauli introduced the four-dimensional  $\sigma$ -matrices and the  $\rho$ -matrix:

$$\sigma_j = \begin{pmatrix} \sigma_j & \mathbf{0} \\ \mathbf{0} & \sigma_j \end{pmatrix}, \quad j = 1, 2, 3, \quad (3.91)$$

where  $\sigma_j$  are the two-dimensional Pauli matrices,  $\mathbf{0}$  is the two-dimensional zero matrix, and  $\sigma_j$  are four-dimensional Pauli matrices. It is clear that the relations among them are

$$\sigma_1\sigma_2 = -\sigma_2\sigma_1 = i\sigma_3, \quad \text{and its cyclic permutations,} \quad (3.92)$$

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \mathbf{1} \quad (\text{four-dimensional unit matrix}). \quad (3.93)$$

Then the  $\rho$ -matrix is defined by

$$\rho_1 = \begin{pmatrix} \mathbf{0} & I \\ I & \mathbf{0} \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} \mathbf{0} & -iI \\ iI & \mathbf{0} \end{pmatrix}, \quad \rho_3 = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & -I \end{pmatrix}, \quad (3.94)$$

where  $I$  is the two-dimensional unit matrix. The properties (3.92) and (3.93) also hold for  $\rho$ :

$$\rho_1\rho_2 = -\rho_2\rho_1 = i\rho_3, \quad \text{and its cyclic permutations,} \quad (3.95)$$

$$\rho_1^2 = \rho_2^2 = \rho_3^2 = \mathbf{1}. \quad (3.96)$$

The three  $\sigma$ -matrices and the three  $\rho$ -matrices commute with each other. It is easily shown that  $[\sigma_i, \rho_j] = 0$  ( $i, j = 1, 2, 3$ ).

Then, in Pauli's representation,  $\alpha$  and  $\beta$  are given by

$$\alpha_j = \rho_1\sigma_j = \begin{pmatrix} \mathbf{0} & \sigma_j \\ \sigma_j & \mathbf{0} \end{pmatrix}, \quad \beta = \rho_3. \quad (3.97)$$

$\alpha$  and  $\beta$  constructed in this way clearly satisfy the conditions (3.88).

### Weyl Equation

So far the discussion about the Dirac equation has been based on the assumption that  $m \neq 0$ . Although in that case the wave function has four components, when  $m = 0$  we can make do with two components. In this case, since

$$H = \alpha \cdot p, \quad (3.98)$$

$\beta$  is not needed and the required commutation relation is only the first equation in (3.88), i.e., the commutation relations among the  $\alpha$ 's. There are two-dimensional representations satisfying this, namely,

$$\alpha_k = \pm \sigma_k . \quad (3.99)$$

Corresponding to this representation, the equation of motion becomes the two-component equation

$$i \frac{\partial}{\partial t} \psi = \pm \boldsymbol{\sigma} \cdot \mathbf{p} \psi . \quad (3.100)$$

This is called *Weyl's equation*. At first, this equation was not taken seriously because it is not compatible with the law of conservation of parity. However, after the proposal of the non-conservation of parity by Lee and Yang in [45], it attracted a lot of attention. Nowadays, it is used as the equation for massless neutrinos. The connection with the non-conservation of parity will be mentioned in the next section.

### 3.6 Relativistic Transformations of Dirac's Wave Function

Field operators generally possess definite properties under the action of the Lorentz transformations. For example, under the Lorentz transformation

$$x_\mu \rightarrow x'_\mu = \sum_\nu a_{\mu\nu} x_\nu , \quad (3.101)$$

a scalar field transforms in such a way that

$$\varphi'(x') = \varphi(x) . \quad (3.102)$$

We have changed  $\varphi$  to  $\varphi'$  because the functional form of the field changes upon a transformation of the coordinates. In other words, in the case of the scalar field, the change of variable and the change of functional form cancel each other, and as a whole the field is invariant. Whether we write  $x$  or  $x'$ , both labels do in fact stand for the same point. On the other hand, when we consider a vector field  $A_\mu(x)$ , its transformation property is

$$A_\mu(x) \rightarrow A'_\mu(x') = \sum_\nu a_{\mu\nu} A_\nu(x) . \quad (3.103)$$

Here, the problem is to see how the four-component Dirac field  $\psi$  transforms under Lorentz transformations. We assume the following linear transformation:

$$\psi \rightarrow \psi' = S\psi . \quad (3.104)$$

More precisely,

$$\psi_r(x) \rightarrow \psi'_r(x') = \sum_s S_{rs} \psi_s(x) . \quad (3.105)$$

We define the matrix  $S$  so that this transformation is compatible with the Dirac equation. Of course,  $S$  must be one of the representations of the Lorentz transformation (3.101). If such a matrix  $S$  exists, then the Dirac equation has exactly the same form in every inertial system, i.e., it is Lorentz covariant.

In order to determine the transformation, we first rewrite Dirac's equation. Multiplying

$$\left( -i \frac{\partial}{\partial t} + \frac{1}{i} \boldsymbol{\alpha} \cdot \nabla + m\beta \right) \psi = 0$$

from the left by  $\beta$ , we obtain

$$\left( \beta \frac{\partial}{\partial x_4} + \frac{1}{i} \beta \boldsymbol{\alpha} \cdot \nabla + m \right) \psi = 0 . \quad (3.106)$$

Then we define the  $\gamma$ -matrices by

$$\gamma_j = \frac{1}{i} \beta \alpha_j , \quad j = 1, 2, 3 , \quad \gamma_4 = \beta . \quad (3.107)$$

The four  $\gamma$ -matrices are Hermitian, and the commutation relations (3.88) can be expressed in the following unified form:

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\delta_{\mu\nu} \mathbf{1} , \quad \mu, \nu = 1, 2, 3, 4 . \quad (3.108)$$

Then, the Dirac equation can be written as

$$\left( \gamma_\mu \frac{\partial}{\partial x_\mu} + m \right) \psi = 0 . \quad (3.109)$$

It remains to prove that, if we start with the equation above and determine  $S$  properly, then the equation

$$\left( \gamma_\mu \frac{\partial}{\partial x'_\mu} + m \right) \psi' = 0 \quad (3.110)$$

will also hold. First, we consider the Lorentz transformation

$$\begin{aligned} x'_\mu &= \sum_\nu a_{\mu\nu} x_\nu, & \sum_\mu a_{\mu\rho} a_{\mu\sigma} &= \delta_{\rho\sigma}, \\ x_\nu &= \sum_\mu a_{\mu\nu} x'_\mu, & \sum_\mu a_{\rho\mu} a_{\sigma\mu} &= \delta_{\rho\sigma}. \end{aligned} \quad (3.111)$$

Here,  $(a_{\mu\nu})$  is an orthogonal matrix, and  $a_{j4}$  and  $a_{4j}$  ( $j = 1, 2, 3$ ) are pure imaginary numbers. Substituting

$$\frac{\partial}{\partial x'_\mu} = \sum_\nu \frac{\partial x_\nu}{\partial x'_\mu} \frac{\partial}{\partial x_\nu} = \sum_\nu a_{\mu\nu} \frac{\partial}{\partial x_\nu}$$

into (3.110) and using (3.104), we obtain

$$\sum_\mu \sum_\nu \gamma_\mu S a_{\mu\nu} \frac{\partial}{\partial x_\nu} \psi + m S \psi = 0.$$

Multiplying from the left by  $S^{-1}$ , we have

$$\sum_\mu \sum_\nu (S^{-1} \gamma_\mu S) a_{\mu\nu} \frac{\partial}{\partial x_\nu} \psi + m \psi = 0.$$

In order for this to coincide with (3.109), we need to require

$$\sum_\mu (S^{-1} \gamma_\mu S) a_{\mu\nu} = \gamma_\nu, \quad \text{or} \quad S^{-1} \gamma_\mu S = \sum_\nu a_{\mu\nu} \gamma_\nu. \quad (3.112)$$

This is one of the conditions for determining  $S$ . This  $S$  is a representation of the Lorentz transformation, as will become clear from the following discussion.

Assuming that, for another Lorentz transformation  $(b_{\mu\nu})$ , the relation

$$T^{-1} \gamma_\mu T = \sum_\nu b_{\mu\nu} \gamma_\nu$$

holds as well as the relation between  $S$  and  $(a_{\mu\nu})$ , we immediately obtain

$$(ST)^{-1} \gamma_\mu (ST) = \sum_\sigma \left( \sum_\nu a_{\mu\nu} b_{\nu\sigma} \right) \gamma_\sigma.$$

Thus, if  $S$  and  $T$  correspond to  $a$  and  $b$ , respectively, then  $ST$  corresponds to  $ab$ . In this sense,  $S$  and  $T$  are representations of the Lorentz transformations  $a$  and  $b$ .

Next, we derive another condition which  $S$  must satisfy. The complex conjugate of Dirac's equation is

$$\left[ (\gamma_\mu)_{rs}^* \frac{\partial}{\partial x_\mu^*} + m \delta_{rs} \right] \psi_s^* = 0, \quad (3.113)$$

where we sum over the index  $s$ . The space-like components of  $x_\mu$  are real numbers, and the time-like component is a pure imaginary number. In addition, the  $\gamma$ -matrices are Hermitian. Taking all this into account, we obtain

$$\sum_{j=1}^3 \frac{\partial \psi_s^*}{\partial x_j} (\gamma_j)_{sr} - \frac{\partial \psi_s^*}{\partial x_4} (\gamma_4)_{sr} + m \psi_r^* = 0.$$

Then, multiplying by  $\gamma_4$  from the right and using its anti-commutativity with  $\gamma_1, \gamma_2,$  and  $\gamma_3$ , we obtain

$$- \frac{\partial \bar{\psi}}{\partial x_\mu} \gamma_\mu + m \bar{\psi} = 0, \quad (3.114)$$

where

$$\bar{\psi} = \psi^\dagger \gamma_4. \quad (3.115)$$

In order for (3.114) to hold in any inertial system,  $\bar{\psi}$  has to transform as

$$\bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} S^{-1}. \quad (3.116)$$

We have considered  $\bar{\psi}$  instead of  $\psi^\dagger$  because the transformation property is easier to handle. However, since  $\psi$  and  $\bar{\psi}$  are not independent, we have to impose a restriction on  $S$  to make (3.104) and (3.114) compatible. Using the Hermitian conjugate of (3.104), we find

$$\bar{\psi}' = \psi'^\dagger \gamma_4 = \psi^\dagger S^\dagger \gamma_4 = \bar{\psi} \gamma_4 S^\dagger \gamma_4.$$

Comparing this equation with (3.116), we arrive at the condition

$$\gamma_4 S^\dagger \gamma_4 = S^{-1}, \quad \text{or} \quad S^\dagger \gamma_4 = \gamma_4 S^{-1}. \quad (3.117)$$

It should be emphasized that  $S^\dagger$  and  $S^{-1}$  cannot be treated as equal because  $S$  is not necessarily unitary. Then the two conditions determining  $S$  are (3.112) and (3.117). In the following, we shall determine the matrix  $S$  for an infinitesimal Lorentz transformation.

We write down the infinitesimal Lorentz transformation as

$$x'_\mu = x_\mu + \sum_\nu \epsilon_{\mu\nu} x_\nu, \quad \epsilon_{\mu\nu} = -\epsilon_{\nu\mu} \text{ infinitesimal.} \quad (3.118)$$

Here  $\epsilon_{\mu\nu}$  is a pure imaginary number if either  $\mu$  or  $\nu$  is 4. Corresponding to (3.118), we assume  $S$  to have the form

$$S = \mathbf{1} + \frac{1}{2} \sum_\mu \sum_\nu \epsilon_{\mu\nu} T_{\mu\nu}, \quad T_{\mu\nu} = -T_{\nu\mu}, \quad (3.119)$$

where  $T_{\mu\nu}$  are  $4 \times 4$  matrices produced by  $\gamma$ -matrices. Substituting the equation above into (3.112), the linear term in  $\epsilon$  is

$$\begin{aligned} \left[ \gamma_\mu, \frac{1}{2} \sum_\lambda \sum_\nu \epsilon_{\lambda\nu} T_{\lambda\nu} \right] &= \sum_\nu \epsilon_{\mu\nu} \gamma_\nu \\ &= \frac{1}{2} \sum_\lambda \sum_\nu \epsilon_{\lambda\nu} (\delta_{\lambda\mu} \gamma_\nu - \delta_{\mu\nu} \gamma_\lambda). \end{aligned}$$

Comparing the anti-symmetrized coefficients of the anti-symmetric tensor  $\epsilon_{\lambda\nu}$ , we deduce that

$$\gamma_\mu T_{\lambda\nu} - T_{\lambda\nu} \gamma_\mu = \delta_{\lambda\mu} \gamma_\nu - \delta_{\nu\mu} \gamma_\lambda. \quad (3.120)$$

This is the condition for determining the anti-symmetric tensor  $T_{\lambda\nu}$ . The tensor which satisfies this equation is

$$T_{\lambda\nu} = \frac{1}{4} (\gamma_\lambda \gamma_\nu - \gamma_\nu \gamma_\lambda). \quad (3.121)$$

It can be checked that this equation satisfies not only (3.112), but also (3.117).

Substituting (3.119) into (3.117) yields

$$\sum_\mu \sum_\nu \epsilon_{\mu\nu}^* T_{\mu\nu}^\dagger = - \sum_\mu \sum_\nu \epsilon_{\mu\nu} \gamma_4 T_{\mu\nu} \gamma_4. \quad (3.122)$$

We separate the cases when  $\mu$  and  $\nu$  contain 4 or not:

$$\epsilon_{ik} T_{ik}^\dagger = -\epsilon_{ik} \gamma_4 T_{ik} \gamma_4 = -\epsilon_{ik} T_{ik},$$

$$-\epsilon_{4k} T_{4k}^\dagger = -\epsilon_{4k} \gamma_4 T_{4k} \gamma_4 = \epsilon_{4k} T_{4k}.$$



Since both cases turn out to give

$$T_{\mu\nu}^\dagger = -T_{\mu\nu} , \quad (3.123)$$

we see that the condition can be satisfied. Thus it has been proven that for the infinitesimal Lorentz transformation,  $T_{\mu\nu}$ , i.e., the matrix  $S$ , exists. In addition, this tells us that

$$\text{Tr}(T_{\mu\nu}) = 0 . \quad (3.124)$$

It also gives the following restriction on  $S$ :

$$\det S = 1 . \quad (3.125)$$

Generalizing, it can be shown that  $S$  exists for the proper Lorentz transformations generated by infinitesimal Lorentz transformations. Actually, (3.125) also holds for finite transformations.<sup>2</sup>

### Space Inversion

One example of a special Lorentz transformation that does not belong to the proper Lorentz transformations is space inversion. In this case,  $S$  must be determined separately. Space inversion is defined by

$$x'_j = -x_j , \quad j = 1, 2, 3 , \quad x'_4 = x_4 . \quad (3.126)$$

The corresponding conditions which  $S$  must satisfy are

$$\gamma_j S = -S \gamma_j , \quad \gamma_4 S = S \gamma_4 , \quad S^\dagger \gamma_4 = \gamma_4 S^{-1} . \quad (3.127)$$

Up to an undetermined phase factor, the solution is

$$S = \gamma_4 . \quad (3.128)$$

Thus we can say that Dirac's equation preserves parity. However, this statement does not hold for the two-component Weyl equation. We shall discuss this below.

### Relativistic Covariance of Weyl's Equation

Although in Sect. 3.5 we used boldface to distinguish Pauli's matrix in two dimensions from the one in four dimensions, from now on we shall use boldface only for three-dimensional vectors and normal characters for their components. In the following, we shall prove the relativistic covariance of the Weyl equation

---

<sup>2</sup> To derive the explicit form of  $S$  corresponding to finite transformations, the reader is referred to [45].

similarly to that of the Dirac equation. We start with Weyl's equation (3.100):

$$i \frac{\partial}{\partial t} \psi = \pm \frac{1}{i} \boldsymbol{\sigma} \cdot \nabla \psi . \quad (3.129)$$

The positive sign corresponds to the right-handed neutrino and the negative one to the left-handed neutrino, as will be explained in the next section. Then we introduce the matrices  $\Gamma_\mu$ :

$$\Gamma_j = \sigma_j , \quad j = 1, 2, 3, \quad \Gamma_4 = \pm i . \quad (3.130)$$

Using the  $\Gamma$ -matrix, Weyl's equation can be written in the form

$$\Gamma_\mu \frac{\partial \psi}{\partial x_\mu} = 0 , \quad \frac{\partial \psi^\dagger}{\partial x_\mu} \Gamma_\mu = 0 . \quad (3.131)$$

Thus, the Lorentz transformation of the Weyl spinors becomes

$$\psi \rightarrow \psi' = S \psi , \quad \psi^\dagger \rightarrow \psi'^\dagger = \psi^\dagger S^\dagger . \quad (3.132)$$

The big difference with Dirac's equation is that in Weyl's equation there is no mass term. The covariance of the two equations in (3.131) (in other words, the condition that they hold in the same form in any inertial system) can be expressed as follows:

$$S^\dagger \Gamma_\mu S = \sum_\nu a_{\mu\nu} \Gamma_\nu . \quad (3.133)$$

Assuming that  $S$  has the form (3.119) for the infinitesimal Lorentz transformation (3.118),

$$\frac{1}{2} \epsilon_{\mu\nu}^* T_{\mu\nu}^\dagger \Gamma_\lambda + \frac{1}{2} \Gamma_\lambda \epsilon_{\mu\nu} T_{\mu\nu} = \epsilon_{\lambda\rho} \Gamma_\rho = \frac{1}{2} \epsilon_{\mu\nu} (\delta_{\mu\lambda} \Gamma_\nu - \delta_{\nu\lambda} \Gamma_\mu) . \quad (3.134)$$

From this equation,  $T_{\mu\nu}$  can be obtained immediately:

$$T_{12} = \pm T_{34} = \frac{i}{2} \sigma_3 , \quad T_{23} = \pm T_{14} = \frac{1}{2} \sigma_1 , \quad T_{31} = \pm T_{24} = \frac{i}{2} \sigma_2 . \quad (3.135)$$

Again we find the condition

$$\text{Tr}(T_{\mu\nu}) = 0 . \quad (3.136)$$

In this case the matrix  $S$  is unimodular, i.e.,

$$\det S = 1 . \quad (3.137)$$

For the space inversion (3.126),  $S$  must satisfy the following four equations:

$$S^\dagger \sigma_j S = -\sigma_j, \quad j = 1, 2, 3, \quad S^\dagger S = \mathbf{1}. \quad (3.138)$$

However, there exists no two-dimensional matrix satisfying these four equations. Thus, Weyl's equation does not maintain its form under space inversion. In other words, this equation is not defined with respect to space inversion. Hence, parity is not conserved. For this reason Weyl's equation was long considered unphysical. However, once the non-conservation of parity proposed by Lee and Yang [45] had been confirmed experimentally, Weyl's equation began to bask in the limelight as the equation of the massless neutrino. The theory of neutrinos described by this equation is called the *two-component theory of the neutrino*.

### 3.7 Solutions of the Free Dirac Equation

To quantize the Dirac field, we must first derive the complete set of solutions of the free Dirac equation. In Pauli's representation, the Hamiltonian is

$$H = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m = \rho_1 \boldsymbol{\sigma} \cdot \mathbf{p} + \rho_3 m, \quad (3.139)$$

with the  $\rho$ -matrices defined by (3.94). We begin by deriving the invariant quantities for the Dirac equation, in other words, the conserved quantities:

$$[H, H] = 0, \quad [p, H] = 0. \quad (3.140)$$

These equations express conservation of energy and momentum. Next, we consider the angular momentum:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}. \quad (3.141)$$

We note that

$$[L_3, \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m] = \boldsymbol{\alpha} \cdot [L_3, \mathbf{p}] = i(\alpha_1 p_2 - \alpha_2 p_1) = i(\boldsymbol{\alpha} \times \mathbf{p})_3.$$

Generalizing this equation we find that the orbital angular momentum is not a conserved quantity:

$$[\mathbf{L}, H] = i\boldsymbol{\alpha} \times \mathbf{p} = i\rho_1 \boldsymbol{\sigma} \times \mathbf{p}. \quad (3.142)$$

The same happens for spins:

$$[\sigma_3, \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m] = \rho_1 [\sigma_3, \boldsymbol{\sigma}] \cdot \mathbf{p} = 2i\rho_1(\sigma_2 p_1 - \sigma_1 p_2) = -2i\rho_1(\boldsymbol{\sigma} \times \mathbf{p})_3,$$

or

$$[\boldsymbol{\sigma}, H] = -2i\rho_1\boldsymbol{\sigma} \times \mathbf{p} . \quad (3.143)$$

Combining (3.142) and (3.143), we find that the total angular momentum is a conserved quantity:

$$[\mathbf{J}, H] = 0 , \quad \text{where } \mathbf{J} = \mathbf{L} + \frac{1}{2}\boldsymbol{\sigma} . \quad (3.144)$$

In contrast with the non-relativistic theory, even for the free particle, the spin and angular momentum are not conserved separately. Only their sum is conserved.

The *helicity* is another important conserved quantity. That is,

$$h = \boldsymbol{\sigma} \cdot \mathbf{p}/p , \quad \text{or } \lambda = h/2 , \quad (3.145)$$

is conserved, and the eigenvalues of  $h$  are  $\pm 1$ . To be precise,  $\lambda$  is the helicity, but here we call  $h$  helicity as well.

The motion of a free Dirac particle is determined by specifying its momentum and helicity. The wave function with the given momentum  $p$  takes the form

$$\psi(x) = \frac{1}{\sqrt{V}} e^{i\mathbf{p}\cdot\mathbf{x} - iEt} u(p) . \quad (3.146)$$

Here,  $V$  is the volume of the whole system within which the particle is confined, and  $u(p)$  is a four-component spinor.

In contrast with the scalar and the vector, the spinor is a physical quantity obeying the transformation rule (3.104). Substituting (3.146) into the Dirac equation, we find

$$Eu(p) = (\rho_1\boldsymbol{\sigma} \cdot \mathbf{p} + \rho_3m)u(p) \equiv H(p)u(p) . \quad (3.147)$$

In order to obtain the energy eigenvalue  $E$ , we diagonalize  $H(p)$  using a unitary matrix  $U$ . We use the formula

$$\exp\left(-\frac{i}{2}a\rho_2\right)\rho_3\exp\left(\frac{i}{2}a\rho_2\right) = \rho_3 \cos a + \rho_1 \sin a . \quad (3.148)$$

Then we write  $H(p)$  as

$$\rho_1\boldsymbol{\sigma} \cdot \mathbf{p} + \rho_3m = A(\rho_3 \cos a + \rho_1 \sin a) ,$$

where

$$A = \sqrt{(\boldsymbol{\sigma} \cdot \mathbf{p})^2 + m^2} = \sqrt{p^2 + m^2} , \quad \tan a = \boldsymbol{\sigma} \cdot \mathbf{p}/m , \quad (3.149)$$

with  $p$  standing for the absolute value of the vector  $p$ . Then, the matrix  $U$  is

$$U = \exp\left(\frac{i}{2}a\rho_2\right) = \exp\left(\frac{i}{2}\rho_2 \tan^{-1} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{m}\right). \quad (3.150)$$

The diagonal form of  $H(p)$  is

$$UH(p)U^{-1} = \sqrt{p^2 + m^2} \rho_3. \quad (3.151)$$

Using the formula

$$\tan^{-1} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{m} = \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{m} \frac{\tan^{-1}(p/m)}{p/m},$$

we put  $U$  in the form

$$U = \frac{1}{\sqrt{2A}} \left( \sqrt{A+m} + i\rho_2 \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{\sqrt{A+m}} \right). \quad (3.152)$$

Next, we introduce the spinor  $v(p) = Uu(p)$ , such that

$$Ev(p) = \sqrt{p^2 + m^2} \rho_3 v(p). \quad (3.153)$$

Since the eigenvalues of  $\rho_3$  are  $\pm 1$ , the eigenvalues  $E$  turn out to be

$$E = \pm \sqrt{p^2 + m^2}. \quad (3.154)$$

Because each eigenvalue is doubly degenerate, we distinguish the solutions by the helicity:

$$1) \ E = \sqrt{p^2 + m^2}, \quad h = 1 \quad 2) \ E = \sqrt{p^2 + m^2}, \quad h = -1$$

$$v(p) = \begin{pmatrix} \cos \theta/2 \\ e^{i\varphi} \sin \theta/2 \\ 0 \\ 0 \end{pmatrix}, \quad v(p) = \begin{pmatrix} -e^{i\varphi} \sin \theta/2 \\ \cos \theta/2 \\ 0 \\ 0 \end{pmatrix},$$

$$3) \ E = -\sqrt{p^2 + m^2}, \quad h = 1 \quad 4) \ E = -\sqrt{p^2 + m^2}, \quad h = -1$$

$$v(p) = \begin{pmatrix} 0 \\ 0 \\ \cos \theta/2 \\ e^{i\varphi} \sin \theta/2 \end{pmatrix}, \quad v(p) = \begin{pmatrix} 0 \\ 0 \\ -e^{i\varphi} \sin \theta/2 \\ \cos \theta/2 \end{pmatrix}.$$

This completes the solution of the free Dirac equation. The wave function can be normalized by requiring

$$\int_V d^3x \psi^\dagger(x) \psi(x) = 1 . \quad (3.155)$$

Note that there are positive and negative energy eigenvalues. The interpretation of negative energy solutions can be correctly given only after quantization. In addition, the Dirac particle does not obey Bose statistics like the spin-0 scalar particle. It obeys Fermi statistics, being a spin-1/2 particle. What kind of quantization must be chosen to satisfy the statement above? This is the subject of the next section.

### Relation Between Dirac's Equation and Weyl's Equation

Weyl's equation is obtained from Dirac's equation without the mass term:

$$H\psi = \boldsymbol{\alpha} \cdot \mathbf{p}\psi = E\psi . \quad (3.156)$$

We introduce the following Hermitian matrices:

$$\sigma_{\mu\nu} = \frac{1}{2i}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu) = \frac{2}{i}T_{\mu\nu} , \quad (3.157)$$

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4 . \quad (3.158)$$

The matrices  $\sigma_{23}$ ,  $\sigma_{31}$ , and  $\sigma_{12}$  are equivalent to  $\sigma_1$ ,  $\sigma_2$ , and  $\sigma_3$  in (3.91). Then we can write

$$\boldsymbol{\alpha} \cdot \mathbf{p} = -\gamma_5 \boldsymbol{\sigma} \cdot \mathbf{p} , \quad (3.159)$$

where  $\boldsymbol{\sigma}$  stands for the three-dimensional vector given by the three components  $\sigma_1$ ,  $\sigma_2$ , and  $\sigma_3$ . This matrix can be thought of as being either two-dimensional or four-dimensional, as we choose. We then decompose the wave function in the following way:

$$\psi = \psi_L + \psi_R = \frac{1 + \gamma_5}{2}\psi + \frac{1 - \gamma_5}{2}\psi . \quad (3.160)$$

In the case of a massless particle, Dirac's equation (3.156) reduces to the two equations

$$E\psi_L = -\boldsymbol{\sigma} \cdot \mathbf{p}\psi_L , \quad E\psi_R = \boldsymbol{\sigma} \cdot \mathbf{p}\psi_R . \quad (3.161)$$

By restricting to the positive energy solution, we find that, since  $E = p$ ,  $\psi_L$  corresponds to the case  $h = -1$ , called the *left-handed* wave function, with spin antiparallel to its momentum, while  $\psi_R$  corresponds to  $h = 1$ , called the

*right-handed* wave function, with spin parallel to the momentum. In the Weyl representation, the matrix  $\gamma_5$  is diagonal:

$$\gamma_5 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \quad (3.162)$$

Hence,  $\psi_L$  and  $\psi_R$  become the following two-component spinors:

$$\psi_L = \begin{pmatrix} \times \\ \times \\ 0 \\ 0 \end{pmatrix}, \quad \psi_R = \begin{pmatrix} 0 \\ 0 \\ \times \\ \times \end{pmatrix}, \quad (3.163)$$

where  $\times$  stands for the components which are in general not zero. This completes the derivation of the two-component Weyl equation. However, as we have seen above, it can be expressed in a four-component form.

As mentioned before, the neutrino can be considered to obey Weyl's equation. If this is true, is the sign of its helicity positive or negative? In other words, is the neutrino left-handed or right-handed? The helicity of the neutrino was measured by Goldhaber, Grozins, and Sunyar in [79]. The process they used is an elementary reaction called *K-capture* induced by essentially the same interaction as  $\beta$ -decay, namely

$$e^- + p \rightarrow n + \nu. \quad (3.164)$$

Both  $p$  and  $n$  are nucleons in an atomic nucleus. The following two-stage process takes place:

$$e^- + \underset{J=0}{A} \longrightarrow \underset{J=1}{B^*} + \nu, \quad \underset{J=1}{B^*} \longrightarrow \underset{J=0}{B} + \gamma, \quad (3.165)$$

where  $A$  is a spin-0 nucleus. Absorbing the electron on its  $K$ -shell according to (3.164), it becomes the spin-1 excited atomic nucleus  $B^*$  and simultaneously emits a neutrino. Then  $B^*$  reaches its ground state  $B$  by emitting a  $\gamma$ -ray.  $B$  is again in a spin-0 state. If there exists an atomic nucleus which undergoes such a chain of reactions, the helicity of the neutrino can be determined in the following way.

First, note that, when they are emitted in opposite directions, the neutrino  $\nu$  and the  $\gamma$  particle carry the maximum energy, thus satisfying the resonant scattering condition. We will check the conservation of angular momentum in this case. The  $z$ -axis is chosen as the direction in which  $\nu$  is radiated, and we consider  $J_z$  before and after the interaction.

If  $\nu$  is left-handed, so that its spin is antiparallel to the  $z$ -direction, i.e.,  $-1/2$ , we have the following possibilities for  $J_z$ :

$$\begin{array}{l}
 e^- + A \longrightarrow B^* + \nu \longrightarrow \gamma + B + \nu \\
 J_z : \begin{array}{ccccccc}
 \frac{1}{2} & 0 & 1 & -\frac{1}{2} & 1 & 0 & -\frac{1}{2} \\
 -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & \text{forbidden} & 0 & -\frac{1}{2}
 \end{array}
 \end{array}$$

In the analysis of the conservation law above, the case where the electron has  $J_z = -1/2$  is forbidden, because the  $z$  component of the spin of the photon emitted in the  $-z$  direction cannot be zero. This is a feature of massless particles, which will be explained later. Therefore, the only allowed case is when the photon is emitted in the  $-z$  direction with  $J_z = 1$ , i.e., when the photon is left-handed (meaning that it has left-circular polarization). Thus, if  $\nu$  is left-handed, the photon emitted with the maximum energy will also be left-handed. Similarly, if  $\nu$  is right-handed, then the photon will also be right-handed.

The atomic nuclei used in the actual experiments were

$$A = {}_{63}^{152}\text{Eu}(0^-), \quad B = {}_{62}^{152}\text{Sm}(0^+), \quad B^* = {}_{62}^{152}\text{Sm}(1^-), \quad (3.166)$$

where the numbers 0 or 1 in brackets represent the spins of the nuclei and the plus and minus signs indicate the parity. By measuring the polarization of the photons with the highest energies, it was deduced that the neutrino is *left-handed*.

### 3.8 Quantization of the Dirac Field

To quantize the Dirac field, we start from the Lagrangian density in the form

$$\mathcal{L} = -\bar{\psi} \left( \gamma_\mu \frac{\partial}{\partial x_\mu} + m \right) \psi, \quad (3.167)$$

and derive the Dirac equation. We must treat  $\psi$  and  $\bar{\psi}$  as independent quantities when deriving the field equation from this Lagrangian density using the variational principle. As a consequence, the following equations are obtained:

$$\left( \gamma_\mu \frac{\partial}{\partial x_\mu} + m \right) \psi = 0, \quad -\frac{\partial \bar{\psi}}{\partial x_\mu} \gamma_\mu + m \bar{\psi} = 0. \quad (3.168)$$

In addition, the quantity canonically conjugate to  $\psi_r$ , denoted  $\pi_r^\dagger$ , is

$$\pi_r^\dagger = \frac{\partial \mathcal{L}}{\partial \dot{\psi}_r} = i(\bar{\psi} \gamma_4)_r = i\psi_r^\dagger. \quad (3.169)$$



The four-dimensional continuity equation related to the law of conservation of probability, which we have discussed before, is

$$\partial_\mu j_\mu = 0, \quad \text{where } j_\mu = i\bar{\psi}\gamma_\mu\psi. \quad (3.170)$$

Combining the equations (3.104), (3.116), and (3.112), it is easily shown that  $j_\mu$  transforms as a four-vector under Lorentz transformations. In addition, it will be explained later that the canonical algebra for Dirac fields obeying Fermi statistics must be chosen as the following anti-commutation relation instead of (3.9):

$$[\psi_r(x), \pi_s^\dagger(x')]_+ = i\delta_{rs}\delta^3(x-x'), \quad \text{where } x_0 = x'_0. \quad (3.171)$$

Here,  $[\ , \ ]_+$  denotes the *anti-commutator*, sometimes written as  $\{ \ , \ }$ , and defined by

$$[A, B]_+ = \{A, B\} = AB + BA. \quad (3.172)$$

Sometimes, the commutator  $[\ , \ ]$  is denoted, in turn, by  $[\ , \ ]_-$ .

For  $x_0 = x'_0$ , we can write the equal-time anti-commutators in the form

$$\begin{aligned} [\psi_r(x), \psi_s^\dagger(x')]_+ &= \delta_{rs}\delta^3(x-x'), \\ [\psi_r(x), \psi_s(x')]_+ &= [\psi_r^\dagger(x), \psi_s^\dagger(x')]_+ = 0. \end{aligned} \quad (3.173)$$

Then, the Hamiltonian density becomes

$$\begin{aligned} \mathcal{H} &= \sum_r \pi_r^\dagger \dot{\psi}_r - \mathcal{L} = i\bar{\psi}\gamma_4\dot{\psi} + \bar{\psi}(\gamma_\mu\partial_\mu + m)\psi \\ &= \psi^\dagger(\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m)\psi, \quad \text{where } \mathbf{p} = -i\nabla. \end{aligned} \quad (3.174)$$

Note the appearance of the Dirac Hamiltonian sandwiched between  $\psi$  and  $\psi^\dagger$ . The Hamiltonian itself is given by the space integration:

$$H = \int d^3x \mathcal{H}(x) = \int d^3x \psi^\dagger(x) \left( \frac{1}{i} \boldsymbol{\alpha} \cdot \nabla + \beta m \right) \psi(x). \quad (3.175)$$

This completes the application of the canonical formalism.

To quantize, we introduce the Fourier expansion

$$\psi_r(x) = \frac{1}{\sqrt{V}} \sum_{\mathbf{q}} \left[ e^{i\mathbf{q}\cdot\mathbf{x} - iE_{\mathbf{q}}x_0} \sum_h u_r^{(+)}(\mathbf{q}, h) a_+(\mathbf{q}, h) + e^{i\mathbf{q}\cdot\mathbf{x} + iE_{\mathbf{q}}x_0} \sum_h u_r^{(-)}(\mathbf{q}, h) a_-(\mathbf{q}, h) \right], \quad (3.176)$$

where  $E_q = \sqrt{q^2 + m^2}$ ,  $q = |\mathbf{q}|$ , and  $u^{(+)}$  and  $u^{(-)}$  are the positive and negative energy solutions, respectively, satisfying the equations

$$(\boldsymbol{\alpha} \cdot \mathbf{q} + \beta m)u^{(\pm)}(\mathbf{q}, h) = \pm E_q u^{(\pm)}(\mathbf{q}, h). \quad (3.177)$$

Substituting the Fourier expansion above into the Hamiltonian (3.175), we obtain

$$H = \sum_{\mathbf{q}} E_q \left[ \sum_h a_+^\dagger(\mathbf{q}, h) a_+(\mathbf{q}, h) - \sum_h a_-^\dagger(\mathbf{q}, h) a_-(\mathbf{q}, h) \right]. \quad (3.178)$$

In addition, the following conserved quantity (which is in fact proportional to the electric charge) is also important:

$$N = \int d^3x j_0(x) = \sum_{\mathbf{q}} \sum_h \left[ a_+^\dagger(\mathbf{q}, h) a_+(\mathbf{q}, h) + a_-^\dagger(\mathbf{q}, h) a_-(\mathbf{q}, h) \right]. \quad (3.179)$$

Next, substituting the Fourier expansion (3.176) into (3.173), we find

$$\begin{aligned} [a_+(\mathbf{q}, h), a_+^\dagger(\mathbf{q}', h')]_+ &= \delta_{h, h'} \delta_{\mathbf{q}, \mathbf{q}'}, \\ [a_-(\mathbf{q}, h), a_-^\dagger(\mathbf{q}', h')]_+ &= \delta_{h, h'} \delta_{\mathbf{q}, \mathbf{q}'}, \\ \text{other anti-commutators} &= 0. \end{aligned} \quad (3.180)$$

Looking at the expression (3.178) for the Hamiltonian, the energy in the system is apparently not positive-definite.

Although the question regarding the negative energy was originally considered as a drawback of the theory, Dirac was able to get round it by assuming that the vacuum is the state filled completely with negative energy states, taking advantage of the fact that the electron obeys Fermi statistics. However, this idea is unnatural if we think of, say, the electric field created by the infinite number of electrons in the negative energy states. A much clearer solution is achieved by quantizing the field. To explain this, we use the following trick:

$$- \sum_h a_-^\dagger(\mathbf{q}, h) a_-(\mathbf{q}, h) = \sum_h a_-(\mathbf{q}, h) a_-^\dagger(\mathbf{q}, h) - 2. \quad (3.181)$$

Neglecting the constant 2, as we did in (3.66), the Hamiltonian becomes positive-definite:

$$H = \sum_{\mathbf{q}} E_q \left[ \sum_h a_+^\dagger(\mathbf{q}, h) a_+(\mathbf{q}, h) + \sum_h a_-(\mathbf{q}, h) a_-^\dagger(\mathbf{q}, h) \right]. \quad (3.182)$$

Although the term we neglected is negative infinity, it is still a constant, viz.,

$$-2 \sum_{\mathbf{q}} E_{\mathbf{q}} . \quad (3.183)$$

This corresponds to neglecting the total energy of the electrons in the negative energy states in Dirac's hole theory. Let us reconstruct the interpretation in the hole theory by reinterpreting the operators.  $a_{+}^{\dagger}$  and  $a_{+}$  are the operators creating and annihilating, respectively, the electron with positive energy (the Dirac particle), while  $a_{-}^{\dagger}$  and  $a_{-}$  are the operators creating and annihilating, respectively, particles with negative energy. However, creating the electron with negative energy is equivalent to annihilating a particle with positive energy. In addition, creating some particle with an additive quantum number is equivalent to annihilating another with the opposite quantum number. Here we consider that creating the electron with negative energy is equivalent to annihilating the positive-energy particle with the opposite quantum numbers. We make the replacement

$$\begin{aligned} a_{+}(\mathbf{q}, h) &\rightarrow a(\mathbf{q}, h) , & a_{+}^{\dagger}(\mathbf{q}, h) &\rightarrow a^{\dagger}(\mathbf{q}, h) , \\ a_{-}(\mathbf{q}, h) &\rightarrow b^{\dagger}(-\mathbf{q}, h) , & a_{-}^{\dagger}(\mathbf{q}, h) &\rightarrow b(-\mathbf{q}, h) . \end{aligned} \quad (3.184)$$

The helicity is a product of the additive quantum numbers  $\mathbf{p}$  and  $\boldsymbol{\sigma}$ , so its sign is unchanged.

Since the particle whose additive quantum numbers all have opposite sign to the electron has a positive electric charge, it is called a *positron*. We define the operators

$$N_{+}(\mathbf{q}, h) = a^{\dagger}(\mathbf{q}, h)a(\mathbf{q}, h) , \quad N_{-}(\mathbf{q}, h) = b^{\dagger}(\mathbf{q}, h)b(\mathbf{q}, h) . \quad (3.185)$$

Hence,  $H$  and  $N$  can be written as

$$H = \sum_{\mathbf{q}} \sum_h E_{\mathbf{q}} [N_{+}(\mathbf{q}, h) + N_{-}(\mathbf{q}, h)] , \quad (3.186)$$

$$N = \sum_{\mathbf{q}} \sum_h [N_{+}(\mathbf{q}, h) - N_{-}(\mathbf{q}, h)] , \quad (3.187)$$

where  $b$  and  $b^{\dagger}$  are the creation and annihilation operators of the positron, respectively.  $N_{\pm}$  are the particle number operators and, unlike the case of Bose statistics, their eigenvalues can only take the values 0 and 1, a feature of Fermi statistics. This can be shown using the anti-commutation relations (3.180). Fixing  $\mathbf{q}$  and  $h$ ,

$$N^2 = a^{\dagger} a a^{\dagger} a = a^{\dagger} (1 - a^{\dagger} a) a = a^{\dagger} a - (a^{\dagger})^2 a^2 = a^{\dagger} a = N ,$$

i.e., from the anti-commutation relations, we find

$$[N_{\pm}(\mathbf{q}, h)]^2 = N_{\pm}(\mathbf{q}, h) . \quad (3.188)$$

This means that the only permissible eigenvalues are 0 and 1, reflecting Pauli's exclusion principle, namely that one state can be occupied by at most one particle. For this reason, the field of particles obeying Fermi statistics must be quantized using anti-commutation relations.

Using the new operators, the Fourier expansion of the Dirac field can be written as

$$\psi_r(x) = \frac{1}{\sqrt{V}} \sum_{\mathbf{q}} \sum_h \left[ e^{iq \cdot x} u_r(\mathbf{q}, h) a(\mathbf{q}, h) + e^{-iq \cdot x} v_r(\mathbf{q}, h) b^\dagger(\mathbf{q}, h) \right] , \quad (3.189)$$

where  $q \cdot x = \mathbf{q} \cdot \mathbf{x} - E_{\mathbf{q}} x_0$  is the four-dimensional scalar product, and we have simply written  $\mathbf{q}$  in  $a$  or  $b$  as  $q$ . In addition, we have used the notation

$$u_r(\mathbf{q}, h) = u_r^{(+)}(\mathbf{q}, h) , \quad v_r(\mathbf{q}, h) = u_r^{(-)}(-\mathbf{q}, h) . \quad (3.190)$$

### Casimir Operator

When experiments are carried out, very often only the momentum and energy are measured, and not the spin or the helicity. In such cases, we must sum over the spin states, and for this purpose, it is useful to define the so-called *Casimir operator*. We shall write the electron and the positron, or more generally the particle state with four-momentum  $p$  and the antiparticle state with four-momentum  $\bar{p}$ , as  $|p, h\rangle$  and  $|\bar{p}, h\rangle$ , respectively. Here  $h$  denotes the helicity. In this case, from the Fourier expansion, we obtain

$$\langle 0 | \psi(x) | p, h \rangle = \frac{1}{\sqrt{V}} e^{ip \cdot x} u(p, h) , \quad (3.191a)$$

$$\langle p, h | \bar{\psi}(x) | 0 \rangle = \frac{1}{\sqrt{V}} e^{-ip \cdot x} \bar{u}(p, h) , \quad (3.191b)$$

where  $|0\rangle$  stands for the vacuum state. For the antiparticle state, we find

$$\langle 0 | \bar{\psi}(x) | \bar{p}, h \rangle = \frac{1}{\sqrt{V}} e^{i\bar{p} \cdot x} \bar{v}(\bar{p}, h) , \quad (3.192a)$$

$$\langle \bar{p}, h | \psi(x) | 0 \rangle = \frac{1}{\sqrt{V}} e^{-i\bar{p} \cdot x} v(\bar{p}, h) . \quad (3.192b)$$

From the orthogonality of the solutions of (3.147) and the normalization of the solution based on (3.155), we have

$$\begin{aligned} \sum_s \left[ \sum_{h'} u_r^{(+)}(\mathbf{p}, h') u_s^{(+)*}(\mathbf{p}, h') \right] u_s^{(+)}(\mathbf{p}, h') &= u_r^{(+)}(\mathbf{p}, h) , \\ \sum_s \left[ \sum_{h'} u_r^{(+)}(\mathbf{p}, h') u_s^{(+)*}(\mathbf{p}, h') \right] u_s^{(-)}(\mathbf{p}, h') &= 0 . \end{aligned} \quad (3.193)$$

These relations tell us that the term in the brackets is equivalent to the projection operator which selects the positive energy solution. Hence,

$$\sum_{h'} u_r^{(+)}(\mathbf{p}, h') u_s^{(+)*}(\mathbf{p}, h') = \left( \frac{H(p) + E_p}{2E_p} \right)_{rs} = \left( \frac{\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m + E_p}{2E_p} \right)_{rs} . \quad (3.194)$$

We rewrite the above equation as

$$\sum_h u_r(p, h) \bar{u}_s(p, h) = \left( \frac{-i\mathbf{p} \cdot \boldsymbol{\gamma} + m}{2E_p} \right)_{rs} . \quad (3.195)$$

Similarly, for the negative energy solution,

$$\sum_{h'} u_r^{(-)}(-\mathbf{p}, h') u_s^{(-)*}(-\mathbf{p}, h') = \left( \frac{-\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m - E_p}{-2E_p} \right)_{rs} , \quad (3.196)$$

i.e.,

$$\sum_h v_r(\bar{\mathbf{p}}, h) \bar{v}_s(\bar{\mathbf{p}}, h) = \left( \frac{-i\bar{\mathbf{p}} \cdot \boldsymbol{\gamma} - m}{2E_p} \right)_{rs} . \quad (3.197)$$

This projection operator is called Casimir operator. Using this operator we can perform all the calculations without using an explicit representation of the Dirac spinor.

### 3.9 Charge Conjugation

In the quantization of the complex scalar field and the Dirac field, the concept of particle and antiparticle arises. We will now focus on the symmetry between them. Such a symmetry always exists in the case of free fields. Since exchanging a particle with an antiparticle means a flip in the sign of the electric charge, such a

transformation is called *charge conjugation* and, following (3.78), it is denoted by  $\mathcal{C}$ . For the Dirac field, the  $\mathcal{C}$ -operation interchanges  $\psi$  and  $\bar{\psi}$ , i.e., in general,

$$\psi_r(x) \rightarrow \mathcal{C}^{-1} \psi_r(x) \mathcal{C} = C_{rs} \bar{\psi}_s(x), \quad (3.198)$$

where  $C$  is a  $4 \times 4$  matrix. In addition, we have summed over the spinor index  $s$ . In the following, we will study the properties of the matrix  $C$ .

First, we consider that four-momentum and spin or helicity are unchanged when interchanging a particle and an antiparticle. From this property and (3.187), we obtain

$$\mathcal{C}^{-1} N \mathcal{C} = -N. \quad (3.199)$$

Thus, the same equation holds for the density as well:

$$\mathcal{C}^{-1} j_\mu \mathcal{C} = -j_\mu. \quad (3.200)$$

In quantized theories, it is more convenient to define  $j_\mu(x)$  by

$$j_\mu(x) = i\bar{\psi}(x)\gamma_\mu\psi(x) \rightarrow \frac{i}{2}[\bar{\psi}(x), \gamma_\mu\psi(x)] = \frac{i}{2}(\gamma_\mu)_{rs}[\bar{\psi}_r(x), \psi_s(x)]. \quad (3.201)$$

With this redefinition, the vacuum expectation value of  $j_\mu$  vanishes.

Then we write the Dirac equation (3.168) as

$$(\gamma \cdot \partial + m)\psi = 0, \quad (\gamma^T \cdot \partial - m)\bar{\psi}^T = 0, \quad (3.202)$$

where  $\gamma^T$  stands for the transpose of the matrix  $\gamma_\mu$ . If Dirac's equation does not change form under the charge conjugation operation, then

$$(\gamma \cdot \partial + m)\mathcal{C}^{-1}\psi\mathcal{C} = (\gamma \cdot \partial + m)C\bar{\psi}^T = 0.$$

Consequently,

$$(C^{-1}\gamma_\mu C \cdot \partial_\mu + m)\bar{\psi}^T = 0. \quad (3.203)$$

This equation is equivalent to the second equation in (3.202) if

$$C^{-1}\gamma_\mu C = -\gamma_\mu^T. \quad (3.204)$$

Clearly, the matrices  $-\gamma_\mu^T$  satisfy the commutation relations (3.108), just like  $\gamma_\mu$ . We skip the details, but given an explicit representation of the  $\gamma$ -matrices, the one obtained by replacing  $\gamma_\mu$  by  $-\gamma_\mu^T$  is also a representation of the Dirac matrices and

it is equivalent to the original one. As a result, we can choose the matrix  $C$  as a unitary matrix. In addition, from (3.204), we have

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu . \quad (3.205)$$

The equations derived by transposing (3.204) are

$$\begin{aligned} C^T \gamma_\mu^T (C^T)^{-1} &= -\gamma_\mu , \\ (C^T)^{-1} \gamma_\mu C^T &= -\gamma_\mu^T . \end{aligned} \quad (3.206)$$

Substituting the right-hand side of the latter relation into the left-hand side of (3.205), we find

$$(C^T C^{-1}) \gamma_\mu (C^T C^{-1})^{-1} = \gamma_\mu . \quad (3.207)$$

All  $4 \times 4$  matrices can be expressed as polynomials of the four  $\gamma$ -matrices. The equation above shows that  $C^T C^{-1}$  commutes with all the  $\gamma$ -matrices, whence it commutes with all the  $4 \times 4$  matrices. Such a matrix must therefore be proportional to the unit matrix:

$$C^T C^{-1} = a \mathbf{1} . \quad (3.208)$$

The coefficient  $a$  does not depend on which representation we choose. This can be proved as follows.

Considering a unitary matrix  $S$ , we change the representation of  $\gamma$  as follows:

$$\gamma_\mu' = S \gamma_\mu S^{-1} . \quad (3.209)$$

In this case, solving (3.204) in the new representation, the matrix  $C'$  will be given by

$$C' = S C S^T . \quad (3.210)$$

Therefore,

$$(C')^T (C')^{-1} = S C^T S^T (S^T)^{-1} C^{-1} S^{-1} = a \mathbf{1} = C^T C^{-1} .$$

Since the coefficient  $a$  is shared by all the equivalent representations, it is enough to determine it for one representation, say the Pauli representation. In this representation,

$$\gamma_\mu^T = \gamma_\mu , \quad \mu = 2, 4 , \quad \gamma_\mu^T = -\gamma_\mu , \quad \mu = 1, 3 , \quad (3.211)$$

and (3.204) is satisfied by choosing  $C = \gamma_2\gamma_4$ . In this case,

$$\begin{aligned} C^T &= (\gamma_2\gamma_4)^T = \gamma_4^T \gamma_2^T = \gamma_4\gamma_2 = -\gamma_2\gamma_4 = -C, \\ C^T C^{-1} &= -1. \end{aligned} \quad (3.212)$$

Thus, we have found that  $a = -1$  and the matrix  $C$  is anti-symmetric.

The charge conjugation transformation of  $\bar{\psi}$  can be derived from the charge conjugation of  $\psi$ . Although we skip the intermediate calculations, using (3.198), (3.204), the unitarity of  $C$ , and the hermiticity of  $\gamma_\mu$ , we find

$$\mathcal{C}^{-1} \bar{\psi}_r \mathcal{C} = (C^{-1})_{rs} \psi_s(x). \quad (3.213)$$

### Charge Conjugation of the Bilinear Forms

Since the  $\gamma$ -matrices are four-dimensional matrices, there exist 16 linearly independent polynomials that can be created using them. Denoting a four-dimensional matrix by  $O$ , we consider the bilinear form

$$[\bar{\psi}(x), O\psi(x)] = O_{rs} [\bar{\psi}_r(x), \psi_s(x)]. \quad (3.214)$$

From the transformation property of this bilinear form under Lorentz transformations, the matrix  $O$  can be categorized into the following five types:

$$\begin{aligned} &1 \text{ (scalar)}, \quad \gamma_\mu \text{ (vector)}, \quad \sigma_{\mu\nu} \text{ (anti-symmetric tensor)}, \\ &\gamma_\mu\gamma_5 \text{ (axial vector)}, \quad \gamma_5 \text{ (pseudo scalar)}. \end{aligned} \quad (3.215)$$

The number of matrices in all categories is  $1 + 4 + 6 + 4 + 1 = 16$ , so they exhaust the linearly independent  $4 \times 4$  matrices.

We now examine the transformation properties of the bilinear forms (3.214) under charge conjugation. Combining (3.198) and (3.213),

$$\mathcal{C}^{-1} [\bar{\psi}(x), O\psi(x)] \mathcal{C} = [\bar{\psi}(x), O'\psi(x)], \quad (3.216)$$

where

$$O' = (C^{-1}OC)^T = CO^T C^{-1} \equiv \epsilon O. \quad (3.217)$$

Calculating  $\epsilon$  for the various cases using (3.205), we obtain Table 3.1.

**Table 3.1** Sign changes under charge conjugation

$O$	1	$\gamma_\mu$	$\sigma_{\mu\nu}$	$\gamma_\mu\gamma_5$	$\gamma_5$
$\epsilon$	1	-1	-1	1	1



### 3.10 Quantization of the Complex Vector Field

So far we have studied the quantization of the spin-0 scalar field and the spin-1/2 spinor field, and in the following we will examine the quantization of a vector field. A real vector field and a complex vector field do not differ essentially, so we consider the latter.

We denote a complex vector field and its conjugate by  $\varphi_\mu$  and  $\varphi_\mu^\dagger$ , respectively. As quantized operators, they are Hermitian conjugates of each other. When using the Pauli metric convention, as we do now, the  $i$  attributed to quantum theory changes its sign under the Hermitian conjugation operator  $\dagger$ , but the one attributed to relativity theory does not change its sign. Therefore,  $\varphi_j$  and  $\varphi_j^\dagger$  are complex conjugates for  $j = 1, 2, 3$ , but when we write

$$\varphi_4 = i\varphi_0, \quad \varphi_4^\dagger = i\varphi_0^\dagger, \quad (3.218)$$

it turns out that  $\varphi_0$  and  $\varphi_0^\dagger$  are complex conjugates, and  $\varphi_4$  and  $\varphi_4^\dagger$  are not. Bearing this in mind, we will discuss the conditions that a free vector field must satisfy:

$$K\varphi_\mu \equiv (\square - m^2)\varphi_\mu = 0, \quad (3.219)$$

$$\partial_\mu\varphi_\mu = 0. \quad (3.220)$$

The first is the Klein–Gordon equation for each component. The second is called the *irreducibility condition*. If this four-divergence does not vanish, it represents a scalar field, and then  $\varphi_\mu$  represents a field describing a mixture of spin-0 and spin-1 particles.

Let us derive the Lagrangian density satisfying the two conditions above. If we try the form

$$\mathcal{L} = -(\partial_\nu\varphi_\mu^\dagger \cdot \partial_\nu\varphi_\mu + m^2\varphi_\mu^\dagger \cdot \varphi_\mu),$$

by analogy with the scalar field Lagrangian, we get into trouble, because although (3.219) holds, (3.220) does not. In addition, the Hamiltonian density derived from this Lagrangian density is not positive-definite, as can be seen here:

$$\begin{aligned} \mathcal{H} &= \pi_\mu^\dagger\pi_\mu + \nabla\varphi_\mu^\dagger \cdot \nabla\varphi_\mu + m^2\varphi_\mu^\dagger\varphi_\mu \\ &= \sum_{j=1}^3 \left( \pi_j^\dagger\pi_j + \nabla\varphi_j^\dagger \cdot \nabla\varphi_j + m^2\varphi_j^\dagger\varphi_j \right) - \left( \pi_0^\dagger\pi_0 + \nabla\varphi_0^\dagger \cdot \nabla\varphi_0 + m^2\varphi_0^\dagger\varphi_0 \right). \end{aligned}$$

We have to find another form, and we will try the one suggested by the Lagrangian of the electromagnetic field:

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\varphi_\nu^\dagger - \partial_\nu\varphi_\mu^\dagger)(\partial_\mu\varphi_\nu - \partial_\nu\varphi_\mu) - m^2\varphi_\mu^\dagger\varphi_\mu. \quad (3.221)$$

Additionally, we define the field strength  $F_{\mu\nu}$  by

$$F_{\mu\nu} = \partial_\mu \varphi_\nu - \partial_\nu \varphi_\mu . \quad (3.222)$$

The Euler–Lagrange equation derived from the variational principle is

$$\frac{\partial \mathcal{L}}{\partial \varphi_\nu^\dagger} - \frac{\partial}{\partial x_\mu} \left( \frac{\partial \mathcal{L}}{\partial \varphi_{\nu,\mu}^\dagger} \right) = -m^2 \varphi_\nu + \partial_\mu F_{\mu\nu} = 0 ,$$

whence

$$\partial_\mu F_{\mu\nu} = m^2 \varphi_\nu , \quad \partial_\mu F_{\mu\nu}^\dagger = m^2 \varphi_\nu^\dagger . \quad (3.223)$$

In the following, we assume  $m \neq 0$ . Since  $F_{\mu\nu}$  is an anti-symmetric tensor,

$$m^2 \partial_\nu \varphi_\nu = \partial_\nu \partial_\mu F_{\mu\nu} = 0 .$$

Thus, from the assumption that  $m \neq 0$ , we can derive (3.220). Additionally, using (3.220) in (3.223), we find that (3.219) is also satisfied:

$$\partial_\mu (\partial_\mu \varphi_\nu - \partial_\nu \varphi_\mu) = \square \varphi_\nu = m^2 \varphi_\nu .$$

Next, we define  $\pi$  as follows:

$$\begin{aligned} \pi_\nu^\dagger &= \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\nu} = \frac{1}{i} (\partial_\nu \varphi_4^\dagger - \partial_4 \varphi_\nu^\dagger) = \frac{1}{i} F_{\nu 4}^\dagger , \\ \pi_\nu &= \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\nu^\dagger} = \frac{1}{i} (\partial_\nu \varphi_4 - \partial_4 \varphi_\nu) = \frac{1}{i} F_{\nu 4} . \end{aligned} \quad (3.224)$$

Here we encounter a formal difficulty, i.e.,

$$\pi_4 = \pi_4^\dagger = 0 . \quad (3.225)$$

It turns out that we cannot define the canonically conjugate momentum for the fourth component. However, noticing that the fourth component is not independent of the other three components, it becomes clear that this is not a fundamental difficulty. From the equation of motion (3.223) and the definition of the canonically conjugate momentum (3.224), we see how the fourth component can be expressed in terms of the other components:

$$\varphi_4 = \frac{i}{m^2} \operatorname{div} \boldsymbol{\pi} , \quad \varphi_4^\dagger = \frac{i}{m^2} \operatorname{div} \boldsymbol{\pi}^\dagger . \quad (3.226)$$

In addition, the time derivative of  $\varphi_j$  can be written in terms of the three components  $\pi_j$  as follows:

$$\dot{\varphi}_j = \pi_j + i \frac{\partial \varphi_4}{\partial x_j}, \quad \dot{\varphi}_j^\dagger = \pi_j^\dagger + i \frac{\partial \varphi_4^\dagger}{\partial x_j}. \quad (3.227)$$

Thus, considering only the three space components as independent quantities, the Hamiltonian density can be written as

$$\mathcal{H} = \sum_{j=1}^3 (\pi_j^\dagger \dot{\varphi}_j + \pi_j \dot{\varphi}_j^\dagger) - \mathcal{L}. \quad (3.228)$$

However, to construct the Hamiltonian we use

$$H = \int d^3x \mathcal{H}(x) = \int d^3x \mathcal{H}'(x), \quad (3.229)$$

where  $\mathcal{H}'$  is defined by adding a three-dimensional divergence term to  $\mathcal{H}$ :

$$\mathcal{H}' = \mathcal{H} - i \operatorname{div}(\boldsymbol{\pi}^\dagger \varphi_4 + \boldsymbol{\pi} \varphi_4^\dagger). \quad (3.230)$$

The advantage of using  $\mathcal{H}'$  is that it makes manifest the positivity of the Hamiltonian:

$$\mathcal{H}' = \boldsymbol{\pi}^\dagger \cdot \boldsymbol{\pi} + \frac{1}{m^2} (\operatorname{div} \boldsymbol{\pi}^\dagger)(\operatorname{div} \boldsymbol{\pi}) + m^2 \boldsymbol{\varphi}^\dagger \cdot \boldsymbol{\varphi} + (\operatorname{rot} \boldsymbol{\varphi}^\dagger) \cdot (\operatorname{rot} \boldsymbol{\varphi}). \quad (3.231)$$

### Canonical Quantization and Canonical Equations

Starting from the Lagrangian density (3.221), we introduce the following equal-time commutation relations:

$$[\pi_j^\dagger(x), \varphi_k(x')] = [\pi_j(x), \varphi_k^\dagger(x')] = -i \delta_{jk} \delta^3(x - x'), \quad \text{where } x_0 = x'_0. \quad (3.232)$$

All other commutators are set to zero. Then we derive the canonical equations:

$$\begin{aligned} \dot{\varphi}_j(x) &= i \int d^3x' [\mathcal{H}(x'), \varphi_j(x)] \\ &= \int d^3x' \left[ \pi_j(x') \delta^3(x - x') + \frac{1}{m^2} \operatorname{div} \boldsymbol{\pi}(x') \frac{\partial}{\partial x'_j} \delta^3(x - x') \right], \end{aligned}$$

whence

$$\dot{\boldsymbol{\varphi}}(x) = \boldsymbol{\pi}(x) - \frac{1}{m^2} \text{grad}(\text{div } \boldsymbol{\pi}) . \quad (3.233)$$

Similarly, we obtain

$$\dot{\boldsymbol{\pi}}(x) = -m^2 \boldsymbol{\varphi}(x) - \text{rot rot } \boldsymbol{\varphi}(x) . \quad (3.234)$$

Combining the two equations above and using (3.226), we can derive the equation of motion (3.219) for  $\varphi_\mu$ . In addition, differentiating (3.226) with respect to time and combining the result with equation (3.224), we obtain the condition (3.220). Thus, the canonical equations (3.233) and (3.234) are equivalent to the original equations.

We can also derive the important equal-time commutation relation

$$\begin{aligned} [\dot{\varphi}_j(x), \varphi_k^\dagger(x')] &= \left[ \pi_j(x) - \frac{1}{m^2} \frac{\partial}{\partial x_j} \text{div } \boldsymbol{\pi}(x), \varphi_k^\dagger(x') \right] \\ &= -i \left( \delta_{jk} - \frac{1}{m^2} \frac{\partial^2}{\partial x_j \partial x_k} \delta^3(x - x') \right) . \end{aligned} \quad (3.235)$$

### Fourier Expansion

Similarly to what we have done so far, we introduce the Fourier expansion for vector fields:

$$\begin{aligned} \varphi_j(x) &= \sum_{\mathbf{p}} \frac{1}{\sqrt{2p_0V}} \left[ e^{i\mathbf{p}\cdot\mathbf{x}} a_j(\mathbf{p}) + e^{-i\mathbf{p}\cdot\mathbf{x}} b_j^\dagger(\mathbf{p}) \right] , \\ \varphi_j^\dagger(x) &= \sum_{\mathbf{p}} \frac{1}{\sqrt{2p_0V}} \left[ e^{i\mathbf{p}\cdot\mathbf{x}} b_j(\mathbf{p}) + e^{-i\mathbf{p}\cdot\mathbf{x}} a_j^\dagger(\mathbf{p}) \right] , \quad j = 1, 2, 3 . \end{aligned} \quad (3.236)$$

We can write the Fourier expansions for  $\varphi_4$  and  $\varphi_4^\dagger$  in terms of  $a_j$  and  $b_j$  above because they can be expressed in terms of the other components, but here we skip this.

Following the same method as before, we can determine the commutation relations between coefficients. Here we only give the result:

$$[a_j(\mathbf{p}), a_k^\dagger(\mathbf{q})] = [b_j(\mathbf{p}), b_k^\dagger(\mathbf{q})] = \delta_{\mathbf{p},\mathbf{q}} \left( \delta_{jk} + \frac{p_j p_k}{m^2} \right) . \quad (3.237)$$

All other commutators are zero. Then, including the fourth component, we rewrite the Fourier expansions:

$$\begin{aligned}\varphi_\mu(x) &= \sum_{\mathbf{p}} \frac{1}{\sqrt{2p_0V}} \left[ e^{ip \cdot x} a_\mu(\mathbf{p}) + e^{-ip \cdot x} b_\mu^\dagger(\mathbf{p}) \right], \\ \varphi_\mu^\dagger(x) &= \sum_{\mathbf{p}} \frac{1}{\sqrt{2p_0V}} \left[ e^{ip \cdot x} b_\mu(\mathbf{p}) + e^{-ip \cdot x} a_\mu^\dagger(\mathbf{p}) \right].\end{aligned}\quad (3.238)$$

Using the irreducibility condition (3.220), we find that

$$p_\mu a_\mu(\mathbf{p}) = p_\mu a_\mu^\dagger(\mathbf{p}) = p_\mu b_\mu(\mathbf{p}) = p_\mu b_\mu^\dagger(\mathbf{p}) = 0. \quad (3.239)$$

From these relations, it turns out that the zeroth component of  $a$  and  $b$  can be expressed in terms of the other three components:

$$a_0(\mathbf{p}) = \frac{p_j}{p_0} a_j(\mathbf{p}), \quad \text{and similar equations.} \quad (3.240)$$

Combining (3.237) with the above equation,

$$[a_\mu(\mathbf{p}), a_\nu^\dagger(\mathbf{q})] = [b_\mu(\mathbf{p}), b_\nu^\dagger(\mathbf{q})] = \delta_{\mathbf{p},\mathbf{q}} \left( \delta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2} \right). \quad (3.241)$$

All other commutation relations are again zero.

We choose three unit vectors orthogonal to each other as

$$\mathbf{e}_1 = \mathbf{p}/p, \quad \mathbf{e}_2, \quad \mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2, \quad (3.242)$$

and we expand  $\mathbf{a}(\mathbf{p})$  as

$$\mathbf{a}(\mathbf{p}) = \mathbf{e}_1 a^{(1)}(\mathbf{p}) + \mathbf{e}_2 a^{(2)}(\mathbf{p}) + \mathbf{e}_3 a^{(3)}(\mathbf{p}). \quad (3.243)$$

In this new coordinate system,

$$\begin{aligned}[a^{(1)}(\mathbf{p}), a^{(1)\dagger}(\mathbf{p})] &= 1 + \frac{p^2}{m^2} = \frac{p_0^2}{m^2}, \\ [a^{(2)}(\mathbf{p}), a^{(2)\dagger}(\mathbf{p})] &= [a^{(3)}(\mathbf{p}), a^{(3)\dagger}(\mathbf{p})] = 1,\end{aligned}\quad (3.244)$$

where  $a^{(1)}$  corresponds to the longitudinal component and  $a^{(2)}$  and  $a^{(3)}$  to the transverse components. Here we change the definition of  $a^{(j)}(\mathbf{p})$ . Instead of (3.243), we set

$$\mathbf{a}(\mathbf{p}) = \mathbf{e}_1 \left( \frac{p_0}{m} \right) a^{(1)}(\mathbf{p}) + \mathbf{e}_2 a^{(2)}(\mathbf{p}) + \mathbf{e}_3 a^{(3)}(\mathbf{p}). \quad (3.245)$$

Consequently, this new operator satisfies the simple commutation relation

$$[a^{(j)}(\mathbf{p}), a^{(k)}(\mathbf{p})] = \delta_{jk} . \quad (3.246)$$

Furthermore, using this new operator, the energy and momentum of the field are

$$H = \sum_{\mathbf{p}} \sum_j p_0 [n_+^{(j)}(\mathbf{p}) + n_-^{(j)}(\mathbf{p})] , \quad (3.247)$$

$$P_k = \sum_{\mathbf{p}} \sum_j p_k [n_+^{(j)}(\mathbf{p}) + n_-^{(j)}(\mathbf{p})] , \quad (3.248)$$

$$n_+^{(j)}(\mathbf{p}) = a^{(j)\dagger}(\mathbf{p})a^{(j)}(\mathbf{p}) , \quad n_-^{(j)}(\mathbf{p}) = b^{(j)\dagger}(\mathbf{p})b^{(j)}(\mathbf{p}) , \quad (3.249)$$

where we have assumed that the momentum  $\mathbf{P}$  is given by interpreting  $\alpha$  as the three components of the vector field in (3.30).  $n_+$  and  $n_-$  are the particle number operators. Their eigenvalues are non-negative integers, 0, 1, 2, ..., reflecting the Bose statistics.

In addition, the operation of charge conjugation connects  $\varphi_\mu(x)$  and  $\varphi_\mu^\dagger(x)$  by a relation similar to (3.78).

In this chapter, considering the fields with spin 0, 1/2, and 1 as quantum-mechanical operators, we have quantized them by using the canonical commutation relations. A field satisfying the so-called *wave equation* as equation of motion can be considered as a classical wave. However, as a consequence of the quantization, the eigenvalues of its energy and momentum coincide with those for particles of the same mass, which implies that fields possess particle-like aspects. Intuitively speaking, in the free state of motion, wave-like aspects described by the wave equation show up strongly, and when the state of motion is changed via interactions, the change in the four-momentum has particle-like aspects.

The dual particle–wave nature arising in quantum mechanics can be understood only after quantizing fields.

# Chapter 4

## Invariant Functions and Quantization of Free Fields



In the last chapter, we quantized free fields using the canonical quantization procedure. However, we did this by fixing a time axis, and we should ask whether the result is independent of this choice. To answer this question, we need to study the unequal-time commutation relations. By introducing several kinds of Lorentz-invariant Green functions, we will show that free fields can be quantized without fixing a direction of time.

### 4.1 Unequal-Time Commutation Relations for Real Scalar Fields

Using the previously obtained Fourier expansion for the real scalar field, one can derive the commutation relations among two fields at unequal times. Using (3.51), (3.52), and (3.53),

$$\begin{aligned} [\varphi(x), \varphi(y)] &= \sum_{\mathbf{p}} \sum_{\mathbf{q}} \frac{1}{\sqrt{2p_0V}} \frac{1}{\sqrt{2q_0V}} \left( e^{ip \cdot x - iq \cdot y} [a(\mathbf{p}), a^\dagger(\mathbf{q})] \right. \\ &\quad \left. + e^{-ip \cdot x + iq \cdot y} [a^\dagger(\mathbf{p}), a(\mathbf{q})] \right) \\ &= \frac{1}{2p_0V} \left[ e^{ip \cdot (x-y)} - e^{-ip \cdot (x-y)} \right]. \end{aligned} \tag{4.1}$$

It is clear from (3.38) that the possible values of the momentum in a cube with fixed volume  $V$  are quantized by periodic conditions. Then, when  $V$  or  $L$  is large, the sum over momenta is replaced by an integral, i.e.,

$$\sum_p \rightarrow \frac{V}{(2\pi)^3} \int d^3 p . \quad (4.2)$$

Making this replacement in (4.1), the right-hand side becomes

$$\frac{1}{(2\pi)^3} \int \frac{d^3 p}{2p_0} \left[ e^{ip \cdot (x-y)} - e^{-ip \cdot (x-y)} \right] \equiv i\Delta(x-y) . \quad (4.3)$$

This function was introduced by Pauli and Jordan [80], but the sign is different from the original, i.e.,

$$\Delta(x) = -\frac{i}{(2\pi)^3} \int \frac{d^3 p}{2p_0} \left( e^{ip \cdot x} - e^{-ip \cdot x} \right) . \quad (4.4)$$

In the following, we list some properties of this function.

1.  $\Delta(x)$  is a Lorentz-invariant function, i.e., for the proper Lorentz transformation,

$$x_\mu \rightarrow x'_\mu = \sum_\nu a_{\mu\nu} x_\nu ,$$

the following equality holds:

$$\Delta(x') = \Delta(x) . \quad (4.5)$$

**Proof**  $\delta(p^2 + m^2)$  is clearly Lorentz invariant, and using the formula

$$\delta(ab) = |b|^{-1} \delta(a) + |a|^{-1} \delta(b) ,$$

one finds

$$\delta(p^2 + m^2) = \frac{1}{2\sqrt{p^2 + m^2}} \left[ \delta \left( p_0 - \sqrt{p^2 + m^2} \right) + \delta \left( p_0 + \sqrt{p^2 + m^2} \right) \right] . \quad (4.6)$$



The first term is nonzero in the region  $p_0 \geq m$  and the second in the region  $p_0 \leq -m$ . Since these two regions do not overlap, each term is invariant under the proper Lorentz transformation. We now introduce the following discontinuous functions:

$$\theta(p_0) = \begin{cases} 1 & \text{for } p_0 > 0, \\ 0 & \text{for } p_0 < 0, \end{cases} \quad (4.7)$$

$$\epsilon(p_0) = \theta(p_0) - \theta(-p_0) = p_0/|p_0|. \quad (4.8)$$

From the above, we note that both  $\theta(p_0)\delta(p^2 + m^2)$  and  $\epsilon(p_0)\delta(p^2 + m^2)$  are Lorentz invariant, whence the following function is also Lorentz invariant:

$$\begin{aligned} & \int d^4 p \epsilon(p_0) \delta(p^2 + m^2) e^{ip \cdot x} \\ &= \frac{1}{2} \int \frac{dp_0 d^3 p}{\sqrt{p^2 + m^2}} \left[ \delta\left(p_0 - \sqrt{p^2 + m^2}\right) - \delta\left(p_0 + \sqrt{p^2 + m^2}\right) \right] e^{ipx - ip_0 x_0}. \end{aligned}$$

Integrating this equation over  $p_0$  and inverting the integration variable according to  $p \rightarrow -p$  in the second term, the above equation becomes

$$\int d^4 p \epsilon(p_0) \delta(p^2 + m^2) e^{ip \cdot x} = \int \frac{d^3 p}{2p_0} \left( e^{ip \cdot x} - e^{-ip \cdot x} \right).$$

In this equation,  $p_0$  stands for  $\sqrt{p^2 + m^2}$ , and so is positive-definite. This equation coincides with (4.4) up to the coefficient. Therefore,  $\Delta(x)$  is Lorentz invariant.

2.  $\Delta(x)$  is an odd function, i.e.,

$$\Delta(-x) = -\Delta(x). \quad (4.9)$$

3. If  $x$  is a space-like vector, i.e.,  $\mathbf{x}^2 > x_0^2$ , then one has

$$\Delta(x) = 0. \quad (4.10)$$

**Proof** If  $x$  is a space-like vector, then there exists a coordinate system with  $x'_0 = 0$ . However, the four-vector  $x' = (\mathbf{x}', 0)$  can be transformed into  $(-\mathbf{x}', 0)$  by a spatial rotation. This rotation is, of course, a Lorentz transformation. Thus, from properties (1) and (2),

$$\Delta(x) = \Delta(x') = \Delta(-x') = -\Delta(x') = 0.$$

This tells us that  $\varphi(x)$  and  $\varphi(y)$  commute if  $x$  and  $y$  are spatially separated. According to Einstein, no action can ever be transmitted faster than the speed of light, whence observations at two spatially separated points can never interfere with

each other. According to quantum mechanics, two physical quantities that can be observed without interfering with each other or getting in each other's way are said to be commutative. This consequence is thus a quantum mechanical expression of Einstein causality. It is called *local commutativity* or *microscopic causality*. In addition, the Lorentz invariance of  $\Delta(x)$  means that the canonical quantization with a fixed choice of time axis is in fact independent of this choice of time axis. Thus, the commutation relation

$$[\varphi(x), \varphi(y)] = i\Delta(x - y) \quad (4.11)$$

does not in effect depend on the way we choose the time axis.

4.  $\Delta(x)$  satisfies the Klein–Gordon equation:

$$K_x \Delta(x) = (\square_x - m^2)\Delta(x) = 0. \quad (4.12)$$

**Proof** We have

$$(\square_x - m^2)[\varphi(x), \varphi(0)] = [(\square_x - m^2)\varphi(x), \varphi(0)] = 0.$$

5. For  $x_0 = 0$ ,

$$\Delta(x) = 0, \quad \frac{\partial}{\partial x_0} \Delta(x) = -\delta^3(x). \quad (4.13)$$

**Proof** The first equation is a special case of property (3). The second can be obtained from the canonical commutation relation:

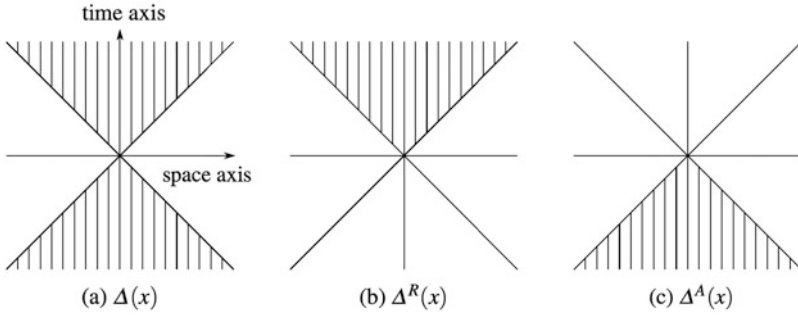
$$[\dot{\varphi}(x), \varphi(0)] = -\delta^3(x), \quad \text{where } x_0 = 0.$$

It can be also derived from the Fourier representation of  $\Delta(x)$ .

These are the properties of the Jordan–Pauli invariant  $\Delta$ -function. When the mass  $m$  vanishes, we write  $\Delta(x)$  as  $D(x)$ .

## 4.2 Various Invariant Functions

In the previous section, we defined the invariant  $\Delta$ -function using the commutation relation. Next, we introduce several kinds of Lorentz-invariant function that are closely associated with it.



**Fig. 4.1** Regions where the  $\Delta$ -function is nonzero

First, since  $\Delta(x)$  vanishes for space-like  $x$ , it is nonzero only in the future and past light cones. We define the retarded  $\Delta$ -function  $\Delta^R$  and the advanced  $\Delta$ -function  $\Delta^A$  as follows:

$$\Delta^R(x) = -\theta(x_0)\Delta(x) = \begin{cases} -\Delta(x) & \text{when } x_0 > 0, \\ 0 & \text{when } x_0 < 0, \end{cases} \quad (4.14)$$

$$\Delta^A(x) = \theta(-x_0)\Delta(x) = \begin{cases} 0 & \text{when } x_0 > 0, \\ \Delta(x) & \text{when } x_0 < 0. \end{cases} \quad (4.15)$$

The regions where these functions are not zero are the shaded areas in Fig. 4.1. Clearly, we have

$$\Delta(x) = \Delta^A(x) - \Delta^R(x), \quad \Delta^A(x) = \Delta^R(-x). \quad (4.16)$$

Next, we define the even function  $\bar{\Delta}$  by summing  $\Delta^A$  and  $\Delta^R(x)$ :

$$\bar{\Delta}(x) = \frac{1}{2}[\Delta^A(x) + \Delta^R(x)] = -\frac{1}{2}\epsilon(x_0)\Delta(x). \quad (4.17)$$

The above functions have been derived by decomposing space-time into the appropriate regions. Furthermore, we can also decompose the function  $\Delta$  in the momentum space in terms of positive and negative frequency modes:

$$\Delta(x) = \Delta^{(+)}(x) + \Delta^{(-)}(x), \quad (4.18)$$

$$i\Delta^{(+)}(x) = \frac{1}{(2\pi)^3} \int \frac{d^3p}{2p_0} e^{ip \cdot x} = \frac{1}{(2\pi)^3} \int d^4p \theta(p_0) \delta(p^2 + m^2) e^{ip \cdot x}, \quad (4.19)$$

$$\begin{aligned}
-i\Delta^{(-)}(x) &= \frac{1}{(2\pi)^3} \int \frac{d^3p}{2p_0} e^{-ip \cdot x} \\
&= \frac{1}{(2\pi)^3} \int d^4p \theta(-p_0) \delta(p^2 + m^2) e^{ip \cdot x} .
\end{aligned} \tag{4.20}$$

Therefore,

$$\Delta^{(+)}(-x) = -\Delta^{(-)}(x) , \quad \Delta^{(-)}(-x) = -\Delta^{(+)}(x) . \tag{4.21}$$

Next, we introduce the even function  $\Delta^{(1)}$ :

$$\Delta^{(1)}(x) = i[\Delta^{(+)}(x) - \Delta^{(-)}(x)] = \frac{1}{(2\pi)^3} \int d^4p \delta(p^2 + m^2) e^{ip \cdot x} . \tag{4.22}$$

From these, we obtain the following equations:

$$\Delta^{(+)}(x) = \frac{1}{2}[\Delta(x) - i\Delta^{(1)}(x)] , \quad \Delta^{(-)}(x) = \frac{1}{2}[\Delta(x) + i\Delta^{(1)}(x)] . \tag{4.23}$$

We observe that  $\Delta^{(+)}$  and  $\Delta^{(-)}$  satisfy the Klein–Gordon equation since  $\Delta(x)$  and  $\Delta^{(1)}(x)$  do. Furthermore,

$$K_x \Delta^R(x) = K_x \Delta^A(x) = K_x \bar{\Delta}(x) = -\delta^4(x) . \tag{4.24}$$

To prove these equations, we can use

$$\Delta^R(x) = \bar{\Delta}(x) - \frac{1}{2}\Delta(x) , \quad \Delta^A(x) = \bar{\Delta}(x) + \frac{1}{2}\Delta(x) , \tag{4.25}$$

and then we only need to prove the last equation in (4.24). For this purpose, we derive the Fourier expression of  $\bar{\Delta}(x)$ . Denoting Cauchy's principal value by  $\mathbf{P}$ , we can write

$$\epsilon(x_0) = \frac{x_0}{|x_0|} = \frac{2}{\pi} \int_0^\infty \frac{d\tau}{\tau} \sin(\tau x_0) = \frac{1}{i\pi} \mathbf{P} \int_{-\infty}^\infty \frac{d\tau}{\tau} e^{i\tau x_0} . \tag{4.26}$$

Therefore,

$$\begin{aligned}
\bar{\Delta}(x) &= \frac{i}{2\pi} \mathbf{P} \int \frac{d\tau}{\tau} e^{i\tau x_0} \frac{(-i)}{(2\pi)^3} \int d^4k \epsilon(k_0) \delta(k^2 + m^2) e^{ik \cdot x} \\
&= \frac{1}{(2\pi)^4} \int d^4p e^{ip \cdot x} \mathbf{P} \int \frac{d\tau}{\tau} \delta[\mathbf{p}^2 - (p_0 + \tau)^2 + m^2] \frac{p_0 + \tau}{|p_0 + \tau|} \\
&= \frac{1}{(2\pi)^4} \mathbf{P} \int \frac{d^4p}{p^2 + m^2} e^{ip \cdot x} ,
\end{aligned} \tag{4.27}$$

where we have taken  $\mathbf{p} = \mathbf{k}$  and  $p_0 = k_0 - \tau$ . Thus,

$$K_x \bar{\Delta}(x) = -\frac{1}{(2\pi)^4} \int d^4 p e^{ip \cdot x} = -\delta^4(x), \quad (4.28)$$

which proves (4.24).

Finally, we introduce Feynman's function  $\Delta_F$ , also known as Stueckelberg's causal function. This is defined by

$$\begin{aligned} \Delta_F &= \frac{1}{2} \Delta^{(1)}(x) - i \bar{\Delta}(x), \\ &= i \theta(x_0) \Delta^{(+)}(x) + i \theta(-x_0) \delta^{(+)}(-x), \\ &= \frac{-i}{(2\pi)^4} \int d^4 p e^{ip \cdot x} \left[ \frac{P}{p^2 + m^2} + i\pi \delta(p^2 + m^2) \right], \\ &= \frac{-i}{(2\pi)^4} \lim_{\epsilon \rightarrow +0} \int d^4 p e^{ip \cdot x} \frac{1}{p^2 + m^2 - i\epsilon}. \end{aligned} \quad (4.29)$$

From now on, we will omit the limit symbol, taking it as understood. Clearly, this function satisfies

$$K_x \Delta_F(x) = i \delta^4(x). \quad (4.30)$$

### 4.3 Unequal-Time Commutation Relations of Free Fields

Up to now, we have investigated the equal-time commutation relations associated with the canonical quantization. In fact, using the invariant function we can obtain the unequal-time commutation relation directly.

For the real scalar field, as mentioned in Sect. 4.1, we have

$$[\varphi(x), \varphi(y)] = i \Delta(x - y). \quad (4.31)$$

Similarly, for the complex scalar field,

$$[\varphi(x), \varphi^\dagger(y)] = [\varphi^\dagger(x), \varphi(y)] = i \Delta(x - y), \quad (4.32)$$

$$[\varphi(x), \varphi(y)] = [\varphi^\dagger(x), \varphi^\dagger(y)] = 0. \quad (4.33)$$

Next, we derive the unequal-time anti-commutation relation for spinor fields. We solve this as an initial value problem of a differential equation, i.e., we start with the equation

$$\left( \gamma_\mu \frac{\partial}{\partial x_\mu} + m \right) \{ \psi(x), \bar{\psi}(y) \} = 0. \quad (4.34)$$

As an initial condition, we use the equal-time anti-commutation relation

$$\{\psi_r(x), \bar{\psi}_s(y)\} = (\gamma_4)_{rs} \delta^3(x - y), \quad \text{when } x_0 = y_0. \quad (4.35)$$

The solution of (4.34) is given uniquely by

$$\{\psi_r(x), \bar{\psi}_s(y)\} = -i \left( \gamma_\mu \frac{\partial}{\partial x_\mu} - m \right)_{rs} \Delta(x - y) = i S_{rs}(x - y). \quad (4.36)$$

Setting  $x_0 = y_0$  in this equation, the right-hand side becomes

$$-i \gamma_4 \frac{\partial}{\partial x_4} \Delta(x - y) = \gamma_4 \delta^3(x - y),$$

which gives (4.35). Thus, in this case the anti-commutation relations are

$$\begin{aligned} \{\psi_r(x), \bar{\psi}_s(y)\} &= i S_{rs}(x - y), \\ \{\psi_r(x), \psi_s(y)\} &= \{\bar{\psi}_r(x), \bar{\psi}_s(y)\} = 0. \end{aligned} \quad (4.37)$$

Next, we consider the complex vector field. In this case, the commutator satisfies the equations

$$K_x [\varphi_\mu(x), \varphi_\nu^\dagger(y)] = 0, \quad (4.38)$$

$$\frac{\partial}{\partial x_\mu} [\varphi_\mu(x), \varphi_\nu^\dagger(y)] = 0. \quad (4.39)$$

To begin with, we consider only the space components. The time component will be analyzed using the second equation. The initial condition is given by the equal-time commutation relations:

$$[\varphi_j(x), \varphi_k^\dagger(y)] = 0, \quad [\dot{\varphi}_j(x), \varphi_k^\dagger(y)] = -i \left( \delta_{jk} - \frac{1}{m^2} \frac{\partial^2}{\partial x_j \partial x_k} \right) \delta^3(x - y), \quad (4.40)$$

where  $x_0 = y_0$ . The solution which satisfies (4.38), (4.39), and (4.40) is

$$\begin{aligned} [\varphi_\mu(x), \varphi_\nu^\dagger(y)] &= i \left( \delta_{\mu\nu} - i \frac{1}{m^2} \frac{\partial^2}{\partial x_\mu \partial x_\nu} \right) \Delta(x - y), \\ [\varphi_\mu(x), \varphi_\nu(y)] &= [\varphi_\mu^\dagger(x), \varphi_\nu^\dagger(y)] = 0. \end{aligned} \quad (4.41)$$

This implies that, although canonical quantization gives an initial condition to the above differential equation by fixing a time axis, the commutators obtained in the

end are independent of the way we choose a specific inertial system. Thus, the method of canonical quantization yields a Lorentz-invariant quantization, at least for free fields.

In Sect. 2.3, we showed that the canonical equation and the Euler–Lagrange equation coincide as long as the canonical commutation relations hold. Conversely, if we assume that the canonical equation coincides with the Euler–Lagrange equation, does this restrict the commutation relations? Furthermore, can we derive the commutation relations without specifying a time axis as we did in the method of canonical quantization? Using these ideas, we shall now quantize free fields.

## 4.4 Generalities of the Quantization of Free Fields

As we mentioned in the previous section, from the requirement that the canonical equation and Euler’s equation must coincide, we can derive the commutation relations for free fields. In this case, an important assumption is that all free fields satisfy the Klein–Gordon equation as a consequence of the field equation.

Assume that the Lagrangian density has the form

$$\mathcal{L} = \mathcal{L} \left[ \varphi_\alpha, \varphi_\alpha^\dagger, \varphi_{\alpha,\mu}, \varphi_{\alpha,\mu}^\dagger \right], \quad (4.42)$$

where  $\varphi_\alpha$  is a complex field. We will deal with the changes needed to handle a real field whenever necessary. In the case of free fields,  $\mathcal{L}$  is a bilinear form in  $\varphi$ ,  $\varphi^\dagger$ , and their derivatives. The Euler–Lagrange equation can therefore be written in the form

$$[\mathcal{L}]_{\varphi_\alpha^\dagger} = D_{\alpha\beta}(\partial)\varphi_\beta = 0, \quad (4.43)$$

where  $D$  is a matrix which includes differential operators. In the case of the complex scalar field, we have

$$[\mathcal{L}]_{\varphi^\dagger} = K\varphi = (\square - m^2)\varphi = 0, \quad (4.44)$$

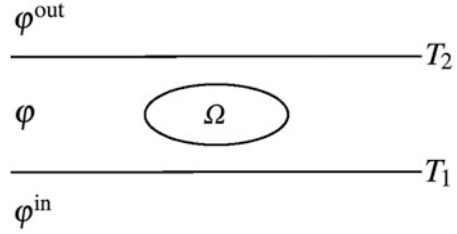
so in this case the one-dimensional matrix  $D$  is given by

$$D(\partial) = K = \square - m^2. \quad (4.45)$$

As a perturbation, we add to the Lagrangian density of the above free field an interaction term between  $\varphi_\alpha(x)$  and an external complex field  $Q(x)$ , and assume equivalence of the canonical equation and the Euler–Lagrange equation:

$$\mathcal{L}(x) \rightarrow \mathcal{L}(x) - Q(x)\varphi_\alpha^\dagger(x)\varphi_\alpha(x), \quad (4.46)$$

**Fig. 4.2** Relation between  $T_1$ ,  $T_2$ , and  $\Omega$



where we sum over  $\alpha$ . From the variational principle,

$$D_{\alpha\beta}(\partial)\varphi_\beta - Q\varphi_\alpha = 0. \quad (4.47)$$

If  $\varphi$  is a real scalar field, then  $Q$  can be replaced by  $2Q$  in the above equation. We will return to this difference later.

Next, we assume that  $Q$  is nonzero only in a certain finite space-time region  $\Omega$ , while

$$Q = 0 \text{ when } t > T_2 \text{ or } t < T_1. \quad (4.48)$$

We denote  $\varphi_\alpha$  before  $T_1$  by  $\varphi_\alpha^{\text{in}}$ , and  $\varphi_\alpha$  after  $T_2$  by  $\varphi_\alpha^{\text{out}}$ . This anteroposterior relation is illustrated in Fig. 4.2. Since the external field vanishes before  $T_1$  and after  $T_2$ , we obtain

$$D_{\alpha\beta}(\partial)\varphi_\beta^{\text{in}} = 0 \text{ for } t < T_1, \quad (4.49)$$

$$D_{\alpha\beta}(\partial)\varphi_\beta^{\text{out}} = 0 \text{ for } t > T_2. \quad (4.50)$$

If  $Q$  is zero in the whole of space-time, then of course we obtain

$$\varphi_\alpha = \varphi_\alpha^{\text{in}} = \varphi_\alpha^{\text{out}}.$$

We next try to solve for  $\varphi_\alpha$  in terms of  $\varphi_\alpha^{\text{in}}$ , and set

$$\varphi_\alpha = \varphi_\alpha^{\text{in}} + \varphi'_\alpha. \quad (4.51)$$

It is clear that  $\varphi'_\alpha = 0$  for  $t < T_1$ . Applying  $D_{\alpha\beta}$  to the above equation, we find

$$D_{\alpha\beta}(\partial)\varphi_\beta = D_{\alpha\beta}(\partial)\varphi_\beta^{\text{in}} + D_{\alpha\beta}(\partial)\varphi'_\beta = Q\varphi_\alpha.$$

Therefore, from (4.49),

$$D_{\alpha\beta}(\partial)\varphi'_\beta = Q\varphi_\alpha. \quad (4.52)$$



Since we have the boundary condition that  $\varphi'$  is 0 in the region  $t < T_1$ , we can integrate the above equation using the retarded  $\Delta$ -function, i.e., defining  $\Delta^R$  by

$$D_{\alpha\beta}(\partial)\Delta_{\beta\gamma}^R(x) = -\delta_{\alpha\gamma}\delta^4(x), \quad \Delta_{\alpha\beta}^R = 0 \text{ when } x_0 < 0, \quad (4.53)$$

the quantity  $\varphi'_\alpha$  can be written in the form

$$\varphi'_\alpha(x) = - \int_{\Omega} d^4x' \Delta_{\alpha\beta}^R(x-x') Q(x') \varphi_\beta(x'). \quad (4.54)$$

Thus, we find that  $\varphi'_\alpha$  vanishes in the region  $x_0 < T_1$  thanks to the properties of  $\Delta^R$  and  $\Omega$ .

Rewriting (4.51) for  $\varphi$ , we obtain the integral equation

$$\varphi_\alpha(x) = \varphi_\alpha^{\text{in}} - \int_{\Omega} d^4x' \Delta_{\alpha\beta}^R(x-x') Q(x') \varphi_\beta(x'). \quad (4.55)$$

In addition, using  $\varphi_\alpha^{\text{out}}$  and  $\Delta^A$  defined by changing the boundary condition in (4.53), we have

$$\varphi_\alpha(x) = \varphi_\alpha^{\text{out}}(x) - \int_{\Omega} d^4x' \Delta_{\alpha\beta}^A(x-x') Q(x') \varphi_\beta(x'). \quad (4.56)$$

The above are consequences of the Euler–Lagrange equation.

We now discuss the canonical equation. To begin with, the Hamiltonian density is

$$\mathcal{H} = \sum_{\alpha} \left( \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_{\alpha}} \dot{\varphi}_{\alpha} + \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_{\alpha}^{\dagger}} \dot{\varphi}_{\alpha}^{\dagger} \right) - \mathcal{L}. \quad (4.57)$$

Introducing  $Q$ , the modification corresponding to (4.46) leads to

$$\mathcal{H}(x) \rightarrow \mathcal{H}(x) + Q(x) \varphi_{\alpha}^{\dagger}(x) \varphi_{\alpha}(x) \equiv \mathcal{H}(x) + \mathcal{H}'(x). \quad (4.58)$$

The Hamiltonian of this system is

$$H = \int_{x_0=t} d^3x [\mathcal{H}(x) + \mathcal{H}'(x)] \equiv \mathcal{H}(x) + \mathcal{H}'(x). \quad (4.59)$$

When  $Q(x)$  is nonzero,  $H$  depends explicitly on  $t$ . When a physical quantity  $F$  depends on  $t$  both implicitly via canonical variables and explicitly through some external field, its total derivative is given by

$$\frac{dF}{dt} = \frac{\partial F}{\partial(t)} + i[H(t), F], \quad (4.60)$$

where the first term on the right-hand side is the derivative with respect to the explicit  $t$ -dependence through the external field.

In particular, for a field, we have

$$\frac{\partial \varphi_\alpha(x)}{\partial x_0} = i[H(x_0), \varphi_\alpha(x)]. \quad (4.61)$$

Since  $Q = 0$  for  $x_0 < T$ , we obtain

$$\frac{\partial \varphi_\alpha^{\text{in}}(x)}{\partial x_0} = i[H_0[\varphi_\alpha^{\text{in}}(x), \dots], \varphi_\alpha^{\text{in}}(x)], \quad (4.62)$$

where  $H_0[\dots]$  stands for a functional of  $\varphi_\alpha^{\text{in}}(x)$ . When  $Q$  is nonzero, we take  $\varphi_\alpha$  as follows (by analogy with the super multi-time theory to be discussed later):

$$\varphi_\alpha(x) = U(x_0)^{-1} \varphi_\alpha^{\text{in}}(x) U(x_0). \quad (4.63)$$

The reason why we adopt this form is that  $\varphi_\alpha$  and  $\varphi_\alpha^{\text{in}}$  satisfy the same equal-time commutation relation. Since  $\varphi_\alpha$  is equal to  $\varphi_\alpha^{\text{in}}$  for  $x_0 < T_1$ ,

$$U = 1 \quad \text{when } x_0 < T_1. \quad (4.64)$$

Substituting (4.63) into (4.61), the right-hand side becomes

$$\begin{aligned} & iU(x_0)^{-1} [H_0[\varphi_\alpha^{\text{in}}, \dots], \varphi_\alpha^{\text{in}}(x)] U(x_0) + iU(x_0)^{-1} [H'[\varphi_\alpha^{\text{in}}, \dots], \varphi_\alpha^{\text{in}}(x)] U(x_0) \\ & = U(x_0)^{-1} \frac{\partial \varphi_\alpha^{\text{in}}(x)}{\partial x_0} U(x_0) + iU(x_0)^{-1} [H'[\varphi_\alpha^{\text{in}}(x), \dots], \varphi_\alpha^{\text{in}}(x)] U(x_0). \end{aligned} \quad (4.65)$$

On the other hand, the left-hand side becomes

$$\begin{aligned} \frac{\partial}{\partial x_0} [U(x_0)^{-1} \varphi_\alpha^{\text{in}}(x) U(x_0)] & = U(x_0)^{-1} \frac{\partial \varphi_\alpha^{\text{in}}(x)}{\partial x_0} U(x_0) + U(x_0)^{-1} \varphi_\alpha^{\text{in}}(x) \frac{\partial U(x_0)}{\partial x_0} \\ & = -U(x_0)^{-1} \frac{\partial U(x_0)}{\partial x_0} U(x_0)^{-1} \varphi_\alpha^{\text{in}}(x) U(x_0). \end{aligned} \quad (4.66)$$

In deriving this equation, we have used the following formula:

$$\frac{\partial}{\partial x_0} U(x_0)^{-1} = -U(x_0)^{-1} \frac{\partial U(x_0)}{\partial x_0} U(x_0)^{-1}.$$

Equating (4.65) and (4.66), we obtain

$$\left[ iH'[\varphi_\alpha^{\text{in}}(x), \dots] + \frac{\partial U(x_0)}{\partial x_0} U(x_0)^{-1}, \varphi_\alpha^{\text{in}}(x) \right] = 0. \quad (4.67)$$

This equation can be satisfied by setting the first factor in the commutator equal to zero, viz.,

$$i \frac{\partial U(x_0)}{\partial x_0} = H'[\varphi_\alpha^{\text{in}}(x), \dots] U(x_0) , \quad (4.68)$$

where

$$H'[\varphi_\alpha^{\text{in}}(x), \dots] = \int d^3x Q(x) \varphi_\alpha^{\dagger \text{in}}(x) \varphi_\alpha^{\text{in}}(x) . \quad (4.69)$$

As a consequence, we have shown that, if  $U(x_0)$  satisfies (4.68),  $\varphi_\alpha(x)$  can be written in the form (4.63). The boundary condition which  $U(x_0)$  must satisfy is (4.64). Then, taking into account (4.64) and integrating (4.68), we obtain the integral equation

$$U(x_0) = 1 - i \int_{-\infty}^{x_0} d^4x' Q(x') \varphi_\beta^{\dagger \text{in}}(x') \varphi_\beta^{\text{in}}(x') U(x'_0) . \quad (4.70)$$

We now compare the two integral equations (4.55) and (4.70). Assuming that  $Q$  is small, we consider the power series expansion with respect to  $Q$ , keeping terms up to the second order in  $Q$ . The solutions of the two integral equations are

$$\varphi_\alpha(x) = \varphi_\alpha^{\text{in}}(x) - \int_{\Omega} d^4x' \Delta_{\alpha\beta}^{\text{R}}(x - x') Q(x') \varphi_\beta^{\text{in}}(x') + \text{second order terms in } Q , \quad (4.71)$$

$$U(x_0) = 1 - i \int_{-\infty}^{x_0} d^4x' Q(x') \varphi_\beta^{\dagger \text{in}}(x') \varphi_\beta^{\text{in}}(x') + \text{second order terms in } Q . \quad (4.72)$$

In the same approximation,

$$\begin{aligned} \varphi_\alpha(x) &= U(x_0)^{-1} \varphi_\alpha^{\text{in}}(x) U(x_0) \\ &= \varphi_\alpha^{\text{in}}(x) - i \int_{-\infty}^{\infty} d^4x' \theta(x_0 - x'_0) [\varphi_\alpha^{\text{in}}(x), \varphi_\beta^{\dagger \text{in}}(x') \varphi_\beta^{\text{in}}(x')] Q(x') + \dots . \end{aligned} \quad (4.73)$$

Requiring that (4.71) coincide with (4.73),

$$\theta(x_0 - x'_0) [\varphi_\alpha^{\text{in}}(x), \varphi_\beta^{\dagger \text{in}}(x') \varphi_\beta^{\text{in}}(x')] = -i \Delta_{\alpha\beta}^{\text{R}}(x - x') \varphi_\beta^{\text{in}}(x') . \quad (4.74)$$

Taking  $Q \rightarrow 0$ ,  $\varphi_\alpha^{\text{in}}$  becomes  $\varphi_\alpha$ , and we obtain for the free field

$$\theta(x_0 - x'_0) [\varphi_\alpha(x), \varphi_\beta^{\dagger}(x') \varphi_\beta(x')] = -i \Delta_{\alpha\beta}^{\text{R}}(x - x') \varphi_\beta(x') . \quad (4.75)$$

In a similar manner, we obtain

$$\theta(x'_0 - x_0)[\varphi_\alpha(x), \varphi_\beta^\dagger(x')\varphi_\beta(x')] = i\Delta_{\alpha\beta}^A(x - x')\varphi_\beta(x'). \quad (4.76)$$

We define the  $\Delta$ -function:

$$\Delta_{\alpha\beta}(x) = \Delta_{\alpha\beta}^A(x) - \Delta_{\alpha\beta}^R(x). \quad (4.77)$$

Adding (4.75) and (4.76), we find the following commutation relation:

$$[\varphi_\alpha(x), \varphi_\beta^\dagger(x')\varphi_\beta(x')] = i\Delta_{\alpha\beta}(x - x')\varphi_\beta(x'). \quad (4.78)$$

In the above, we considered a complex field. For real fields, a factor of 2 appears, as mentioned earlier:

$$[\varphi_\alpha(x), \varphi_\beta(x')\varphi_\beta(x')] = 2i\Delta_{\alpha\beta}(x - x')\varphi_\beta(x'). \quad (4.79)$$

From the two equations above, we can derive the commutators among two fields. In the case of the complex field, the results are as follows:

1. Solution for Bose statistics:

$$\begin{aligned} [\varphi_\alpha(x), \varphi_\beta(x')] &= [\varphi_\alpha^\dagger(x), \varphi_\beta^\dagger(x')] = 0, \\ [\varphi_\alpha(x), \varphi_\beta^\dagger(x')] &= i\Delta_{\alpha\beta}(x - x'). \end{aligned} \quad (4.80)$$

2. Solution for Fermi statistics:

$$\begin{aligned} \{\varphi_\alpha(x), \varphi_\beta(x')\} &= \{\varphi_\alpha^\dagger(x), \varphi_\beta^\dagger(x')\} = 0, \\ \{\varphi_\alpha(x), \varphi_\beta^\dagger(x')\} &= i\Delta_{\alpha\beta}(x - x'). \end{aligned} \quad (4.81)$$

Assuming Bose statistics for the real field,

$$[\varphi_\alpha(x), \varphi_\beta(x')] = i\Delta_{\alpha\beta}(x - x'). \quad (4.82)$$

Regarding the applicability of Bose or Fermi statistics, we have the *spin–statistics theorem* due to Pauli [81, 82], which establishes a connection between spin and statistics. If we assume the Lorentz invariance of the theory, microscopic causality, and a positive-definite metric, particles with integer spin must obey Bose statistics, while particles with half-integer spin must obey Fermi statistics. We will discuss the notion of a positive-definite metric in the next chapter.

At the beginning of this section, we assumed that the fields satisfy the Klein–Gordon equation as a consequence of the field equation (4.43). This means that mathematically there exists a matrix of differential operators  $C$  satisfying

$$D_{\alpha\beta}(\partial)C_{\beta\gamma}(\partial) = C_{\alpha\beta}(\partial)D_{\beta\gamma}(\partial) = \delta_{\alpha\gamma}(\square - m^2). \quad (4.83)$$

Therefore, (4.53) and the corresponding solution for the equation of  $\Delta^A$  are given by

$$\Delta_{\alpha\beta}^R(x) = C_{\alpha\beta}(\partial)\Delta^R(x), \quad \Delta_{\alpha\beta}^A(x) = C_{\alpha\beta}(\partial)\Delta^A(x). \quad (4.84)$$

Thus,

$$\Delta_{\alpha\beta}(x) = C_{\alpha\beta}(\partial)\Delta(x). \quad (4.85)$$

Using this formula, free fields can be quantized directly.

**One-Component Real Scalar Field** In this case, it is clear that

$$C(\partial) = 1, \quad [\varphi(x), \varphi(x')] = i\Delta(x - x'). \quad (4.86)$$

**Faddeev–Popov Ghost Field** Suppose two real scalar fields  $c$  and  $\bar{c}$  obey Fermi statistics and are described by the Lagrangian density

$$\mathcal{L} = i\partial_\mu \bar{c} \cdot \partial_\mu c. \quad (4.87)$$

In this case,

$$D(\partial) = -i\square, \quad C(\partial) = i. \quad (4.88)$$

The commutation relations are therefore

$$\{c(x), c(x')\} = \{\bar{c}(x), \bar{c}(x')\} = 0, \quad \{c(x), \bar{c}(x')\} = -D(x - x'). \quad (4.89)$$

This case violates Pauli’s theorem for spin and statistics, so at least one of the assumptions leading to that theorem must be relaxed. In fact, in this case we drop the assumption of a positive-definite metric and introduce an indefinite metric. This field is then referred to as a *ghost field*. As will be discussed later, it plays an important role in non-Abelian gauge field theory. In general, indefinite metrics arise in gauge field theories, as will be discussed in detail in the next chapter.

**Complex Vector Field** From the Lagrangian density (3.221),

$$[\mathcal{L}]_{\varphi_\alpha}^\dagger = \partial_\beta(\partial_\beta \varphi_\alpha - \partial_\alpha \varphi_\beta) - m^2 \varphi_\alpha = D_{\alpha\beta}(\partial)\varphi_\beta. \quad (4.90)$$

Therefore,

$$D_{\alpha\beta}(\partial) = \delta_{\alpha\beta}(\square - m^2) - \partial_\alpha \partial_\beta . \quad (4.91)$$

From this, we obtain  $C$  as follows:

$$C_{\alpha\beta}(\partial) = \delta_{\alpha\beta} - \frac{1}{m^2} \partial_\alpha \partial_\beta . \quad (4.92)$$

This is the well-known form, reproducing the equation (4.41).

Compared to the method of canonical quantization, it is clear from the example of the vector field that, using this approach, the unequal-time commutation relations can be obtained rather easily.

# Chapter 5

## Indefinite Metric and the Electromagnetic Field



One may say that the history of quantum field theory began with the quantization of the electromagnetic field. However, the electromagnetic field cannot be treated as a simple field because one feature of this vector field, which is in fact a *gauge field*, is an indefinite metric. We thus begin by discussing this metric. For more detail, the reader is referred to the book *Quantum Field Theory* by Noboru Nakanishi, written in Japanese [83].

### 5.1 Indefinite Metric

A key tool for formulating quantum mechanics is Hilbert space, which is a vector space. We start with a brief review of finite-dimensional vector spaces  $\mathcal{V}$ , which will be intuitive rather than mathematically rigorous:

1. Vectors form a group under addition.
2. Vectors can be multiplied by complex numbers and the distributive property holds.
3. The number of linearly independent vectors defines the dimension of the vector space.

In quantum physics, vectors are written in the form “bra”  $\langle x|$  or “ket”  $|x\rangle$ . We define an inner product which is used to introduce a metric.

The inner product has the following properties:

$$\langle y|x\rangle = \langle x|y\rangle^* , \tag{5.1}$$

$$\langle y|ax + bx'\rangle = a\langle y|x\rangle + b\langle y|x'\rangle , \tag{5.2}$$

$$\langle ay + by'|x\rangle = a^*\langle y|x\rangle + b^*\langle y'|x\rangle , \tag{5.3}$$

where  $a$  and  $b$  stand for complex numbers. The third equation follows from the first two. In mathematics, (5.2) and (5.3) are often the other way round, i.e.,  $a^*$  and  $b^*$  appear in (5.2), while  $a$  and  $b$  appear in (5.3). The inner product is said to be positive-definite if

$$\langle x|x \rangle \geq 0. \quad (5.4)$$

Then,  $\langle x|x \rangle = 0$  and  $|0\rangle = 0$  are equivalent. In this case, the space is Euclidean. An indefinite-metric space is one in which we relax the condition (5.4). Then, (5.1) implies that  $\langle x|x \rangle$  is real but not necessarily positive.

### Linear Operators

If  $T$  is such that, for an arbitrary  $|x\rangle \in \mathcal{V}$ ,  $T|x\rangle \in \mathcal{V}$ , then we call  $T$  an *operator on*  $\mathcal{V}$ . If  $T$  satisfies the linearity condition

$$T|ax + by\rangle = aT|x\rangle + bT|y\rangle, \quad (5.5)$$

it is called a *linear operator*.

Given a set of linearly independent vectors, i.e., a basis  $\{e_j\}$ , we can construct a representation of  $T$ . An arbitrary  $|x\rangle \in \mathcal{V}$  can be uniquely expressed in the form

$$|x\rangle = \sum_j x_j |e_j\rangle. \quad (5.6)$$

We can then write  $T|e_j\rangle$  in the form

$$T|e_j\rangle = \sum_k |e_k\rangle t_{kj}. \quad (5.7)$$

The matrix  $t$  with components  $t_{kj}$  is called a *representation of*  $T$ . Denoting the corresponding representation of  $S$  by  $s$ ,

$$\begin{aligned} ST|e_j\rangle &= \sum_k S|e_k\rangle t_{kj} \\ &= \sum_k \sum_l |e_l\rangle s_{lk} t_{kj} \\ &= \sum_l |e_l\rangle \left( \sum_k s_{lk} t_{kj} \right), \end{aligned} \quad (5.8)$$

whence the matrix  $st$  represents the operator  $ST$ .

A matrix  $u$  such that  $\det u \neq 0$  can be used to transform to another basis:

$$|e'_j\rangle = \sum_k |e_k\rangle u_{kj}. \quad (5.9)$$



Denoting the representation of  $T$  in the new basis by  $\tilde{t}$ , we have

$$T|e'_j\rangle = \sum_k |e'_k\rangle \tilde{t}_{kj} . \quad (5.10)$$

The transformation law for the representation  $t$  is then given by the similarity transformation

$$\tilde{t} = u^{-1} t u . \quad (5.11)$$

### Metric Matrix

Given a basis  $\{e_j\}$ , we define the metric matrix by

$$\eta_{mn} = \langle e_m | e_n \rangle . \quad (5.12)$$

From the property (5.1) of the inner product, we then have  $\eta_{nm} = \eta_{mn}^*$ , i.e.,  $\eta^\dagger = \eta$ , which shows that  $\eta$  is a Hermitian matrix. Under the change of basis (5.9), this matrix transforms by

$$\tilde{\eta} = u^\dagger \eta u . \quad (5.13)$$

Note the difference with (5.11).

### Degeneracy

If there exists a vector  $|k\rangle$  satisfying  $\langle l|k\rangle = 0$  for any vector  $|l\rangle$  in  $\mathcal{V}$ , the vector space is said to be degenerate.

### Hermitian Conjugate

For a given linear operator  $T$  and arbitrary vectors  $|k\rangle$  and  $|l\rangle$ , the relation

$$\langle k|T^\dagger|l\rangle = \langle l|T|k\rangle^* \quad (5.14)$$

defines the Hermitian conjugate  $T^\dagger$  of  $T$ . If the matrices representing  $T$  and  $T^\dagger$  are  $s$  and  $t$ , respectively, then

$$\eta s = t^\dagger \eta . \quad (5.15)$$

If in particular  $\mathcal{V}$  is non-degenerate and  $\det \eta \neq 0$ , then we obtain

$$s = \eta^{-1} t^\dagger \eta . \quad (5.16)$$

Thus, in general, the matrix representing a Hermitian operator, i.e., one with  $T^\dagger = T$ , will not be a Hermitian matrix. In an indefinite-metric space, we use the following terminology for operators:

$$\begin{aligned} U^\dagger &= U^{-1} && \text{unitary matrix,} \\ P^2 &= P && \text{projection operator,} \\ K^n &= 0 \quad (K^{n-1} \neq 0) && \text{nilpotent operator.} \end{aligned}$$

In a non-degenerate vector space, the eigenvalues of  $\eta$  can be normalized to  $\pm 1$  by a suitable transformation of the basis.

## 5.2 Generalized Eigenstates

When the metric is indefinite, things happen that would not be expected with a positive-definite metric. In the Lorentz-invariant formulation of gauge field theories, which include the electromagnetic field as a typical case, the indefinite metric is inevitable, so we discuss its properties here.

If we have

$$T|k\rangle = \lambda|k\rangle, \quad (5.17)$$

then, as in the case of the positive-definite metric,  $|k\rangle$  is called an eigenstate of  $T$  and  $\lambda$  its eigenvalue. In addition, if  $\mathscr{W}$  is a subspace of  $\mathscr{V}$  with the property

$$|k\rangle \in \mathscr{W} \implies T|k\rangle \in \mathscr{W}, \quad (5.18)$$

then  $\mathscr{W}$  is called an *invariant subspace* of  $T$ . For any matrix  $t$  representing  $T$ , we can define

$$f(x) = \det(x\mathbf{I} - t), \quad (5.19)$$

where  $\mathbf{I}$  is the unit matrix in the space of the representation  $t$ . This polynomial in  $x$  is called the *eigen polynomial* or *characteristic polynomial* of  $T$  and is independent of the choice of basis.

**Cayley–Hamilton Theorem** This states that  $T$  satisfies its own characteristic polynomial, i.e.,

$$f(T) = 0. \quad (5.20)$$

**Proof** We define the adjugate matrix  $\tilde{s}$  of the matrix  $s$  by

$$\tilde{s}_{mn} = \frac{\partial}{\partial s_{nm}} \det s . \quad (5.21)$$

Clearly,

$$\sum_m s_{nm} \tilde{s}_{ml} = \delta_{nl} \det s . \quad (5.22)$$

We define the primary expression  $s_{nm}(x)$  by

$$s_{nm}(x) = x \delta_{nm} - t_{nm} . \quad (5.23)$$

Therefore,

$$\sum_m s_{nm}(T) \tilde{s}_{ml}(T) = \delta_{nl} f(T) , \quad (5.24)$$

$$T|m\rangle = \sum_n |n\rangle t_{nm} . \quad (5.25)$$

Rewriting (5.25),

$$\sum_n (\delta_{nm} T - t_{nm}) |n\rangle = \sum_n s_{nm}(T) |n\rangle = 0 .$$

Thus,

$$\begin{aligned} \sum_{n,m} \tilde{s}_{ml}(T) s_{nm}(T) |n\rangle &= \sum_{n,m} s_{nm}(T) \tilde{s}_{ml}(T) |n\rangle \\ &= \sum_n \delta_{ln} f(T) |n\rangle \\ &= f(T) |l\rangle \\ &= 0 . \end{aligned} \quad (5.26)$$

Here  $|l\rangle$  is an arbitrary basis vector. Since this equation holds for every element of the basis, this proves (5.20).  $\blacksquare$

Next, we factorize  $f(x)$ . Considering  $\lambda_i \neq \lambda_j$  for  $i \neq j$ , we obtain

$$f(x) = \prod_{j=1}^r (x - \lambda_j)^{N_j} , \quad \sum_{j=1}^r N_j = N , \quad (5.27)$$

where  $N$  is the dimension of the vector space. We know that  $f(T) = 0$ , and consider the polynomial with the lowest degree satisfying  $g(T) = 0$ :

$$g(x) = \prod_{j=1}^r (x - \lambda_j)^{n_j}, \quad 0 < n_j \leq N_j. \quad (5.28)$$

**Decomposition Theorem** The vector space  $\mathcal{V}$  can be decomposed as

$$\mathcal{V} = \sum_{j=1}^r \bigoplus \mathcal{V}(\lambda_j), \quad (5.29)$$

where  $|k\rangle \in \mathcal{V}(\lambda_j)$  means that  $(T - \lambda_j)^{n_j}|k\rangle = 0$ .

**Proof** First, we define  $g_1(x), g_2(x), \dots, g_r(x)$  by

$$g_j(x) = \frac{g(x)}{(x - \lambda_j)^{n_j}} = \prod_{i \neq j} (x - \lambda_i)^{n_i}. \quad (5.30)$$

Since  $g_1(x), g_2(x), \dots, g_r(x)$  are disjoint, there exist polynomials  $h_j(x)$  satisfying

$$\sum_{j=1}^r h_j(x)g_j(x) = 1. \quad (5.31)$$

We now set

$$P_j(T) = h_j(T)g_j(T), \quad \sum_{j=1}^r P_j(T) = 1. \quad (5.32)$$

Additionally, if  $i \neq j$ ,  $g_i(x)g_j(x)$  can be divided by  $g(x)$ , so

$$g_i(T)g_j(T) = 0.$$

Hence, if  $i \neq j$ ,  $P_i(T)P_j(T) = 0$ . Multiplying (5.32) by  $P_j(T)$ ,

$$P_j^2(T) = P_j(T). \quad (5.33)$$

This shows that  $P_j(T)$  is a projection operator, and  $\mathcal{V}(\lambda_j)$  can be defined by  $P_j(T)\mathcal{V} = \mathcal{V}(\lambda_j)$ , which proves the decomposition theorem. ■

The orthogonality condition for eigenvectors is complicated in the case of an indefinite metric. Even if  $H$  is a Hermitian operator, it is not always true that its eigenvalues are real. If  $h$  is the matrix representation of  $H$ , (5.15) implies

$$\eta h = h^\dagger \eta = (\eta h)^\dagger . \quad (5.34)$$

Thus, the product  $\eta h$  is a Hermitian matrix. Here we can choose  $\eta$  to satisfy  $\eta^2 = \mathbf{1}$  and (5.13) allows us to select a proper basis when  $\mathcal{V}$  is not degenerate. As a consequence,

$$\det(x\mathbf{I} - h) = \det \eta \cdot \det(x\eta - \eta h) , \quad (5.35)$$

which shows that the characteristic polynomial of  $H$  is then a real polynomial. Note again that the characteristic polynomial is independent of the choice of basis. The eigenvalues of  $H$  are thus either real numbers or complex conjugate pairs of complex numbers. In the following, we take  $H$  to be a Hermitian operator and consider the subspace  $\mathcal{V}(\lambda)$  defined above.

First, considering the two vectors which belong to  $\mathcal{V}(\lambda_j)$  and  $\mathcal{V}(\lambda_k)$ , i.e., the ket vectors  $|e_j\rangle$  and  $|e_k\rangle$ , we have

$$(H - \lambda_j)^{n_j} |e_j\rangle = 0 , \quad \langle e_k | (H - \lambda_k^*)^{n_k} = 0 . \quad (5.36)$$

If  $\lambda_j \neq \lambda_k^*$ , then  $(x - \lambda_j)^{n_j}$  and  $(x - \lambda_k^*)^{n_k}$  are disjoint, so there exist polynomials  $f(x)$  and  $g(x)$  satisfying

$$f(x)(x - \lambda_j)^{n_j} + g(x)(x - \lambda_k^*)^{n_k} = 1 . \quad (5.37)$$

We thus obtain the orthogonality relation

$$\langle e_k | e_j \rangle = \langle e_k | f(H)(H - \lambda_j)^{n_j} |e_j\rangle + \langle e_k | (H - \lambda_k^*)^{n_k} g(H) |e_j\rangle = 0 . \quad (5.38)$$

For complex eigenvalues  $\lambda_j \neq \lambda_j^*$ , we then have

$$\langle e_j | e_j \rangle = 0 , \quad (5.39)$$

so the norm of this vector is zero.

If  $\lambda_j$  is a real number, there must exist a vector  $|x_j\rangle$  in  $\mathcal{V}(\lambda_j)$  such that

$$\langle x_j | (H - \lambda_j)^{n_j-1} |x_j\rangle \neq 0 . \quad (5.40)$$

Otherwise, we would always have  $(H - \lambda_j)^{n_j-1} = 0$  on  $\mathcal{V}(\lambda_j)$ , and this would contradict the condition that  $n_j$  is the smallest such number. Therefore, the following ket vectors are linearly independent:

$$|x_j\rangle , \quad (H - \lambda_j)|x_j\rangle , \quad \dots , \quad (H - \lambda_j)^{n_j-1}|x_j\rangle . \quad (5.41)$$

### Multipole State

For a real eigenvalue  $\lambda$ , when there is no  $|x'\rangle$  such that  $|x\rangle = (H - \lambda)|x'\rangle$  and

$$(H - \lambda)^m |x\rangle = 0, \quad \langle x|(H - \lambda)^{m-1}|x\rangle \neq 0 \quad (m \geq 2), \quad (5.42)$$

we call  $|x\rangle$  an  $m$ th order multipole state, where “state” refers to the quantum-mechanical state.

Here we normalize  $|x\rangle$  so that

$$\langle x|(H - \lambda)^{m-1}|x\rangle = \epsilon, \quad \epsilon = \pm 1, \quad (5.43)$$

and then consider a linear combination of the linearly independent vectors in (5.41):

$$|z\rangle = |x\rangle + \sum_{k=1}^{m-1} c_k (H - \lambda)^k |x\rangle. \quad (5.44)$$

Choosing  $c_k$  suitably, we can arrange for the following equation to hold:

$$\langle z|(H - \lambda)^k |z\rangle = 0, \quad k = 0, 1, \dots, m-2. \quad (5.45)$$

Combining (5.43) and (5.45), we have

$$\langle z|(H - \lambda)^k |z\rangle = \epsilon \delta_{k, m-1}. \quad (5.46)$$

We now construct the  $m$  linearly independent vectors in (5.41). Here we choose the set starting from  $|z\rangle$  instead of  $|x\rangle$ , i.e.,

$$|z\rangle, \quad (H - \lambda)|z\rangle, \quad \dots, \quad (H - \lambda)^{m-1}|z\rangle. \quad (5.47)$$

For this basis,  $\eta$  and the representation  $h$  of  $H$  are given by

$$\eta = \begin{pmatrix} \mathbf{0} & \epsilon \\ \cdot & \cdot \\ \epsilon & \mathbf{0} \end{pmatrix}, \quad \eta^2 = 1, \quad (5.48)$$

$$h = \begin{pmatrix} \lambda & & & \mathbf{0} \\ 1 & \lambda & & \\ & 1 & \cdot & \cdot \\ & & \cdot & \lambda \\ \mathbf{0} & & & 1 & \lambda \end{pmatrix}, \quad h^\dagger = \eta h \eta. \quad (5.49)$$

In addition, the projection operator for such a subspace  $\mathcal{V}$  is given by

$$P(\mathcal{V}) = \epsilon \sum_{k=0}^{m-1} (H - \lambda)^k |z\rangle\langle z| (H - \lambda)^{m-k-1} . \quad (5.50)$$

The above works when  $\lambda$  is real.

When  $\lambda$  is complex, we consider  $\mathcal{V}(\lambda) \oplus \mathcal{V}(\lambda^*)$ . Using  $\eta$  and  $h$  in (5.48) and (5.49), the metric matrix and the matrix representation of  $H$  become

$$\begin{pmatrix} 0 & \eta \\ \eta & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} h & 0 \\ 0 & h^* \end{pmatrix} , \quad (5.51)$$

respectively.

### Subspace Describing Physical States

In various kinds of physical theories, especially those involving gauge fields, the indefinite metric is a necessary feature of the mathematical formulation. However, our observable world can be described by a positive-definite metric. The norm of the state vector corresponding to the probability is always positive. We thus pick out a subspace  $\mathcal{V}_{\text{phys}}$  from the state-vector space  $\mathcal{V}$ , viz.,

$$\mathcal{V}_{\text{phys}} \subset \mathcal{V} , \quad (5.52)$$

and interpret only the former as observable. An arbitrary vector  $|x\rangle$  in this subspace has the property

$$\langle x|x\rangle \geq 0 . \quad (5.53)$$

The difference with (5.4) is that the equation  $\langle x|x\rangle = 0$  does not necessarily require  $|x\rangle = 0$ , i.e., we accept that there can be a nonzero state vector with zero norm. Setting aside the issue of defining this subspace, what we can say is that, even if the parent space  $\mathcal{V}$  is non-degenerate, this subspace is degenerate.

To begin with, consider a vector  $|y\rangle$  with zero norm which belongs to this subspace, and another, arbitrary vector  $|x\rangle$  in this subspace, i.e.,

$$|x\rangle, |y\rangle \in \mathcal{V}_{\text{phys}} , \quad \langle y|y\rangle = 0 . \quad (5.54)$$

Since  $|x\rangle + a|y\rangle$  also belongs to this subspace, its norm satisfies

$$\langle x|x\rangle + 2\text{Re}(a\langle x|y\rangle) + |a|^2\langle y|y\rangle = \langle x|x\rangle + 2\text{Re}(a\langle x|y\rangle) \geq 0 .$$

In order for this inequality to hold for an arbitrary complex number  $a$ , we must have

$$\langle x|y\rangle = 0 . \quad (5.55)$$

It turns out that the subspace generated by  $|x\rangle$  and  $|y\rangle$  is degenerate. Indeed, the vector  $|y\rangle$  with zero norm is orthogonal to all the vectors in this subspace.

We thus consider the following direct-product decomposition:

$$\mathcal{V}_{\text{phys}} = \mathcal{V}_+ \oplus \mathcal{V}_0 , \quad (5.56)$$

where all the vectors belonging to  $\mathcal{V}_+$  have positive norm and the vectors belonging to  $\mathcal{V}_0$  have zero norm. Note that this decomposition is not unique. Although  $\mathcal{V}_0$  can be determined uniquely,  $\mathcal{V}_+$  cannot. This is because if  $|x\rangle$  has positive norm under the condition (5.54), then  $|x\rangle + a|y\rangle$  has the same positive norm, so the part of the vectors in  $\mathcal{V}_0$  becomes indeterminate. Consider  $|x'_i\rangle = |x_i\rangle + |y_i\rangle$ , where  $|x_i\rangle$  is a positive-norm vector,  $|y_i\rangle$  is a zero-norm vector, and both belong to  $\mathcal{V}_{\text{phys}}$ . Then,  $\langle x'_1|x'_2\rangle = \langle x_1|x_2\rangle$ , so if the metric is given,  $\mathcal{V}_+$  cannot be determined uniquely. This is because  $\mathcal{V}_{\text{phys}}$  is degenerate.

We now consider an operator  $T$  having  $\mathcal{V}_{\text{phys}}$  as an invariant subspace:

$$T\mathcal{V}_{\text{phys}} \subset \mathcal{V}_{\text{phys}} . \quad (5.57)$$

It is easy to show that

$$\langle x'_1|T|x'_2\rangle = \langle x_1|T|x_2\rangle . \quad (5.58)$$

Therefore, with respect to such an operator,  $|x'_i\rangle$  and  $|x_i\rangle$  are equivalent, so, formally, we can write

$$|x'_i\rangle \equiv |x_i\rangle \pmod{\mathcal{V}_0} . \quad (5.59)$$

An operator satisfying (5.57) is said to be an observable. Then we consider the following quotient space:

$$\mathcal{H}_{\text{phys}} = \mathcal{V}_{\text{phys}}/\mathcal{V}_0 . \quad (5.60)$$

If we now express a vector belonging to an equivalence class as a ket vector  $|x\rangle$ , we have

$$\langle x|x\rangle \geq 0 , \quad (5.61)$$

where equality holds only for

$$|x\rangle \equiv 0 \pmod{\mathcal{V}_0} .$$



In this sense, the quotient space  $\mathcal{H}_{\text{phys}}$  has a positive-definite metric, which allows a physical interpretation. Regarding the observable  $T$ , the matrix element can be defined by

$$\langle x_1 | T | x_2 \rangle = \langle x_1 + y_1 | T | x_2 + y_2 \rangle, \quad |y_1\rangle, |y_2\rangle \in \mathcal{V}_0. \quad (5.62)$$

$T$  is then viewed as an operator in  $\mathcal{H}_{\text{phys}}$ . Introducing a suitable operator  $Q$ ,  $\mathcal{V}_{\text{phys}}$  can be defined for practical purposes by

$$\mathcal{V}_{\text{phys}} = \{ |x\rangle \mid Q|x\rangle = 0, |x\rangle \in \mathcal{V} \}. \quad (5.63)$$

### 5.3 Free Electromagnetic Field in the Fermi Gauge

In order to quantize the electromagnetic field, we consider the Lagrangian density constructed by rewriting (3.221) in terms of a real vector field, setting the mass  $m$  to zero:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu}, \quad \text{where } F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (5.64)$$

In this case, the electric and magnetic fields are defined by

$$F_{4k} = -F_{k4} = iE_k, \quad F_{ij} = -F_{ji} = \epsilon_{ijk} H_k, \quad (5.65)$$

where  $\epsilon_{ijk}$  is 1,  $-1$ , or 0 depending on whether  $(ijk)$  is an even permutation of (123), an odd permutation of (123), or neither. Then (5.64) can be written in the form

$$\mathcal{L} = \frac{1}{2} (\mathbf{E}^2 - \mathbf{H}^2). \quad (5.66)$$

Using the variational principle to derive the equation of motion from this Lagrangian density, we obtain the source-free Maxwell equation:

$$\partial_\mu F_{\mu\nu} = \square A_\nu - \partial_\nu \partial_\mu A_\mu = 0. \quad (5.67)$$

Given any solution  $A_\mu$ , other solutions can be obtained by the gauge transformation

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \Lambda, \quad (5.68)$$

where  $\Lambda$  is an arbitrary real function of  $x_\mu$ . The fact that the solution allows such a gauge transformation comes from the fact that the operator  $D_{\mu\nu}(\partial)$  appearing previously has no inverse. Although we quantized the massive gauge field by

expressing the fourth component in terms of the other components, we cannot do this when the mass is zero. That is, if  $m \neq 0$ , we obtain the field equation

$$\partial_\mu A_\mu = 0 , \quad (5.69)$$

but if  $m = 0$ , then we cannot derive this.

Moreover, the quantity  $\pi_j$  canonically conjugate to  $A_j$  is

$$\pi_j = \frac{\partial \mathcal{L}}{\partial \dot{A}_j} = iF_{4j} = -E_j , \quad (5.70)$$

but there is clearly no way to define a quantity canonically conjugate to  $A_4$ . Here we have a novel situation, not encountered previously. One way to deal with it is to remove the fourth component via a gauge transformation and quantize only the remaining three components. However, with such a quantization method, Lorentz invariance is unclear, and it is better to find other ways where invariance is manifest. One such method is due to Fermi.

### Fermi's Method

Fermi introduced the Lagrangian density

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu} - \frac{1}{2}(\partial_\mu A_\mu)^2 . \quad (5.71)$$

Because of the additional term, the canonical conjugate quantity can be defined even for the fourth component. In fact, instead of (5.70), we have

$$\pi_j = iF_{4j} , \quad \pi_4 = i(\partial_\mu A_\mu) . \quad (5.72)$$

In addition, taking the Euler derivative, the field equation becomes

$$[\mathcal{L}]_{A_\mu} = \frac{\partial \mathcal{L}}{\partial A_\mu} - \frac{\partial}{\partial x_\nu} \left( \frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} \right) = D_{\mu\nu}(\partial)A_\nu = 0 , \quad (5.73)$$

where

$$D_{\mu\nu}(\partial) = \delta_{\mu\nu} \square . \quad (5.74)$$

Therefore, (5.73) turns out to be the massless Klein–Gordon equation, i.e.,

$$\square A_\mu = 0 . \quad (5.75)$$

Next, we define the inverse  $C$  of the operator  $D$  such that

$$D_{\alpha\beta}C_{\beta\gamma} = \delta_{\alpha\gamma} \square . \quad (5.76)$$

Clearly,

$$C_{\beta\gamma} = \delta_{\beta\gamma} . \quad (5.77)$$

The four-dimensional commutation relation is

$$[A_\alpha(x), A_\beta(y)] = iC_{\alpha\beta}D(x-y) = i\delta_{\alpha\beta}D(x-y) , \quad (5.78)$$

where the  $\Delta$ -function for  $m = 0$  is denoted by  $D$ . In this case,

$$D(x) = -\frac{1}{2\pi}\epsilon(x_0)\delta(x^2) , \quad D^{(1)}(x) = \frac{1}{2\pi^2}P\frac{1}{x^2} . \quad (5.79)$$

From this commutation relation, we obtain the following equal-time canonical commutator:

$$[A_\mu(x), \pi_\nu(y)] = i\delta_{\mu\nu}\delta^3(x-y) , \quad \text{where } x_0 = y_0 . \quad (5.80)$$

To derive this, we used the definition of  $\pi_\nu$  in (5.72) and the following property of the  $D$ -function (4.13):

$$\frac{\partial}{\partial x_0}D(x) = -\delta^3(x) , \quad \text{where } x_0 = 0 . \quad (5.81)$$

Additionally, it is straightforward to show that

$$[\partial_\mu A_\mu(x), \partial_\nu A_\nu(y)] = -i\Box_x D(x-y) = 0 . \quad (5.82)$$

The field equation here is the Klein–Gordon equation given by (5.75), and we require (5.69) to get Maxwell’s equation. However, when  $m \neq 0$ , (5.69) cannot be derived from the field equation. The method introduced by Fermi involves restricting the observable physical states by imposing an additional condition. His condition for an observable state  $|\Psi\rangle$  is

$$\langle\Psi|\partial_\mu A_\mu(x)|\Psi\rangle = 0 . \quad (5.83)$$

The set of such states forms a vector space. Since this condition is bilinear in  $|\Psi\rangle$ , we will introduce its linearization later on.

When the state is restricted in this way, Maxwell’s equation holds in terms of expectation values:

$$\langle\Psi|\partial_\mu F_{\mu\nu}|\Psi\rangle = \langle\Psi|\Box A_\nu(x)|\Psi\rangle - \partial_\nu\langle\Psi|\partial_\mu A_\mu(x)|\Psi\rangle = 0 . \quad (5.84)$$

The above can be achieved using the following Lagrangian density, which differs from (5.71) only by a divergence:

$$\mathcal{L}' = -\frac{1}{2}(\partial_\nu A_\mu)(\partial_\nu A_\mu) . \quad (5.85)$$

The energy and the momentum of this system are given by

$$H = \frac{1}{2} \int d^3x [ \dot{A}_\mu^2 + (\nabla A_\mu)^2 ] , \quad (5.86)$$

$$\mathbf{P} = - \int d^3x \dot{A}_\mu \nabla A_\mu . \quad (5.87)$$

These are common to both Lagrangians.

### Fourier Expansion

First, we implement the Fourier expansion of  $A_\mu$ , although in the following, we change the variable of the Fourier coefficient from the three-dimensional vector  $\mathbf{k}$  to the four-dimensional vector  $k$  :

$$A_\mu(x) = \sum_k \frac{1}{\sqrt{2k_0V}} [ e^{ik \cdot x} A_\mu(k) + e^{-ik \cdot x} A_\mu^\dagger(k) ] . \quad (5.88)$$

We then expand the Fourier coefficients  $A_\mu(k)$  in terms of polarization vectors:

$$A_\mu(k) = \sum_\lambda e_\mu^{(\lambda)} a(k, \lambda) , \quad A_\mu^\dagger(k) = \sum_\lambda e_\mu^{(\lambda)} a^\dagger(k, \lambda) . \quad (5.89)$$

Here we define four polarization vectors such that

$$\begin{aligned} e_4^{(1)} = e_4^{(2)} = e_4^{(3)} = 0 , \quad e_4^{(4)} = 1 , \\ e^{(3)} = e^{(1)} \times e^{(2)} = \frac{\mathbf{k}}{k_0} , \quad e^{(1)} \cdot k = e^{(2)} \cdot k = 0 . \end{aligned} \quad (5.90)$$

Thus,  $e^{(1)}$  and  $e^{(2)}$  correspond to a transverse wave,  $e^{(3)}$  to a longitudinal wave, and  $e^{(4)}$  to a scalar wave. These four vectors form a complete orthogonal system:

$$e_\mu^{(\lambda)} e_\mu^{(\lambda')} = \delta_{\lambda\lambda'} , \quad e_\mu^{(\lambda)} e_\nu^{(\lambda)} = \delta_{\mu\nu} . \quad (5.91)$$

We can now write the Fourier expansion in the form

$$A_\mu(x) = \sum_k \sum_\lambda \frac{1}{\sqrt{2k_0V}} [ e_\mu^{(\lambda)} a(k, \lambda) e^{ik \cdot x} + e_\mu^{(\lambda)} a^\dagger(k, \lambda) e^{-ik \cdot x} ] . \quad (5.92)$$

In the Pauli metric, the adjoint operator  $a^\dagger(k, \lambda)$  of  $a(k, \lambda)$  is the Hermitian conjugate for  $\lambda = 1, 2, 3$ , while  $ia(k, 4)$  and  $ia^\dagger(k, 4)$  are Hermitian conjugates for  $\lambda = 4$ . The commutation relation is

$$[a(k, \lambda), a^\dagger(k', \lambda')] = \delta_{\lambda\lambda'} \delta_{\mathbf{k}, \mathbf{k}'} . \quad (5.93)$$

The  $a$ 's commute among themselves and so do the  $a^\dagger$ 's.

Up to c-numbers, the energy and momentum of the whole system are given by

$$H = \sum_{\mathbf{k}} k_0 \sum_{\lambda} a^\dagger(k, \lambda) a(k, \lambda) , \quad (5.94)$$

$$P_j = \sum_{\mathbf{k}} k_j \sum_{\lambda} a^\dagger(k, \lambda) a(k, \lambda) . \quad (5.95)$$

For  $\lambda = 1, 2, 3$ , the particle number operator is

$$n(k, \lambda) = a^\dagger(k, \lambda) a(k, \lambda) . \quad (5.96)$$

Incidentally,  $a^\dagger(k, 4)$  is not the Hermitian conjugate of  $a(k, 4)$ , but a so-called *adjoint operator*. Setting

$$a(k, 4) = ia(k, 0) , \quad a^\dagger(k, 4) = ia^\dagger(k, 0) , \quad (5.97)$$

then  $a^\dagger(k, 0)$  becomes the Hermitian conjugate of  $a(k, 0)$ . As is clear from the commutation relation

$$[a(k, 0), a^\dagger(k, 0)] = -1 , \quad (5.98)$$

the scalar wave has indefinite metric. If  $a(k, 0)$  and  $a^\dagger(k, 0)$  are interpreted as an annihilation operator and a creation operator, respectively, then for the vacuum  $|0\rangle$ , we obtain

$$a(k, 0)|0\rangle = 0 . \quad (5.99)$$

We introduce the  $n$ -particle state of the scalar wave by

$$|n\rangle = \frac{1}{\sqrt{n!}} [a^\dagger(k, 0)]^n |0\rangle . \quad (5.100)$$

Its norm is given by

$$\langle n|n\rangle = (-1)^n \langle 0|0\rangle = (-1)^n .$$

This shows why the indefinite metric turns up. The particle number operator of the scalar wave is now given by

$$n(k, 0) = a^\dagger(k, 0)a(k, 0) . \quad (5.101)$$

When  $k$  is fixed, there are four polarization states and the metric matrix  $\eta(k)$  corresponding to the four states is

$$\eta(k) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} . \quad (5.102)$$

It is clear that the indefinite metric of the one-particle state for the scalar wave has its origin in the Minkowski metric.

## 5.4 Lorenz Condition and Physical State Space

In the previous section, we imposed the additional condition (5.83) to pick out physically observable states. It has been written in bilinear form, so we now seek a way to rewrite it as a linear form.

Since the free field can be decomposed into a positive frequency part  $A_\mu^{(+)}(x)$  corresponding to the annihilation operators and a negative frequency part  $A_\mu^{(-)}(x)$  corresponding to the creation operators, we can write

$$A_\mu(x) = A_\mu^{(+)}(x) + A_\mu^{(-)}(x) . \quad (5.103)$$

The vacuum  $|0\rangle$  then satisfies

$$A_\mu^{(+)}(x)|0\rangle = 0 . \quad (5.104)$$

Restricting the system to the free electromagnetic field, it turns out that the state-vector space  $\mathcal{V}$  can be created by applying  $A_\mu^{(-)}(x)$  to the vacuum  $|0\rangle$ . Since the space  $\mathcal{V}$  has indefinite metric, in order to obtain a physical interpretation, we must introduce the subspace  $\mathcal{V}_{\text{phys}}$  discussed in Sect. 5.2.

One possibility here is to choose the Coulomb gauge, in which the scalar wave reproducing the negative term in the metric matrix is dropped at the outset, but what we would like is rather to eliminate the indefiniteness of the metric while maintaining the manifest covariance of the theory. This can be done by adopting the *Lorenz condition*. In (5.63), we mentioned a suitable operator  $Q$  to select  $\mathcal{V}_{\text{phys}}$ , but

in the following we shall see that there may exist more than one operator with the property (5.63). Then we define the subspace  $\mathcal{V}_{\text{phys}}$  as follows:

$$\mathcal{V}_{\text{phys}} = \{|\Phi\rangle \mid \partial_\mu A_\mu^{(+)}(x)|\Phi\rangle = 0, \quad |\Phi\rangle \in \mathcal{V}\}, \quad (5.105)$$

where  $x$  runs over all the points in four-dimensional space-time. The extra condition appearing here, i.e.,

$$\partial_\mu A_\mu^{(+)}(x)|\Phi\rangle = 0, \quad (5.106)$$

is called the *Lorenz condition*. It is clear that a state which satisfies this condition also obeys (5.83), i.e.,

$$\langle\Phi|\partial_\mu A_\mu(x)|\Phi\rangle = 0.$$

What we need to prove first is that arbitrary vectors which belong to  $\mathcal{V}_{\text{phys}}$  have nonzero norm, i.e.,

$$\langle\Phi|\Phi\rangle \geq 0. \quad (5.107)$$

From the Fourier expansion (5.92),

$$\partial_\mu A_\mu^{(+)}(x) = i \sum_k \frac{1}{\sqrt{2k_0V}} e^{ik \cdot x} \sum_\lambda k_\mu e_\mu^{(\lambda)} a(k, \lambda). \quad (5.108)$$

The sum over  $\lambda$  is

$$\sum_\lambda k_\mu e_\mu^{(\lambda)} a(k, \lambda) = |k| [a(k, 3) - a(k, 0)], \quad (5.109)$$

where  $a(k, 3)$  is the operator annihilating the longitudinal wave. We then introduce the following operators:

$$a(k) = \frac{1}{\sqrt{2}} [a(k, 3) - a(k, 0)], \quad b(k) = \frac{1}{\sqrt{2}} [a(k, 3) + a(k, 0)]. \quad (5.110)$$

Therefore, the Lorenz condition ends up in the form

$$a(k)|\Phi\rangle = 0. \quad (5.111)$$

For a given  $k$ , we obtain the commutation relations

$$[a, a^\dagger] = [b, b^\dagger] = 0, \quad [a, b^\dagger] = [b, a^\dagger] = 1. \quad (5.112)$$

These operators commute with  $a(k, 1)$ ,  $a(k, 2)$ ,  $a^\dagger(k, 1)$ , and  $a^\dagger(k, 2)$ .  $\mathcal{V}$  can be generated by applying polynomials in  $a^\dagger(k, 1)$ ,  $a^\dagger(k, 2)$ ,  $a^\dagger(k)$ , and  $b^\dagger(k)$  to the vacuum state. On the other hand, it is clear from (5.112) that  $\mathcal{V}_{\text{phys}}$  can be generated by applying the operators  $a^\dagger(k, 1)$ ,  $a^\dagger(k, 2)$ , and  $a^\dagger(k)$  to  $|0\rangle$ . We expand the state  $|\Phi_j\rangle$  belonging to  $\mathcal{V}_{\text{phys}}$  in terms of the number of  $a^\dagger$  operators:

$$|\Phi_j\rangle = |\Phi_0\rangle + |\Phi_1\rangle + |\Phi_2\rangle + \cdots, \quad (5.113)$$

where  $|\Phi_j\rangle$  denotes the state containing transverse photons and  $j$   $a$ -photons. The inner product of two physical states is

$$\langle\Phi'|\Phi\rangle = \langle\Phi'_0|\Phi_0\rangle, \quad (5.114)$$

since the norm is zero unless the number of  $a$ -photons is zero. For example, using (5.112), the inner product of states  $a^\dagger(k)|\Phi_T\rangle$  and  $a^\dagger(k')|\Phi'_T\rangle$  containing single  $a$ -photons, while both  $|\Phi_T\rangle$  and  $|\Phi'_T\rangle$  contain only transverse photons, is

$$\langle\Phi'_T|a(k')a^\dagger(k)|\Phi_T\rangle = \langle\Phi'_T|a^\dagger(k)a(k')|\Phi_T\rangle = 0.$$

This is because when the annihilation operator  $a(k')$  acts on the state  $|\Phi_T\rangle$ , it destroys it. This is the reason why (5.114) holds. If  $\Phi$  belongs to  $\mathcal{V}_{\text{phys}}$ , then (5.114) implies

$$\langle\Phi|\Phi\rangle = \langle\Phi_0|\Phi_0\rangle \geq 0. \quad (5.115)$$

If this vanishes, it does not mean  $|\Phi\rangle = 0$ , but rather  $|\Phi_0\rangle = 0$ . In the decomposition (5.56), it turns out that  $\mathcal{V}_+$  consists of state vectors containing only transverse photons, while  $\mathcal{V}_0$  consists of state vectors containing  $a$ -photons. However, since a time axis is chosen in the decomposition (5.56), the vector space  $\mathcal{V}_+$  is not Lorentz invariant.

Note, however, that the quotient space  $\mathcal{H}_{\text{phys}}$  defined by (5.60) is Lorentz invariant. Moreover, (5.114) can also be written

$$\langle\Phi'|\Phi\rangle = \langle\Phi'|P(\mathcal{V}_+)|\Phi\rangle, \quad \text{where } |\Phi\rangle, |\Phi'\rangle \in \mathcal{V}_{\text{phys}}. \quad (5.116)$$

Here  $P(\mathcal{V}_+)$  is the operator projecting onto the subspace  $\mathcal{V}_+$ . Although  $\mathcal{V}_+$  is not Lorentz invariant, (5.116) is independent of the choice of  $\mathcal{V}_+$ , so this inner product is itself Lorentz invariant, and so too is  $\mathcal{V}_{\text{phys}}$ .

We now give some examples of observables, as defined in Sect. 5.2. For instance, the following are observables:

$$F_{\mu\nu}(x), \quad \partial_\mu A_\mu(x). \quad (5.117)$$



This is because, as is easily shown, we have

$$[F_{\mu\nu}(x), \partial_\rho A_\rho^{(+)}(y)] = 0, \quad [\partial_\mu A_\mu(x), \partial_\rho A_\rho^{(+)}(y)] = 0. \quad (5.118)$$

Thus, if  $|\Phi\rangle$  belongs to  $\mathcal{V}_{\text{phys}}$ , then both  $F_{\mu\nu}(x)|\Phi\rangle$  and  $\partial_\mu A_\mu(x)|\Phi\rangle$  belong to  $\mathcal{V}_{\text{phys}}$ . Therefore, if  $|\Phi\rangle$  and  $|\Phi'\rangle$  belong to  $\mathcal{V}_{\text{phys}}$ , we have

$$\langle\Phi'|F_{\mu\nu}(x)F_{\rho\sigma}(y)|\Phi\rangle = \langle\Phi'|F_{\mu\nu}(x)P(\mathcal{V}_+)F_{\rho\sigma}(y)|\Phi\rangle. \quad (5.119)$$

In the presence of interactions, an  $S$ -matrix can be defined. This is also an observable, satisfying

$$[S, \partial_\rho A_\rho^{(+)}(x)] = [S, \partial_\rho A_\rho(x)] = 0. \quad (5.120)$$

## 5.5 Free Electromagnetic Field: Generalization of Gauge Choices

So far we have discussed the quantization of the electromagnetic field in the Fermi gauge. It would be easier if the story were self-contained in this gauge, but later, when we come to consider renormalization for instance, it is useful to have the gauge choice generalized in advance. Therefore, generalizing (5.71), we introduce the Lagrangian density

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu} - \frac{1}{2\alpha}(\partial_\mu A_\mu)^2. \quad (5.121)$$

The parameter  $\alpha$  is called a *gauge parameter*. The value  $\alpha = 1$  corresponds to the Fermi gauge. We then consider the Euler derivative:

$$[\mathcal{L}]_{A_\mu} = D_{\mu\nu}(\partial)A_\nu, \quad (5.122)$$

where

$$D_{\mu\nu}(\partial) = \delta_{\mu\nu}\square + \frac{1-\alpha}{\alpha}\partial_\mu\partial_\nu. \quad (5.123)$$

In order to quantize the field, we define the differential operator  $C$  to be such that

$$C_{\mu\sigma}D_{\sigma\nu} = D_{\mu\sigma}C_{\sigma\nu} = \delta_{\mu\nu}\square. \quad (5.124)$$

From this, it follows that

$$C_{\mu\nu}(\partial) = \delta_{\mu\nu} - (1-\alpha)\frac{\partial_\mu\partial_\nu}{\square}. \quad (5.125)$$

One feature of the general gauge is that  $\square$  appears in the denominator, and this yields the dipole state or dipole ghost discussed in Sect. 5.2. The reason it is called a ghost is that it is related to the indefinite metric, as mentioned at the end of Sect. 4.4. Using this  $C$ , we obtain the commutation relation

$$\begin{aligned} [A_\mu(x), A_\nu(y)] &= iC_{\mu\nu}(\partial)D(x-y) \\ &= i\delta_{\mu\nu}D(x-y) + i(1-\alpha)\partial_\mu\partial_\nu E(x-y), \end{aligned} \quad (5.126)$$

where  $E$  is defined by

$$\square E(x) = -D(x), \quad E(x) = -\frac{1}{8\pi}\epsilon(x_0)\theta(-x^2). \quad (5.127)$$

This function can be also written in the form

$$E(x) = -\left[ \frac{\partial}{\partial m^2} \Delta(x, m^2) \right]_{m^2=0}, \quad (5.128)$$

where, as for the function  $\Delta(x)$  defined in Chap. 4, we have written the mass dependence  $m^2$  explicitly as  $\Delta(x, m^2)$ . We will use this definition later, too.

Although the field equation is different from both Maxwell's equation and the Klein–Gordon equation, the latter can be obtained by imposing the Lorenz condition, which is the same as the Fermi gauge.

From the field equation  $D_{\mu\nu}(\partial)A_\nu = 0$ , we have

$$\partial_\mu D_{\mu\nu}(\partial)A_\nu = \frac{1}{\alpha}\square\partial_\nu A_\nu = 0, \quad \square(\partial_\nu A_\nu) = 0. \quad (5.129)$$

Combining the field equation and (5.129),

$$\square D_{\mu\nu}(\partial)A_\nu = \square^2 A_\mu + \frac{1-\alpha}{\alpha}\partial_\mu\square(\partial_\nu A_\nu) = \square^2 A_\mu = 0. \quad (5.130)$$

Note that  $\square A_\mu \neq 0$ . This is clear from the commutation relation

$$\begin{aligned} [\square A_\mu(x), A_\nu(y)] &= i\square C_{\mu\nu}(\partial)D(x-y) \\ &= -i(1-\alpha)\partial_\mu\partial_\nu D(x-y) \neq 0, \end{aligned} \quad (5.131)$$

where we have assumed that  $\alpha$  is different from 1. Therefore,

$$\square A_\mu \neq 0, \quad \square^2 A_\mu = 0. \quad (5.132)$$

These two equations imply the existence of the dipole ghost, as we shall now show.

### Dipole Ghost

Recall that multipole states with indefinite metric were defined by using a Hermitian operator. In the following, we choose as Hermitian operator the mass operator:

$$H = -P_\mu^2 . \quad (5.133)$$

Then it can be proven that the state  $A_\alpha(x)|0\rangle$  includes a dipole state. To show this, we use (3.29) and  $P_\mu|0\rangle = 0$  to obtain

$$\begin{aligned} P_\mu A_\alpha(x)|0\rangle &= [P_\mu, A_\alpha(x)]|0\rangle + A_\alpha(x)P_\mu|0\rangle \\ &= i\partial_\mu A_\alpha(x)|0\rangle . \end{aligned}$$

Therefore, repeating the transformation above,

$$H A_\alpha(x)|0\rangle = (-P_\mu^2)A_\alpha(x)|0\rangle = \square A_\alpha(x)|0\rangle . \quad (5.134)$$

As is clear from the fact that the vacuum expectation value of (5.131) is nonzero, this state does not vanish. However, the norm of this state is zero, as we will show below. From (5.132),

$$\begin{aligned} \langle 0|\square A_\alpha(x)\square A_\beta(y)|0\rangle &= \langle 0|A_\alpha(x)(-P_\mu^2)(-P_\nu^2)A_\beta(y)|0\rangle \\ &= \langle 0|A_\alpha(x)\square^2 A_\beta(y)|0\rangle \\ &= 0 . \end{aligned} \quad (5.135)$$

Consequently,

$$H A_\alpha(x)|0\rangle \neq 0 , \quad H^2 A_\alpha(x)|0\rangle = 0 . \quad (5.136)$$

It is obvious from the definition (5.42) that this implies the existence of the dipole ghost for the eigenvalue  $\lambda = 0$  of  $H$ .

### Lorenz Condition

When a dipole ghost exists, we cannot carry out the Fourier expansion. In fact, the Fourier expansion cannot be carried out for the solution of (5.132). This makes it harder to obtain an intuitive understanding of the dipole ghost. In order to investigate its structure intuitively, we consider

$$D_{\mu\nu}^{-1} = \frac{C_{\mu\nu}}{\square} = \frac{\delta_{\mu\nu}}{\square} - (1 - \alpha) \frac{\partial_\mu \partial_\nu}{\square^2} . \quad (5.137)$$

The main issue is the treatment of the second term. We first write the above equation in the form

$$D_{\mu\nu}^{-1} = \frac{\delta_{\mu\nu}}{\square} - (1 - \alpha) \partial_\mu \partial_\nu \lim_{\epsilon \rightarrow 0} \frac{1}{\square(\square - \epsilon^2)} . \quad (5.138)$$

In later discussions, we will treat  $\epsilon$  as nonzero, but small, i.e.,

$$D_{\mu\nu}^{-1} = \frac{\delta_{\mu\nu}}{\square} - (1 - \alpha) \frac{1}{\epsilon^2} \left( \frac{1}{\square - \epsilon^2} - \frac{1}{\square} \right) \partial_\mu \partial_\nu . \quad (5.139)$$

For the inverse  $D(\partial, \epsilon)$  of this expression, we have

$$D_{\mu\nu}(\partial, \epsilon) = \delta_{\mu\nu} \square + \frac{\square}{\alpha \square - \epsilon^2} (1 - \alpha) \partial_\mu \partial_\nu . \quad (5.140)$$

Then the commutation relation corresponding to (5.137) is

$$[A_\mu(x), A_\nu(y)] = i \left( \delta_{\mu\nu} + \frac{1 - \alpha}{\epsilon^2} \partial_\mu \partial_\nu \right) D(x - y) - i \frac{1 - \alpha}{\epsilon^2} \partial_\mu \partial_\nu \nabla(x - y, \epsilon^2) . \quad (5.141)$$

We now write  $A_\mu$  in such a way as to reproduce this commutation relation:

$$A_\mu(x) = a_\mu(x) + \frac{1 - \alpha}{2\epsilon^2} \partial_\mu \partial_\nu a_\nu(x) + \frac{\sqrt{1 - \alpha}}{\epsilon} \partial_\mu b(x) , \quad (5.142)$$

where

$$[a_\mu(x), a_\nu(y)] = i \delta_{\mu\nu} D(x - y) , \quad [b(x), b(y)] = i \Delta(x - y, \epsilon) . \quad (5.143)$$

Hence,  $a_\mu$  is the electromagnetic field in the Fermi gauge and  $b$  is a real scalar field with mass  $\epsilon$ . Of course, these two fields are independent:

$$[a_\mu(x), b(y)] = 0 . \quad (5.144)$$

The field equations are

$$\square a_\mu = 0 , \quad (\square - \epsilon^2) b = 0 . \quad (5.145)$$

From this, we can derive the following equation of motion:

$$D_{\mu\nu}(\partial, \epsilon) A_\nu = 0 . \quad (5.146)$$

In the discussion above, we have assumed  $1 - \alpha > 0$ , but if  $1 - \alpha < 0$ , the coefficient of  $\partial_\mu b$  in (5.142) becomes  $\sqrt{\alpha - 1}/\epsilon$ . It turns out that  $b$  obeys the indefinite metric. In the following, we discuss the case where  $1 - \alpha > 0$ . Now,

$$\partial_\mu A_\mu = \partial_\mu a_\mu + \epsilon \sqrt{1 - \alpha} b . \quad (5.147)$$

As long as  $\epsilon \neq 0$  in the Fourier expansion,  $a_\mu$  and  $b$  are fields of different masses and the Lorenz condition

$$\partial_\mu A_\mu^{(+)}(x)|\Phi\rangle = 0 \quad (5.148)$$

can be decomposed into the following two equations:

$$\partial_\mu a_\mu^{(+)}(x)|\Phi\rangle = 0, \quad b^{(+)}(x)|\Phi\rangle = 0. \quad (5.149)$$

As can be seen, the new degree of freedom  $b$  appears in the Fermi gauge, but the second condition above removes it. Thus, in the end, the condition is reduced to the Lorenz condition in the Fermi gauge expressed by the first relation in (5.149). Due to the assumption that  $\epsilon \neq 0$ , we can apply the Fourier expansion in the analysis above.

From the discussion above, the Fermi gauge with  $\alpha = 1$  is obviously rather special. The Landau gauge with  $\alpha = 0$  is also special since  $C$  then consists only of the transverse part.

It was Dirac who introduced the indefinite metric for the first time [84], while Gupta and Bleuler did this when quantizing the electromagnetic field [85, 86].

# Chapter 6

## Quantization of Interacting Systems



So far, we have considered the quantization of free fields, i.e., fields without interactions. In this case, the field equation is linear and can easily be solved. Thanks to this, it can be shown that canonical quantization is a relativistically-invariant quantization method which does not depend on the way we choose our coordinates. However, when quantizing interacting fields, we begin with the equal-time commutation relations and the relativistic invariance is not apparent. In this case, there is a choice between the Schrödinger picture and the Heisenberg picture. But whichever is adopted, it is hard to reach a clear conclusion regarding the issue of relativistic invariance when the canonical quantization method is used. It is the interaction-picture method introduced by Tomonaga [87] and Schwinger [88] that provides a precise solution to this problem.

In this picture, the invariance of both the commutation relations and the Schrödinger equation become clear. Using this manifestly Lorentz-covariant picture, we can then move on to the field theory in terms of Heisenberg's picture. Here, we will explain the supermultiple-time formulation in the interaction picture.

### 6.1 Tomonaga–Schwinger Equation

We begin with Schrödinger's equation for interacting fields:

$$i \frac{\partial}{\partial t} \Psi(t) = H \Psi(t) . \quad (6.1)$$

Here  $H$  is the Hamiltonian of the total system, consisting of a free part and an interaction part:

$$H = \int d^3x \mathcal{H}(\mathbf{x}) = \int d^3x \mathcal{H}_f(\mathbf{x}) + \int d^3x \mathcal{H}_{\text{int}}(\mathbf{x}) = H_f + H_{\text{int}} , \quad (6.2)$$

where  $f$  and  $int$  indicate the free and interaction parts, respectively. In fact, this separation is not trivial, as will be explained later when we discuss renormalization. However, we shall not go into this yet.

Next, we introduce the unitary transformation

$$U(t) = \exp(i t H_f) , \quad (6.3)$$

and define the field operator

$$\varphi_\alpha(x) = \varphi_\alpha(\mathbf{x}, t) = U(t)\varphi_\alpha(\mathbf{x})U(t)^{-1} . \quad (6.4)$$

Then,  $\varphi_\alpha(x)$  satisfies the free field equation and the commutation relations of the free field:

$$D_{\alpha\beta}(\partial)\varphi_\beta(x) = 0 , \quad (6.5)$$

$$[\varphi_\alpha(x), \varphi_\beta^\dagger(y)]_\pm = iC_{\alpha\beta}(\partial)\Delta(x - y) . \quad (6.6)$$

It turns out that the free field obeys the methods of relativistically invariant quantization.

We introduce a new state vector  $\Phi(t)$  by

$$\Phi(t) = U(t)\Psi(t) , \quad (6.7)$$

which satisfies

$$i\frac{\partial}{\partial t}\Phi(t) = \left[ U(t) \int d^3x \mathcal{H}_{int}(\mathbf{x})U(t)^{-1} \right] \Phi(t) . \quad (6.8)$$

We consider  $\mathcal{H}_{int}(x)$  as a polynomial in the field operators  $\varphi_\alpha(x)$ :

$$\begin{aligned} U(t)\mathcal{H}_{int}[\varphi_\alpha(\mathbf{x}), \dots]U(t)^{-1} &= \mathcal{H}_{int}[U(t)\varphi_\alpha(\mathbf{x})U(t)^{-1}, \dots] \\ &= \mathcal{H}_{int}[\varphi_\alpha(x), \dots] . \end{aligned} \quad (6.9)$$

Then we can rewrite (6.8) in the form

$$i\frac{\partial}{\partial t}\Phi(t) = \int d^3x \mathcal{H}_{int}[\varphi_\alpha(x), \dots]\Phi(t) . \quad (6.10)$$

In this picture, the field operator obeys the free field equation, but the state vector obeys Schrödinger's equation with a Hamiltonian which is the interaction part of the whole Hamiltonian. This is called the *interaction picture*. Clearly, for free fields, (6.10) shows that  $\Phi$  is independent of time, and the interaction picture coincides with the Heisenberg picture.

The relativistic invariance of (6.10) is not apparent. This is because we have fixed the time axis. Tomonaga showed how to rewrite this in a Lorentz invariant form using the analogy with Dirac’s multi-time formulation [89]. Here we will derive it directly without explaining Dirac’s multi-time theory.

First, the way the state vector  $\Phi(t)$  has been described means that all observations must be conducted at the same time for the given time axis. This restriction seems much too strong. It would be good if observations at various points in space did not have to be carried out at the same time. For instance, it would be good if observations could be carried out at  $t_1$  for the point  $(x_1, y_1, z_1)$  and at  $t_2$  for the point  $(x_2, y_2, z_2)$ . In this case, if the distance between the two points is space-like, these observations will not interfere with each other, according to Einstein’s causality, i.e.,

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - (t_1 - t_2)^2 > 0 . \tag{6.11}$$

We then consider the time which is not necessarily the same at each point in space, i.e.,

$$t = f(x, y, z) . \tag{6.12}$$

This relation specifies a hypersurface in four-dimensional space-time. In addition, we assume that two points on this hypersurface are spatially separated. We call this a *space-like hypersurface*, and denote it by  $\sigma$ . When we introduce the state vector  $\Phi[\sigma]$ , what type of equation does it obey? This will be a generalization of  $\Phi(t)$  corresponding to  $t = \text{const}$ . To study this, we write (6.10) in the form

$$\Phi(t + dt) = \Phi(t) - idt \int d^3x \mathcal{H}_{\text{int}}[\varphi_\alpha(x), \dots] \Phi(t) . \tag{6.13}$$

Then in going from the hypersurface at  $t$  to the one at  $t + dt$ , we consider that at each point  $(x, y, z)$  in space we only shift the volume element with infinitesimal volume  $d^3x$  by  $dt$ . However, we change it smoothly enough to maintain the condition of being space-like. We do this at all points (see Fig. 6.1). We assume that, for a change  $d\omega = d^3x dt$  in the infinitesimal volume, the state vector is multiplied by the factor

$$1 - id\omega \mathcal{H}_{\text{int}}(x) . \tag{6.14}$$

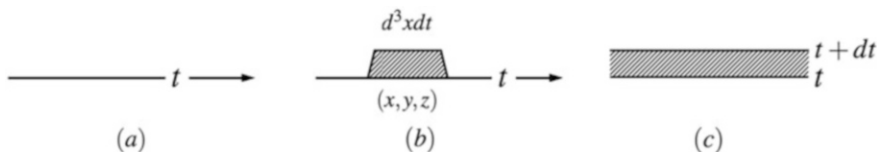


Fig. 6.1 Variation of the hypersurface



Therefore, multiplying by these factors in such a way as to cover all points in space and neglecting infinitesimal terms of order greater than or equal to  $d\omega^2$ , we obtain (6.13). Generalization to the hypersurface is clear. If the hypersurface  $\sigma$  is altered to  $\sigma'$  solely through a variation of the infinitesimal volume  $d\omega$  near the space-time point  $x$ , and if both  $\sigma$  and  $\sigma'$  are space-like hypersurfaces, then

$$\Phi[\sigma'] = [1 - i d\omega \mathcal{H}_{\text{int}}(x)] \Phi[\sigma], \quad (6.15)$$

where  $\sigma'$  has been taken to be a future hypersurface of  $\sigma$ . Here we have used the abbreviation

$$\mathcal{H}_{\text{int}}(x) = \mathcal{H}_{\text{int}}[\varphi_\alpha(x), \dots]. \quad (6.16)$$

Note that the notion of the space-like hypersurface is independent of the choice of coordinate system. The four-dimensional infinitesimal volume is also a Lorentz-invariant quantity and the Hamiltonian density  $\mathcal{H}_{\text{int}}(x)$  is a scalar for simple cases. Therefore, (6.15) is independent of the way we choose the coordinate system. In order to write this equation in a differential form, we define the following functional derivative:

$$\lim_{d\omega \rightarrow 0} \frac{\Phi[\sigma'] - \Phi[\sigma]}{d\omega} = \frac{\delta\Phi[\sigma]}{\delta\sigma(x)}. \quad (6.17)$$

Using this symbol, (6.15) can be written in the form

$$i \frac{\delta\Phi[\sigma]}{\delta\sigma(x)} = \mathcal{H}_{\text{int}}(x) \Phi[\sigma]. \quad (6.18)$$

This is the *Tomonaga–Schwinger equation*. Because we choose a different time at each point in space, this is called the *super multiple-time formulation*.

This theory was first formulated by Tomonaga [87]. It worked well when applied to QED, especially for the renormalized theory. In the United States, its usefulness was also recognized by Schwinger, and it became a central topic in the postwar period.

In simple cases, the Hamiltonian density depends only on the point  $x$  and is independent of  $\sigma$ . Let us consider the integrability condition in this case. We thus rewrite the condition

$$\frac{\delta}{\delta\sigma(x)} \frac{\delta}{\delta\sigma(y)} \Phi[\sigma] = \frac{\delta}{\delta\sigma(y)} \frac{\delta}{\delta\sigma(x)} \Phi[\sigma], \quad (6.19)$$

using (6.18), and find

$$\mathcal{H}_{\text{int}}(y) \mathcal{H}_{\text{int}}(x) - \mathcal{H}_{\text{int}}(x) \mathcal{H}_{\text{int}}(y) = 0, \quad (6.20)$$

where the points  $x$  and  $y$  are on the space-like hypersurface  $\sigma$ . This says that the Hamiltonian densities at two space-like separated points commute, and it is called an *integrability condition*. It is closely related to the notion of microscopic causality.

Next, let us consider an initial-value problem. The state vector is given on the space-like hypersurface  $\sigma_0$ :

$$\Phi[\sigma_0] = \Phi . \quad (6.21)$$

With this initial condition, the state vector on a general space-like hypersurface is given by

$$\Phi[\sigma] = U[\sigma, \sigma_0]\Phi , \quad (6.22)$$

and we choose  $U$  in such a way as to satisfy the initial condition and the Tomonaga–Schwinger (TS) equation. First, since (6.22) satisfies the TS equation,

$$i \frac{\delta}{\delta\sigma(x)} U[\sigma, \sigma_0] = \mathcal{H}_{\text{int}}(x) U[\sigma, \sigma_0] . \quad (6.23)$$

Additionally, corresponding to the initial condition (6.21),

$$U[\sigma_0, \sigma_0] = 1 . \quad (6.24)$$

When it has this form,  $U$  is called a *generalized transformation functional*. This becomes unitary if the Hamiltonian density is Hermitian. To show this, we write the Hermitian conjugate equation of the equation (6.23) as follows:

$$-i \frac{\delta}{\delta\sigma(x)} U^\dagger[\sigma, \sigma_0] = U^\dagger[\sigma, \sigma_0] \mathcal{H}_{\text{int}}(x) . \quad (6.25)$$

Combining (6.23) and (6.25),

$$\frac{\delta}{\delta\sigma(x)} U^\dagger[\sigma, \sigma_0] U[\sigma, \sigma_0] = 0 .$$

Thus, the product  $U^\dagger U$  does not depend on  $\sigma$ , and we can substitute  $\sigma_0$  for  $\sigma$  and use (6.24) to obtain

$$U^\dagger[\sigma, \sigma_0] U[\sigma, \sigma_0] = 1 . \quad (6.26)$$

Additionally, as is clear from the definition (6.22), the following composition law holds:

$$U[\sigma_2, \sigma_0] = U[\sigma_2, \sigma_1] U[\sigma_1, \sigma_0] . \quad (6.27)$$

Since this equation implies that the time evolution of the system is independent of its history, this expresses the causality condition. Moreover, if we set  $\sigma_2 = \sigma_0$  in (6.27), then taking into account (6.26),

$$U[\sigma_0, \sigma_1] = (U[\sigma_1, \sigma_0])^{-1} = U^\dagger[\sigma_1, \sigma_0] . \quad (6.28)$$

Hence,  $U$  is a unitary operator.

In particular, if we choose a space-like hypersurface  $\sigma$  and then introduce a time axis, we can write  $U$  as

$$U(t_2, t_1) .$$

In this case, we can write the expectation value of a given operator in the interaction picture (but not using the bra-ket symbol here):

$$(\Phi(t), \varphi_\alpha(x)\Phi(t)) = (\Phi, U(t, t_0)^{-1}\varphi_\alpha(x)U(t, t_0)\Phi) . \quad (6.29)$$

Due to the time-independence of the state vector  $\Phi$ , this tells us that

$$U(t, t_0)^{-1}\varphi_\alpha(x)U(t, t_0) = \varphi_\alpha^{(H)}(x) \quad (6.30)$$

specifies an operator in the Heisenberg picture. Given the Heisenberg operator  $\varphi_\alpha^{(H)}(x)$ , it turns out that we can define various interaction pictures depending on the choice of  $t_0$ . We write the operators in the interaction picture corresponding to  $t_0 \rightarrow -\infty$  and  $t_0 \rightarrow \infty$  as  $\varphi_\alpha^{\text{in}}$  and  $\varphi_\alpha^{\text{out}}$ , respectively. Then,

$$\begin{aligned} \varphi_\alpha^{(H)}(x) &= U(t, -\infty)^{-1}\varphi_\alpha^{\text{in}}(x)U(t, -\infty) \\ &= U(t, \infty)^{-1}\varphi_\alpha^{\text{out}}(x)U(t, \infty) . \end{aligned} \quad (6.31)$$

Although it is not strictly rigorous, taking  $t \rightarrow \infty$  and  $U(\infty, \infty) = 1$  in the equation above, we obtain

$$\varphi_\alpha^{\text{out}}(x) = S^{-1}\varphi_\alpha^{\text{in}}(x)S , \quad (6.32)$$

$$S = U(\infty, -\infty) . \quad (6.33)$$

The operator  $S$  is called the *S-matrix*.

## 6.2 Retarded Product Expansion of the Heisenberg Operators

As mentioned in the last section, field operators in the interaction picture can be quantized in just the same manner as in the free-field case. So, how can we carry out this quantization in the Heisenberg picture? To do this, we usually express the Heisenberg operator in terms of operators quantized in the interaction picture. Since this problem cannot generally be solved, a perturbation method is often used. There are two such methods: the first is a retarded product expansion and the second is the Yang–Feldman method [90]. Here, we will discuss the first of these methods.

In this section, the Heisenberg operator is denoted by  $\varphi_\alpha$ . We choose  $\varphi_\alpha^{\text{in}}$  as an operator in the interaction picture. Since the relation between them is given by (6.31), we have

$$\varphi_\alpha(x) = U(t, -\infty)^{-1} \varphi_\alpha^{\text{in}}(x) U(t, -\infty), \quad t = x_0. \quad (6.34)$$

We expand this representation perturbatively as a power series in the coupling constants. The interaction Hamiltonian in the interaction picture is

$$H(t) = \int_{x_0=t} d^3x \mathcal{H}_{\text{int}}(x). \quad (6.35)$$

We introduce the quantity

$$\mathcal{O}(t_0) = U(t, t_0)^{-1} \mathcal{O} U(t, t_0). \quad (6.36)$$

Therefore, from the TS equation in the single-time theory, we obtain

$$i \frac{\partial}{\partial t_0} \mathcal{O}(t_0) = [H(t_0), \mathcal{O}(t_0)]. \quad (6.37)$$

We convert this equation into the integral form

$$\mathcal{O}(t_0) = \mathcal{O} - i \int_{t_0}^t dt'_0 [\mathcal{O}(t'_0), H(t'_0)]. \quad (6.38)$$

Then iterating by inserting (6.38) back into  $\mathcal{O}(t'_0)$  in the integrand, we have

$$\begin{aligned} \mathcal{O}(t_0) &= \mathcal{O} - i \int_{t_0}^t dt'_0 [\mathcal{O}, H(t'_0)] + (-i)^2 \int_{t_0}^t dt'_0 \int_{t'_0}^t dt''_0 [[\mathcal{O}, H(t'')], H(t'_0)] + \dots \\ &= \mathcal{O} - i \int_{t_0}^t dt' [\mathcal{O}, H(t')] + (-i)^2 \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' [[\mathcal{O}, H(t')], H(t'')] + \dots \end{aligned} \quad (6.39)$$

We now make the replacement

$$\mathcal{O}(t_0) \rightarrow \varphi_\alpha(x), \quad \mathcal{O} \rightarrow \varphi_\alpha^{\text{in}}(x), \quad t_0 \rightarrow -\infty.$$

This yields the expansion equation

$$\begin{aligned} \varphi_\alpha(x) &= U(t, -\infty)^{-1} \varphi_\alpha^{\text{in}}(\mathbf{x}, t) U(t, -\infty) \\ &= \varphi_\alpha^{\text{in}}(\mathbf{x}, t) - i \int_{-\infty}^t dt' [\varphi_\alpha^{\text{in}}(\mathbf{x}, t), H(t')] \\ &\quad + (-i)^2 \int_{-\infty}^t dt' \int_{-\infty}^{t'} dt'' [[\varphi_\alpha^{\text{in}}(\mathbf{x}, t), H(t')], H(t'')] + \dots \end{aligned} \quad (6.40)$$

Next, we define the retarded product of operators:

$$R[\mathcal{O}(t) : A(t_1), \dots, A(t_n)] = (-i)^n [\dots [\mathcal{O}(t), A(t'_1)] \dots A(t'_n)],$$

where  $t > t'_1 > \dots > t'_n$  and  $(t'_1, \dots, t'_n)$  is a permutation of  $(t_1, \dots, t_n)$ . If  $t$  is smaller than any of the  $t_j$ , we define  $R = 0$ . From the definition,  $R$  is a symmetric function of  $t_1, \dots, t_n$ . Using  $R$ , the general term in (6.40) can be written as

$$\begin{aligned} &(-i)^n \int_{-\infty}^t dt_1 \int_{-\infty}^{t_1} dt_2 \dots \int_{-\infty}^{t_{n-1}} dt_n [\dots [\varphi_\alpha^{\text{in}}(\mathbf{x}, t), H(t_1)] \dots H(t_n)] \\ &= \int_{-\infty}^t dt_1 \int_{-\infty}^{t_1} dt_2 \dots \int_{-\infty}^{t_{n-1}} dt_n R[\varphi_\alpha^{\text{in}}(\mathbf{x}, t) : H(t_1) \dots H(t_n)] \\ &= \frac{1}{n!} \int_{-\infty}^t dt_1 \int_{-\infty}^{t_1} dt_2 \dots \int_{-\infty}^{t_{n-1}} dt_n R[\varphi_\alpha^{\text{in}}(\mathbf{x}, t) : H(t_1) \dots H(t_n)] \\ &= \frac{1}{n!} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \dots \int_{-\infty}^{\infty} dt_n R[\varphi_\alpha^{\text{in}}(\mathbf{x}, t) : H(t_1) \dots H(t_n)]. \end{aligned}$$

Combining this with (6.35), we rewrite the expansion of (6.40) as

$$\varphi_\alpha(x) = \varphi_\alpha^{\text{in}}(x) + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{\infty} d^4x_1 \dots \int_{-\infty}^{\infty} d^4x_n R \dots [\varphi_\alpha^{\text{in}}(x) : \mathcal{H}_{\text{int}}(x_1), \dots, \mathcal{H}_{\text{int}}(x_n) \dots]. \quad (6.41)$$

This is in fact a generalization of the result in Sect. 4.4.

### 6.3 Yang–Feldman Expansion of the Heisenberg Operators

Next, we consider a method for expanding the Heisenberg operator using the field equation. This method was originally introduced by Yang and Feldman [90]. The approach used in Sect. 4.4 is in fact based on this.

First, we separate the Lagrangian density into the free part and the interaction part. As mentioned in the previous section, this separation is not trivial:

$$\mathcal{L} = \mathcal{L}_f + \mathcal{L}_{\text{int}} . \quad (6.42)$$

Next, we write down the Euler–Lagrange equation, where we keep the separation above, i.e.,

$$\begin{aligned} [\mathcal{L}]_{\varphi_\alpha^\dagger} &= [\mathcal{L}_f]_{\varphi_\alpha^\dagger} + [\mathcal{L}_{\text{int}}]_{\varphi_\alpha^\dagger} \\ &= D_{\alpha\beta}(\partial)\varphi_\beta + [\mathcal{L}_{\text{int}}]_{\varphi_\alpha^\dagger} = 0 . \end{aligned} \quad (6.43)$$

In order to solve this equation, recall Fig. 4.2 in Sect. 4.4. We assume that interactions exist only in the finite space-time region  $\Omega$ , i.e., we assume that the coupling constant  $g$  is a function of  $x$ , with  $g \neq 0$  only in  $\Omega$  and  $g = 0$  outside  $\Omega$ . Therefore, in Fig. 4.2 we assume that  $g = 0$  for  $t > T_2$  or  $t < T_1$ . Thus, in these two regions the solution of (6.43) becomes

$$\varphi_\alpha = \begin{cases} \varphi_\alpha^{\text{in}}, & \text{for } t < T_1 \\ \varphi_\alpha^{\text{out}}, & \text{for } t > T_2 . \end{cases} \quad (6.44)$$

At the end, we take the limits  $T_1 \rightarrow -\infty$  and  $T_2 \rightarrow \infty$ . In addition, the two free fields are defined to be connected throughout space-time via the equation

$$D_{\alpha\beta}(\partial)\varphi_\beta^{\text{in}}(x) = D_{\alpha\beta}(\partial)\varphi_\beta^{\text{out}}(x) = 0 . \quad (6.45)$$

In general, we take

$$\varphi_\alpha(x) = \varphi_\alpha^{\text{in}}(x) + \varphi'_\alpha(x) . \quad (6.46)$$

Inserting this in (6.43) and using (6.45),

$$D_{\alpha\beta}(\partial)\varphi'_\beta(x) = -[\mathcal{L}_{\text{int}}]_{\varphi_\alpha^\dagger}(x) . \quad (6.47)$$

In order to integrate this equation, we introduce the retarded  $\Delta$ -function  $\Delta^{\text{R}}$  by (4.53):

$$\varphi'_\alpha(x) = \int_{\Omega} d^4x' \Delta_{\alpha\beta}^{\text{R}}(x - x') [\mathcal{L}_{\text{int}}]_{\varphi_\beta^\dagger}(x') . \quad (6.48)$$

Extending  $\Omega$  into the whole space-time and taking  $g(x) \rightarrow g$ ,

$$\varphi_\alpha(x) = \varphi_\alpha^{\text{in}}(x) + \int d^4x' \Delta_{\alpha\beta}^{\text{R}}(x-x') [\mathcal{L}_{\text{int}}]_{\varphi_\beta^\dagger}(x'). \quad (6.49)$$

This is a non-linear integral equation for  $\varphi_\alpha(x)$ . Similarly, using the advanced  $\Delta$ -function,

$$\varphi_\alpha(x) = \varphi_\alpha^{\text{out}}(x) + \int d^4x' \Delta_{\alpha\beta}^{\text{A}}(x-x') [\mathcal{L}_{\text{int}}]_{\varphi_\beta^\dagger}(x'). \quad (6.50)$$

The quantized field operator can be obtained by solving this integral equation. Taking the difference between (6.49) and (6.50),

$$\varphi_\alpha^{\text{out}}(x) = \varphi_\alpha^{\text{in}}(x) - \int d^4x' \Delta_{\alpha\beta}(x-x') [\mathcal{L}_{\text{int}}]_{\varphi_\beta^\dagger}(x'). \quad (6.51)$$

Alternatively, from (6.32), the  $S$ -matrix is determined by

$$S^{-1} \varphi_\alpha^{\text{in}}(x) S = \varphi_\alpha^{\text{in}}(x) - \int d^4x' \Delta_{\alpha\beta}(x-x') [\mathcal{L}_{\text{int}}]_{\varphi_\beta^\dagger}(x'). \quad (6.52)$$

The equation above is called the *Yang–Feldman equation*. The commutation relations are

$$[\varphi_\alpha^{\text{in}}(x), \varphi_\beta^{\dagger\text{in}}]_{\pm} = [\varphi_\alpha^{\text{out}}(x), \varphi_\beta^{\dagger\text{out}}(y)]_{\pm} = iC_{\alpha\beta}(\partial) \Delta(x-y). \quad (6.53)$$

The sign chosen above depends on the statistics of the field.

## 6.4 Examples of Interactions

So far we have discussed the general situation without specifying the form of interaction. We shall now consider some well known examples.

First, let us consider quantum electrodynamics (QED), which describes the interaction between the electromagnetic field and the electron field. A general theory of electromagnetic interactions was given by Dirac. When the field  $\psi$  corresponds to a particle with an electric charge  $e$ , we only have to make the following replacements in its free-field Lagrangian density:

$$\partial_\mu \psi \rightarrow (\partial_\mu - ieA_\mu) \psi, \quad \partial_\mu \bar{\psi} \rightarrow (\partial_\mu + ieA_\mu) \bar{\psi}, \quad (6.54)$$

where  $e$  is the electric charge of the quantum of the field  $\psi$  and  $-e$  is the electric charge of its antiparticle. We call the interaction induced by such a replacement

a *minimal electromagnetic coupling*. Applying the replacement above for the Lagrangian density of the Dirac field, the interaction part is given by

$$\mathcal{L}_{\text{int}} = j_{\mu} A_{\mu} , \quad (6.55)$$

where  $j_{\mu}$  is a four-current density given by

$$j_{\mu} = ie\bar{\psi}\gamma_{\mu}\psi , \quad \text{or} \quad \frac{1}{2}ie[\bar{\psi}, \gamma_{\mu}\psi] . \quad (6.56)$$

The meaning of the latter case is discussed in Sect. 3.9. We then write the Lagrangian density of the whole system as

$$\mathcal{L} = -\bar{\psi}[\gamma_{\mu}(\partial_{\mu} - ieA_{\mu}) + m]\psi - \frac{1}{4}F_{\mu\nu}F_{\mu\nu} - \frac{1}{2\alpha}(\partial_{\mu}A_{\mu})^2 . \quad (6.57)$$

QED is characterized by what is known as *gauge invariance*, i.e., making the following changes:

$$\begin{aligned} \text{gauge transformation of the first kind} \quad \psi &\rightarrow e^{ie\Lambda}\psi , \quad \bar{\psi} \rightarrow e^{-ie\Lambda}\bar{\psi} , \\ \text{gauge transformation of the second kind} \quad A_{\mu} &\rightarrow A_{\mu} + \partial_{\mu}\Lambda , \end{aligned} \quad (6.58)$$

the Lagrangian (6.57) is invariant, except for the last term. Note that  $\Lambda$  is an arbitrary function of the space-time coordinate.  $D_{\mu\nu}(\partial)$  derived from the free part of the Lagrangian density, i.e., the part that remains when  $e = 0$ , has no inverse, so quantization does not go through. For this reason, we add the last term in (6.57), which is called, a *gauge-fixing term*. With this procedure, although it is possible to quantize, the general gauge invariance is lost. There is only a restricted gauge invariance, i.e., the Lagrangian density (6.57) is invariant only under the condition

$$\square\Lambda = 0 . \quad (6.59)$$

In Chap. 3, we discussed the free Dirac particle. Here we mention a special feature of the electromagnetic field and its interactions. This concerns the magnetic moment of the electron. From the Lagrangian density (6.57), we obtain the following field equations:

$$[\gamma_{\mu}(\partial_{\mu} - ieA_{\mu}) + m]\psi = 0 , \quad \bar{\psi}[\gamma_{\mu}(\overleftarrow{\partial}_{\mu} + ieA_{\mu}) - m] = 0 . \quad (6.60)$$

$$\square A_{\mu} + \frac{1-\alpha}{\alpha}\partial_{\mu}\partial_{\nu}A_{\nu} = -j_{\mu} , \quad (6.61)$$



where  $\overleftarrow{\partial}_\mu$  in the second equation of (6.60) is an operator which differentiates  $\bar{\psi}$  on the left. From this, it is easy to derive the conservation of the current density:

$$\partial_\mu j_\mu = 0 . \quad (6.62)$$

Moreover, taking the divergence of (6.61), we have

$$\square \partial_\mu A_\mu = 0 . \quad (6.63)$$

Combining this with the Yang–Feldman formalism,

$$\partial_\mu A_\mu = \partial_\mu A_\mu^{\text{in}} = \partial_\mu A_\mu^{\text{out}} . \quad (6.64)$$

Since

$$S^{-1} \partial_\mu A_\mu^{\text{in}} S = \partial_\mu A_\mu^{\text{out}} , \quad (6.65)$$

equation (6.64) shows that the  $S$ -matrix and  $\partial_\mu A_\mu$  commute:

$$[S, \partial_\mu A_\mu] = 0 . \quad (6.66)$$

This is (5.120). Therefore, the  $S$ -matrix is an observable quantity.

Multiplying the first equation in (6.60) by the operator

$$\gamma_\lambda (\partial_\lambda - ieA_\lambda) - m ,$$

we obtain

$$[\gamma_\lambda \gamma_\mu (\partial_\lambda - ieA_\lambda) (\partial_\mu - ieA_\mu) - m^2] \psi = 0 . \quad (6.67)$$

Then we rewrite the product of  $\gamma$ -matrices in the form

$$\gamma_\lambda \gamma_\mu = \delta_{\lambda\mu} + i\sigma_{\lambda\mu} .$$

Thus, (6.67) can be written in the form

$$\left[ (\partial_\lambda - ieA_\lambda)^2 - m^2 + \frac{e}{2} \sigma_{\lambda\mu} F_{\lambda\mu} \right] \psi = 0 . \quad (6.68)$$

In the non-relativistic approximation, we can derive the Schrödinger equation. Assuming that the electric field is not a quantized field but a static external field, we replace the differential operators by  $E$  and  $\mathbf{p}$ :

$$(E - e\varphi)^2 \psi = \left[ (\mathbf{p} - e\mathbf{A})^2 + m^2 - \frac{e}{2} \sigma_{\lambda\mu} F_{\lambda\mu} \right] \psi . \quad (6.69)$$

Writing the energy eigenvalue of this system as  $E = m + W$  and introducing the following non-relativistic approximation

$$\left| \frac{W - e\varphi}{2m} \right| \ll 1, \quad (6.70)$$

equation (6.69) can be approximated by

$$W\psi = \left[ \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\varphi - \frac{e}{4m} \sigma_{\lambda\mu} F_{\lambda\mu} \right] \psi. \quad (6.71)$$

If  $W$  stands for the energy in the non-relativistic Schrödinger equation, we pick up only the two large components from the four components of  $\psi$  under this approximation, and the last term in (6.71) becomes

$$- \frac{e}{2m} \boldsymbol{\sigma} \mathbf{H}. \quad (6.72)$$

This means that the magnetic moment of the electron is  $-e/2m$ . This was an empirical fact in non-relativistic quantum mechanics, but in the Dirac theory it is proven from first principles, and this is one of the experimental reasons in favour of the theory.

For the next example, we consider a system involving nucleons and  $\pi$ -mesons. This interaction was first introduced by Yukawa [17]. In the light of the current quark model, it is viewed as a phenomenological theory rather than a fundamental one. According to Yukawa's theory, the force among nucleons, i.e., the nuclear force, is produced by the exchange of  $\pi$ -mesons between nucleons. Empirically, we know that the nuclear force does not depend on the electric charge of the nucleon in a given angular-momentum state. This is referred to as *charge independence*. This property yields a restriction on the interaction between nucleons and  $\pi$ -mesons. For instance, assume that two nucleons are in the  $^1S$ -state. Here the 1 on the upper left means that the total spin derived from the composition of the two nucleon spins is 0, so it belongs to the spin singlet.  $S$  indicates that the relativistic angular momentum is 0. In this state, the potential between two protons, or two neutrons, or a proton and a neutron satisfies the approximate equality

$$V_{pp}(^1S) = V_{nn}(^1S) = V_{np}(^1S). \quad (6.73)$$

On the other hand, due to Pauli's principle, the spin triplet state  $^3S$  exists only in the neutron-proton system. The charge independence reflects a kind of symmetry, and some kind of conserved quantity corresponds to it. This conserved quantity is called *isospin*, because its mathematical properties are similar to those of spin. In the following, we begin by introducing the isospin.

We consider the spin in a two-electron system. Denoting the states with spin-up and spin-down by  $\alpha$  and  $\beta$ , respectively, the wave functions of the spin triplet and the spin singlet are:

$$\text{triplet} \left\{ \begin{array}{l} \alpha(1)\alpha(2), \\ \frac{\alpha(1)\beta(2) + \beta(1)\alpha(2)}{\sqrt{2}}, \\ \beta(1)\beta(2), \end{array} \right. \quad \text{singlet} \frac{\alpha(1)\beta(2) - \beta(1)\alpha(2)}{\sqrt{2}}. \quad (6.74)$$

States in the same multiplet can be mapped to each other by spin rotations.

We consider the proton state  $p$  and the neutron state  $n$  to be two different states of the same nucleon. Then, by analogy with the spin case, isospin states in the two-nucleon system will belong to either the isospin triplet or the isospin singlet. To be more precise, we rewrite  $\alpha$  and  $\beta$  as  $p$  and  $n$ , respectively. The charge independence of the nuclear force means that members of the same multiplet have the same potential. For example, the spin singlet  $^1S$ -state corresponds to the isospin triplet, while  $^3S$ -state corresponds to the isospin singlet:

$$\text{triplet} \left\{ \begin{array}{l} p(1)p(2), \\ \frac{p(1)n(2) + n(1)p(2)}{\sqrt{2}}, \\ n(1)n(2), \end{array} \right. \quad \text{singlet} \frac{p(1)n(2) - n(1)p(2)}{\sqrt{2}}. \quad (6.75)$$

For spin, the Pauli matrices  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  are defined by

$$\begin{aligned} \sigma_x \alpha &= \beta, & \sigma_y \alpha &= i\beta, & \sigma_z \alpha &= \alpha, \\ \sigma_x \beta &= \alpha, & \sigma_y \beta &= -i\alpha, & \sigma_z \beta &= -\beta. \end{aligned} \quad (6.76)$$

For isospin, we also introduce three matrices  $\tau_1$ ,  $\tau_2$ , and  $\tau_3$ , defining them by the following equations:

$$\begin{aligned} \tau_1 p &= n, & \tau_2 p &= in, & \tau_3 p &= p, \\ \tau_1 n &= p, & \tau_2 n &= -ip, & \tau_3 n &= -n. \end{aligned} \quad (6.77)$$

As in the case of ordinary spin, we define the isospin of a system with  $A$  nucleons as

$$I_a = \frac{1}{2} \sum_{k=1}^A \tau_a^{(k)}, \quad \text{where } k \text{ is the number of nucleons.} \quad (6.78)$$

Therefore, as in the case of angular momentum the commutation relation among the three components is

$$[I_a, I_b] = i\epsilon_{abc}I_c . \quad (6.79)$$

Then the requirement that the potential  $V$  between two nucleons be independent of the charge can be expressed by

$$[I_a, V] = 0 . \quad (6.80)$$

In fact, this holds only approximately.

We can consider  $I_1$ ,  $I_2$ , and  $I_3$  as the three components of an abstract vector in isospin space, and we denote it by  $\hat{I}$ . Although this system is invariant under rotations in the isospin space, due to the Coulomb interaction and the small mass difference between the proton and the neutron, this invariance is approximate. The isospin of the nucleon is  $1/2$ , while that of the  $\pi$ -meson with three isospin states  $\pi^+$ ,  $\pi^0$ , and  $\pi^-$  is 1. Hence, if we denote the eigenvalue of  $\hat{I}^2$  by  $I'(I' + 1)$ , it turns out that this multiplet is degenerate with degree  $2I' + 1$  as in the case of the angular momentum. However, as in the case including  $\pi$ -mesons, we have to generalize (6.78).

The spin of the  $\pi$ -meson will be shown later to be zero. According to (3.70), for a three-component real scalar field, the Lagrangian density is

$$\mathcal{L}_\pi = - \sum_{\alpha=1}^3 \frac{1}{2} [(\partial_\lambda \varphi_\alpha)^2 + \mu^2 \varphi_\alpha^2] , \quad (6.81)$$

where  $\varphi_1$ ,  $\varphi_2$ , and  $\varphi_3$  are the three components of the vector  $\hat{\varphi}$  in the isospin space. By (3.69),  $\varphi_1$  and  $\varphi_2$  correspond to the real and imaginary parts of  $\varphi$  and  $\varphi^\dagger$ , respectively.  $\varphi$  is the operator which annihilates  $\pi^+$  and creates  $\pi^-$ , while  $\varphi^\dagger$  is its conjugate operator.  $\varphi_3$  is the operator which creates and annihilates  $\pi^0$ . Since the nucleon belongs to the two-dimensional irreducible representation of the rotation group in the isospin space, it becomes mathematically equivalent to the spinor for ordinary spatial rotations, and is called an *isospinor*. The  $\pi$ -meson is an *isovector*. The Lagrangian density of the nucleon is

$$\mathcal{L}_N = - \sum_{\alpha=1}^2 \bar{\psi}_\alpha (\gamma_\mu \partial_\mu + M) \psi_\alpha , \quad (6.82)$$

where  $\psi_1$  stands for the proton component  $\psi_p$  and  $\psi_2$  for the neutron component  $\psi_n$ . Therefore, referring to (6.77), the  $\tau$ -matrices have the same representation as the Pauli matrices:

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \quad (6.83)$$

Since  $\psi$  is an isospinor,  $\bar{\psi}\tau_\alpha\psi$  becomes an isovector in the isospin space. If we take the scalar product of the isovector and the meson field  $\hat{\phi}$ , the result is an isoscalar which is invariant under rotations in the isospin space. In order to ensure the charge independence of the nuclear force, the interaction in the nucleon–meson system must be an isoscalar. A similar idea was first introduced by Kemmer [27]. For the interaction, we thus adopt a term proportional to

$$\bar{\psi}\mathcal{O}\hat{\tau}\psi\cdot\hat{\phi}. \quad (6.84)$$

Here  $\hat{\tau}\cdot\hat{\phi}$  stands for the scalar product of two isovectors.  $\mathcal{O}$  is a Dirac matrix, and since the  $\pi$ -meson is in fact a pseudo-scalar, we choose it to be  $\mathcal{O} = i\gamma_5$ . Considering  $\hat{\omega}$  to be a vector generating an infinitesimal rotation, (6.84) is invariant under the following infinitesimal transformation in the isospin space:

$$\psi \rightarrow \psi + \frac{i}{2}(\hat{\omega}\cdot\hat{\tau})\psi, \quad \hat{\phi} \rightarrow \hat{\phi} - \hat{\omega}\times\hat{\phi}, \quad (6.85)$$

where  $\times$  denotes the vector product in the isospin space.

The property of charge independence has been observed empirically in many situations. Here we give an example. From the fact that the deuteron  $d$  has isospin 0, we can derive the following relation between two differential cross-sections:

$$\frac{d\sigma(p+p \rightarrow d+\pi^+)}{d\Omega} = 2\frac{d\sigma(n+p \rightarrow d+\pi^0)}{d\Omega}. \quad (6.86)$$

This equation is consistent with experimental results and thus supports the hypothesis of charge independence.

The spin and the parity of the  $\pi$ -meson or pion has been determined as follows. According to the principle of detailed balance, assuming that the mechanical law that governs the nucleon–meson system is time-reversible and comparing with the reverse process,

$$\pi^+ + d \rightleftharpoons p + p, \quad (6.87)$$

the following relation holds:

$$\frac{d\sigma(\rightarrow)}{d\Omega} \bigg/ \frac{d\sigma(\leftarrow)}{d\Omega} = \frac{4}{3(2S+1)} \frac{p^2}{q^2}, \quad (6.88)$$

where  $p$  and  $q$  are the momenta of the proton and the pion in the centre-of-mass system and  $S$  is the spin of the pion. We assume that in any process the initial particle is not polarized. In this equation, it has been taken into account that the spins of the proton and deuteron are  $1/2$  and  $1$ , respectively. Comparing (6.88) with experiment, we find

$$S = 0, \quad (6.89)$$

i.e., the spin of the pion is zero. Moreover, rewriting (6.88) for the total cross-section, the Pauli principle implies a difference of a factor of two:

$$\frac{\sigma(\rightarrow)}{\sigma(\leftarrow)} = \frac{2}{3(2S+1)} \frac{p^2}{q^2}. \quad (6.90)$$

The spin of  $\pi^-$  is also zero because it is the antiparticle of  $\pi^+$ . From the charge invariance, it is easy to understand that the spin of  $\pi^0$  must be zero.  $\pi^0$  decays mainly into two photons:

$$\pi^0 \rightarrow 2\gamma. \quad (6.91)$$

Later, we will use this to prove that the spin of  $\pi^0$  cannot be 1.

The low-energy  $\pi^-$  loses its energy via electromagnetic interactions when traversing a material, and falls to the  $l = 0$  Bohr orbit of the nucleon in a time much shorter than its lifetime. It is subsequently absorbed by the nucleon. In particular, the following absorption process occurs for the deuteron:

$$\pi^- + d \rightarrow n + n. \quad (6.92)$$

Let us compare the angular momenta on each side. Since the spin of the pion is zero and the orbital angular momentum between the pion and the deuteron is zero, the total angular momentum on the left-hand side can be determined by the angular momentum state of the deuteron. The deuteron is a bound state of the proton and the neutron with angular momentum state

$${}^3S_1 + {}^3D_1. \quad (6.93)$$

As mentioned before, the number 3 on the top left indicates that the two spins comprise a parallel triplet term, while  $S$  and  $D$  indicate that the orbital angular momenta are zero and two, respectively. Moreover, the number 1 on the bottom right indicates that the total angular momentum is unity. Since there are two neutrons, which have Fermi statistics, on the right-hand side of (6.92), their wave function must be anti-symmetric. Consequently, the total wave function of the deuteron must be anti-symmetric as well. Thus, when the spin wave functions are symmetric triplets, the angular momentum must be odd, and when the spin wave function is an anti-symmetric singlet, the angular momentum must be even. Writing down the compositions of spins and orbital angular momenta on the lower right, all the possible states are

$${}^1S_0, \quad {}^3P_{0,1,2}, \quad {}^1D_2, \quad {}^3F_{2,3,4}, \quad \dots. \quad (6.94)$$

Choosing the one whose orbital angular momentum is unity, the only angular-momentum state which corresponds to (6.92) is

$${}^3S_1 + {}^3D_1 \rightarrow {}^3P_1 . \quad (6.95)$$

Since the parity of the orbital part is altered from even to odd, it turns out that this change originates in the intrinsic parity of the pion. That is, the (intrinsic) parity of the pion is odd. Here we have assumed that the intrinsic parity of the proton and the neutron is even. Therefore, the pion is a spin-0 and parity-odd particle. In general, this is expressed as  $0^-$ . Equivalently, we say that the pion is a pseudo-scalar.

The interactions in the nucleon–meson system which preserve both parity and isospin can be written as follows:

$$\mathcal{L}_{\text{int}} = -iG\bar{\psi}\gamma_5\hat{\tau}\psi \cdot \hat{\varphi} - \frac{\lambda}{4}(\hat{\varphi}^2)^2 , \quad (6.96)$$

where the first term is the so-called *Yukawa interaction*, and the second term expresses the interaction among mesons, a necessary term for the renormalization discussed later. The real (pseudo) scalar representation has been used here. If the complex representation is used,

$$\bar{\psi}\gamma_5\hat{\tau}\psi \cdot \hat{\varphi} = \sqrt{2}(\bar{\psi}_n\gamma_5\psi_p\varphi^\dagger + \bar{\psi}_p\gamma_5\psi_n\varphi) + (\bar{\psi}_p\gamma_5\psi_p - \bar{\psi}_n\gamma_5\psi_n)\varphi_3 .$$

The full Lagrangian density is therefore

$$\mathcal{L} = -\bar{\psi}(\gamma_\mu\partial_\mu + M)\psi - \frac{1}{2}[(\partial_\lambda\hat{\varphi})^2 + \mu^2\hat{\varphi}^2] - iG\bar{\psi}\gamma_5\hat{\tau}\psi \cdot \hat{\varphi} - \frac{\lambda}{4}(\hat{\varphi}^2)^2 . \quad (6.97)$$

We end with a remark about the nuclear force, which was the key factor for the development of Yukawa's theory. Since it is difficult to discuss this issue from a completely general point of view, we consider the neutral scalar meson. Assuming that the nucleon is much heavier than the meson, the nucleon can be considered to be at rest, so it can be replaced by an external field. This is analogous to replacing the effects of the proton by a Coulomb potential when we treat the proton as being at rest in the problem of the hydrogen atom. This is called a *static approximation*. In this approximation, for the nucleon–meson interaction, i.e.,

$$H_{\text{int}} = \int d^3x g\bar{\psi}(\mathbf{x})\psi(\mathbf{x})\varphi(\mathbf{x}) , \quad (6.98)$$

we consider  $\bar{\psi}(\mathbf{x})\psi(\mathbf{x})$  to be the probability density of the nucleon, and replace it as follows:

$$\bar{\psi}(\mathbf{x})\psi(\mathbf{x}) \rightarrow \sum_{j=1}^N \delta^3(\mathbf{x} - \mathbf{x}_j) , \quad (6.99)$$

where  $\mathbf{x}_j$  is the position coordinate of the  $j$ th nucleon. Hence, the nucleon is not quantized here. Considering a two-body problem, choosing  $N = 2$ , and substituting (6.99) into (6.98), we have

$$H = H_f + g[\varphi(\mathbf{x}_1) + \varphi(\mathbf{x}_2)] \equiv H_f + H' , \quad (6.100)$$

where  $H_f$  is the Hamiltonian of the free meson and we are working in the Schrödinger picture. We now determine the eigenvalues of this Hamiltonian. In particular, we would like to obtain the energy eigenvalue of the ground state. The equation is

$$H\Psi = E\Psi . \quad (6.101)$$

To solve this problem, we use the Fourier representation

$$H_f = \sum_{\mathbf{q}} q_0 c^\dagger(\mathbf{q})c(\mathbf{q}) , \quad (6.102)$$

$$\varphi(\mathbf{x}) = \sum_{\mathbf{q}} \frac{1}{\sqrt{2q_0V}} [c(\mathbf{q})e^{i\mathbf{q}\mathbf{x}} + c^\dagger(\mathbf{q})e^{-i\mathbf{q}\mathbf{x}}] . \quad (6.103)$$

$H_f$  is already diagonal with respect to the particle number, but the interaction part is not. We thus diagonalize  $H$  by a suitable unitary transformation. To do this, we take

$$\Psi = U\Phi , \quad U \equiv \exp(iS) , \quad (6.104)$$

where  $U$  is unitary and  $S$  is Hermitian. Therefore,

$$U^{-1}HU\Phi = E\Phi , \quad (6.105)$$

$$\begin{aligned} U^{-1}HU &= U^{-1}(H_f + H')U \\ &= H_f + i[H_f, S] + \frac{i^2}{2}[[H_f, S], S] + \dots \\ &\quad + H' + i[H', S] + \dots \\ &\quad \vdots \quad \quad \quad \vdots \\ &\quad g^0 \quad g^1 \quad \quad g^2 \end{aligned} \quad (6.106)$$



where, assuming that  $S$  is of order of  $g$ , we have aligned terms with the same power of  $g$  vertically. We then define  $S$  by

$$H' + i[H_f, S] = 0. \quad (6.107)$$

Therefore, (6.106) can be assembled into the following form:

$$U^{-1}HU = H_f + \sum_{n=1}^{\infty} i^n \left[ \frac{1}{n!} - \frac{1}{(n+1)!} \right] \overbrace{[\dots [H', S] \dots S]}^n \text{commutators}. \quad (6.108)$$

Since the second term begins with terms of order  $g^2$ , it turns out that  $H$  has been diagonalized up to order  $g^2$ . In the case of the neutral meson, since the term with  $n = 1$  is a c-number, any term with more than two commutators vanishes. To solve (6.107), we use

$$[H_f, c(\mathbf{q})] = -q_0 c(\mathbf{q}), \quad [H_f, c^\dagger(\mathbf{q})] = q_0 c^\dagger(\mathbf{q}), \quad (6.109)$$

to show that  $S$  is given by

$$S = -ig \sum_{\mathbf{q}} \frac{1}{\sqrt{2q_0V}} \frac{1}{q_0} \left[ c(\mathbf{q})(e^{i\mathbf{q}x_1} + e^{i\mathbf{q}x_2}) - c^\dagger(\mathbf{q})(e^{-i\mathbf{q}x_1} + e^{-i\mathbf{q}x_2}) \right]. \quad (6.110)$$

Thus,

$$\begin{aligned} U^{-1}HU &= H_f + \frac{i}{2}[H', S] \\ &= H_f - g^2 \sum_{\mathbf{q}} \frac{1}{2q_0^2V} (e^{i\mathbf{q}x_1} + e^{i\mathbf{q}x_2})(e^{-i\mathbf{q}x_1} + e^{-i\mathbf{q}x_2}). \end{aligned}$$

Taking the limit  $V \rightarrow \infty$  and making the replacement (4.2),

$$U^{-1}HU = H_f - \frac{g^2}{(2\pi)^3} \int \frac{d^3q}{q_0^2} [1 + \cos \mathbf{q}(\mathbf{x}_1 - \mathbf{x}_2)]. \quad (6.111)$$

The integral splits into the divergent self-energy  $\delta M$  and the potential  $\mathcal{V}(r)$ :

$$\delta M = -\frac{g^2}{(2\pi)^3} \int \frac{d^3q}{2q_0^2}, \quad (6.112)$$

$$\mathcal{V}(r) = -\frac{g^2}{(2\pi)^3} \int \frac{d^3q}{q_0^2} \cos \mathbf{q}(\mathbf{x}_1 - \mathbf{x}_2) = -\frac{g^2}{4\pi} \frac{e^{-mr}}{r}, \quad r = |\mathbf{x}_1 - \mathbf{x}_2|. \quad (6.113)$$

Therefore, the lowest eigenvalue  $E$  of  $H$  can be written

$$E = 2\delta M + \mathcal{V}(r). \quad (6.114)$$

The function  $\mathcal{V}(r)$  is called the *Yukawa potential*.

# Chapter 7

## Symmetries and Conservation Laws



In this chapter, we study invariance under different types of continuous transformation in a given system, showing that such symmetries are closely related to the existence of conservation laws. The key to this is Noether's theorem. We begin by formulating this idea in a point-particle system, and then extend it to the field case.

### 7.1 Noether's Theorem for Point-Particle Systems

We define an action integral  $I$  starting from the Lagrangian  $L(q_j(t), \dot{q}_j(t))$  for a given system:

$$I = \int_{t_1}^{t_2} dt L(q_j(t), \dot{q}_j(t)) . \tag{7.1}$$

We consider the infinitesimal variations

$$t \rightarrow t' = t + \delta t , \quad q_j(t) \rightarrow q'_j(t') = q_j(t) + \delta q_j . \tag{7.2}$$

Regarding the dynamical variables  $q_j$ , we consider two kinds of transformation. The first does not involve any change in the time variable, while the second does:

$$\delta^* q_j(t) = q'_j(t) - q_j(t) , \quad \delta q_j(t) = q'_j(t') - q_j(t) . \tag{7.3}$$

The variation  $\delta^*$  thus commutes with the time derivative and we have the following relation between these two kinds of variation:

$$\delta^* q_j(t) = \delta q_j(t) - \delta t \dot{q}_j(t) . \tag{7.4}$$

We now consider a variation of the action integral corresponding to (7.2):

$$I' = \int_{t'_1}^{t'_2} dt' L(q'_j(t'), \dot{q}'_j(t')) . \quad (7.5)$$

The change in the  $t$ -derivative can be found from

$$\frac{d}{dt'} = \frac{dt}{dt'} \frac{d}{dt} = \left(1 - \frac{d}{dt} \delta t\right) \frac{d}{dt} ,$$

and

$$dt' = \frac{dt'}{dt} dt = \left(1 + \frac{d}{dt} \delta t\right) dt . \quad (7.6)$$

Transforming the integral variable  $t'$  of  $I'$  back to  $t$ , we find the difference from  $I$  to be

$$\begin{aligned} I' - I &= \int_{t_1}^{t_2} dt \left(1 + \frac{d}{dt} \delta t\right) L' - \int_{t_1}^{t_2} dt L \\ &= \int_{t_1}^{t_2} dt \left(L' - L + \frac{d}{dt} \delta t \cdot L\right) . \end{aligned} \quad (7.7)$$

The integrand can be rewritten

$$L' - L + \frac{d}{dt} \delta t \cdot L = \sum_j \left( \frac{\partial L}{\partial q_j} \delta q_j + \frac{\partial L}{\partial \dot{q}_j} \delta \dot{q}_j \right) + \frac{d}{dt} (\delta t L) - \delta t \frac{d}{dt} L .$$

Using (7.4) and taking into account the commutativity of  $\delta^*$  and the time derivative, the left-hand side becomes

$$\sum_j \left[ \frac{\partial L}{\partial q_j} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) \right] \delta^* q_j + \frac{d}{dt} \left( \sum_j \frac{\partial L}{\partial \dot{q}_j} \delta^* q_j + \delta t L \right) . \quad (7.8)$$

Although the first term must vanish, according to the Euler–Lagrange equation, we keep it for the moment:

$$I' - I = \left[ \sum_j \frac{\partial L}{\partial \dot{q}_j} \delta^* q_j + \delta t L \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} dt \sum_j \left[ \frac{\partial L}{\partial q_j} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) \right] \delta^* q_j . \quad (7.9)$$

If we now use the Euler–Lagrange equation, we obtain

$$I' - I = Q(t_2) - Q(t_1) , \quad (7.10)$$

where

$$Q = \sum_j \frac{\partial L}{\partial \dot{q}_j} \delta^* q_j + \delta t L = \sum_j p_j \delta q_j - H \delta t . \quad (7.11)$$

Therefore, if  $I$  is invariant under the infinitesimal transformation (7.2), the equation of motion implies the following conservation law:

$$\frac{d}{dt} Q(t) = 0 . \quad (7.12)$$

This correspondence between the invariance of the action integral under the infinitesimal transformation and the conservation law is called *Noether's theorem*. For this theorem within a point-particle system, the reader is referred to *Introduction to Analytical Mechanics for Learning Quantum Mechanics*, written in Japanese by Yasushi Takahashi [91].

### Time-like Parallel Translation

If we consider

$$\delta t = \epsilon , \quad \delta q_j(t) = 0 , \quad (7.13)$$

where  $\epsilon$  is an infinitesimal parameter independent of  $t$ , the action integral is invariant in the case where  $L$  has no explicit time dependence. Then,

$$Q = -\epsilon H , \quad (7.14)$$

whence  $H$  is conserved, i.e., we obtain the conservation law of energy.

### Spatial Parallel Translation

If we consider

$$\delta t = 0 , \quad \delta \mathbf{x}_j(t) = \boldsymbol{\epsilon} , \quad (7.15)$$

the action integral defined by the Lagrangian for a point-particle system interacting with a two-body central force, viz.,

$$L = \sum_j \frac{1}{2} m_j \dot{\mathbf{x}}_j^2 - \sum_{i>j} V(|\mathbf{x}_i - \mathbf{x}_j|) , \quad (7.16)$$

is invariant. The corresponding conserved quantity is

$$Q = \epsilon \sum_j \mathbf{p}_j = \epsilon \mathbf{P} , \quad \mathbf{p}_j = m_j \dot{\mathbf{x}}_j . \quad (7.17)$$

We thus obtain the conservation of momentum  $\mathbf{P}$ .

## 7.2 Noether's Theorem in Field Theory

It is easy to extend Noether's theorem from point-particle systems to fields. We begin with the action integral

$$I = \int_{\Omega} d^4x \mathcal{L}[\varphi_{\alpha}(x), \varphi_{\alpha,\mu}(x)] , \quad (7.18)$$

where  $\varphi(x)_{\alpha}$  is a generic notation for fields in various representations of the Poincaré algebra, and the subscript  $\alpha$  distinguishes between different fields, but also between the fields and their conjugates. We introduce the infinitesimal variations

$$x_{\mu} \rightarrow x'_{\mu} + \delta x_{\mu} , \quad \varphi_{\alpha}(x) \rightarrow \varphi'_{\alpha}(x') = \varphi_{\alpha}(x) + \delta\varphi_{\alpha}(x) , \quad (7.19)$$

where  $\varphi \rightarrow \varphi'$  stands for the  $x$ -dependence of  $\varphi$  or the variation of the functional form. As in the previous section, we introduce two types of variation:

$$\delta^* \varphi_{\alpha}(x) = \varphi'_{\alpha}(x) - \varphi_{\alpha}(x) , \quad \delta\varphi_{\alpha}(x) = \varphi'_{\alpha}(x') - \varphi_{\alpha}(x) , \quad (7.20)$$

where  $\delta^*$  is the variation used to derive the Euler–Lagrange equation, which was simply written as  $\delta$  in Chap. 3.  $\delta^*$  commutes with the derivative. Therefore, we can consider the variation in Chap. 3 to be  $\delta$  under the restriction  $\delta x_{\mu} = 0$ . The relation between the two variations is

$$\delta^* \varphi_{\alpha}(x) = \delta\varphi_{\alpha}(x) - \varphi_{\alpha,\mu}(x) \delta x_{\mu} . \quad (7.21)$$

Due to the commutativity of  $\delta^*$  with the derivative,

$$\delta\varphi_{\alpha,\mu}(x) = \partial_{\mu} \delta\varphi_{\alpha}(x) - \varphi_{\alpha,\nu}(x) \partial_{\mu} \delta x_{\nu} . \quad (7.22)$$

We can now calculate  $\delta I$ :

$$\delta I = \int_{\Omega'} d^4x' \mathcal{L}[\varphi'_{\alpha}(x'), \varphi'_{\alpha,\mu}(x')] - \int_{\Omega} d^4x \mathcal{L}[\varphi_{\alpha}(x), \varphi_{\alpha,\mu}(x)] . \quad (7.23)$$

In order to make the integration regions coincide, we change the variable  $x'$  back to  $x$  in the first integral. Using

$$d^4 x' = \frac{\partial(x'_0, x'_1, x'_2, x'_3)}{\partial(x_0, x_1, x_2, x_3)} d^4 x, \quad (7.24)$$

we obtain

$$\delta I = \int_{\Omega} d^4 x \left\{ \mathcal{L} \left[ \varphi_{\alpha}(x) + \delta\varphi_{\alpha}(x), \varphi_{\alpha,\mu}(x) + \delta\varphi_{\alpha,\mu}(x) \right] \frac{\partial(x'_0, x'_1, x'_2, x'_3)}{\partial(x_0, x_1, x_2, x_3)} - \mathcal{L} \left[ \varphi_{\alpha}(x), \varphi_{\alpha,\mu}(x) \right] \right\}.$$

For infinitesimal  $\delta x_{\mu}$ , we have

$$\frac{\partial(x'_0, x'_1, x'_2, x'_3)}{\partial(x_0, x_1, x_2, x_3)} = 1 + \partial_{\nu}(\delta x_{\nu}). \quad (7.25)$$

Using this and the summation convention for repeated indices,

$$\delta I = \int_{\Omega} d^4 x \left[ \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha}} \delta\varphi_{\alpha} + \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} \delta\varphi_{\alpha,\mu} + \mathcal{L} \partial_{\nu}(\delta x_{\nu}) \right]. \quad (7.26)$$

Therefore, if the action integral  $I$  is invariant under any choice of  $\Omega$ ,

$$\delta \mathcal{L} + \mathcal{L} \partial_{\nu}(\delta x_{\nu}) = 0, \quad (7.27)$$

where

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha}} \delta\varphi_{\alpha} + \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} \delta\varphi_{\alpha,\mu}. \quad (7.28)$$

Rewriting this in terms of  $\delta^*$  rather than  $\delta$ ,

$$[\mathcal{L}]_{\varphi_{\alpha}} \delta^* \varphi_{\alpha} + \frac{\partial}{\partial x_{\mu}} \left( \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} \delta^* \varphi_{\alpha} + \mathcal{L} \delta x_{\mu} \right) = 0. \quad (7.29)$$

Next, we introduce the canonical energy–momentum tensor

$$\mathcal{T}_{\mu\nu} = -\frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} \varphi_{\alpha,\nu} + \delta_{\mu\nu} \mathcal{L}. \quad (7.30)$$

We write the second term in (7.29) in terms of  $\delta$  once again:

$$[\mathcal{L}]_{\varphi_\alpha} \delta^* \varphi_\alpha + \frac{\partial}{\partial x_\mu} \left( \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} \delta \varphi_\alpha + \mathcal{L}_{\mu\nu} \delta x_\nu \right) = 0. \quad (7.31)$$

This expresses the invariance of the action integral. Using the field equation, we can drop the first term, and this yields the conservation law

$$\frac{\partial}{\partial x_\mu} \left( \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} \delta \varphi_\alpha + \mathcal{T}_{\mu\nu} \delta x_\nu \right) = 0. \quad (7.32)$$

The derivation above is basically the same as the one for a point-particle system. In field theory, Noether's theorem expresses the relationship between invariance and the above conservation law. The conserved four-dimensional current in brackets is called the *Noether current*.

### 7.3 Applications of Noether's Theorem

#### Space-Time Translations

If there is no external field, the Lagrangian density will not depend explicitly on any coordinate. This means that, if the Lagrangian density depends on coordinates only through fields, then the action integral will be invariant under the following transformations:

$$\delta x_\mu = \epsilon_\mu, \quad \delta \varphi_\alpha = 0. \quad (7.33)$$

Inserting (7.33) into (7.32), the conservation law is in this case

$$\partial_\mu (\mathcal{T}_{\mu\nu} \epsilon_\nu) = 0, \quad \text{or} \quad \partial_\mu \mathcal{T}_{\mu\nu} = 0. \quad (7.34)$$

This is because  $\epsilon_\nu$  is an arbitrary infinitesimal parameter. From this, we can obtain the energy–momentum vector as a conserved quantity:

$$P_\nu = \int d^3x \mathcal{T}_{0\nu}(x). \quad (7.35)$$

It is easy to check that this definition coincides with the one given in Sect. 3.2. If the Lagrangian density includes an external field,

$$\partial_\mu \mathcal{T}_{\mu\nu} = \frac{\partial \mathcal{L}}{\partial (x_\nu)}, \quad (7.36)$$

where the derivative with respect to  $(x_\nu)$  picks up only the coordinate dependence introduced through the external field, and involves no derivative of the dynamical



fields under investigation. We can in fact generalize (3.29). If  $\mathcal{O}(x)$  is written in terms of fields and does not depend explicitly on any coordinate,

$$\frac{\partial}{\partial x_\mu} \mathcal{O}[\varphi_\alpha(x)] = i[\mathcal{O}[\varphi_\alpha(x)], P_\mu]. \quad (7.37)$$

On the other hand, if  $\mathcal{O}(x)$  depends not only on the field  $\varphi_\alpha(x)$ , but also explicitly on some coordinate, i.e., if  $\mathcal{O}$  can be written as  $\mathcal{O}[\varphi_\alpha(x), (x_\mu)]$ , then (7.37) changes to

$$\frac{\partial}{\partial x_\mu} \mathcal{O}(x) = i[\mathcal{O}(x), P_\mu] + \frac{\partial}{\partial (x_\mu)} \mathcal{O}(x). \quad (7.38)$$

This equation will be used later.

### Phase Transformations

Next, we consider a phase transformation. Assuming that  $\varphi_\alpha$  is a complex field, the Lagrangian density also includes the complex conjugate  $\varphi_\alpha^*$  of  $\varphi_\alpha$ . Then writing the complex conjugate field explicitly, we consider the Lagrangian density

$$\mathcal{L} = \mathcal{L}[\varphi_\alpha(x), \varphi_\alpha^*(x), \varphi_{\alpha,\nu}(x), \varphi_{\alpha,\mu}^*(x)]. \quad (7.39)$$

We assume that this  $\mathcal{L}$  is invariant under the infinitesimal phase transformation

$$\begin{aligned} \varphi_\alpha &\rightarrow e^{iq_\alpha\lambda} \varphi_\alpha = \varphi_\alpha + iq_\alpha\lambda\varphi_\alpha, \\ \varphi_\alpha^* &\rightarrow e^{-iq_\alpha\lambda} \varphi_\alpha^* = \varphi_\alpha^* - iq_\alpha\lambda\varphi_\alpha^*, \end{aligned} \quad (7.40)$$

that is,

$$\delta\varphi_\alpha = iq_\alpha\lambda\varphi_\alpha, \quad \delta\varphi_\alpha^* = -iq_\alpha\lambda\varphi_\alpha^*. \quad (7.41)$$

Therefore, the Noether current is

$$\lambda \sum_\alpha iq_\alpha \left( \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} \varphi_\alpha - \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}^*} \varphi_\alpha^* \right). \quad (7.42)$$

When  $q_\alpha$  is an electric charge of the quantum of the field  $\varphi_\alpha$ , the conserved current is in fact

$$j_\mu = \sum_\alpha (-iq_\alpha) \left( \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} \varphi_\alpha - \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}^*} \varphi_\alpha^* \right). \quad (7.43)$$

This yields the conservation law.

### Isospin Transformations

Next, we consider an infinitesimal transformation in isospin space. If when we introduce an infinitesimal rotation vector  $\hat{\omega}$  the action integral is invariant under the infinitesimal rotation of the field quantity given by (6.85), then, in terms of the nucleon field and the meson field, Noether's current can be written in the form

$$-\hat{\omega} \cdot \left[ \frac{i}{2} \bar{\psi} \gamma_\mu \hat{\tau} \psi - \hat{\phi} \times \hat{\phi}_{,\mu} \right]. \quad (7.44)$$

Therefore, the conserved isospin current is

$$\hat{j}_\mu = \frac{i}{2} \bar{\psi} \gamma_\mu \hat{\tau} \psi - \hat{\phi} \times \partial_\mu \hat{\phi}. \quad (7.45)$$

The isospin  $\hat{L}$  is then given by

$$\hat{I} = \int d^3x \left[ \frac{1}{2} \psi^\dagger \hat{\tau} \psi + \hat{\phi} \times \hat{\pi} \right], \quad \hat{\pi} = \dot{\hat{\phi}}. \quad (7.46)$$

Various conserved quantities can be derived from the Noether current in this way.

## 7.4 Poincaré Invariance

Combining the previously mentioned parallel translations in space-time and the Lorentz transformations, we obtain the group of Poincaré transformations. Here we first consider the Lorentz transformation. Using an infinitesimal parameter  $\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}$ , the infinitesimal Lorentz transformation can be expressed as

$$\delta x_\mu = \epsilon_{\mu\nu} x_\nu, \quad \delta \varphi_\alpha(x) = \frac{1}{2} \epsilon_{\mu\nu} (T_{\mu\nu})_{\alpha\beta} \varphi_\beta(x). \quad (7.47)$$

This is a generalization of (3.119).  $T_{\mu\nu}$  is a representation of the generator of the Lorentz transformation. We have already encountered the following examples:

$$\text{Spin 0} \quad T_{\mu\nu} = 0, \quad (7.48a)$$

$$\text{Spin 1/2} \quad T_{\mu\nu} = \frac{1}{4} (\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu), \quad (7.48b)$$

$$\text{Spin 1} \quad T_{\mu\nu} = a_{\mu\nu}. \quad (7.48c)$$

However, since the matrix  $a_{\mu\nu}$  in the spin-1 case can be chosen to be the same as the matrix of the coordinate transformation, its components can be determined from

the identity

$$\delta x_\mu = \epsilon_{\mu\nu} x_\nu = \frac{1}{2} \epsilon_{\rho\sigma} (a_{\rho\sigma})_{\mu\nu} x_\nu .$$

Its specific form is then

$$(a_{\rho\sigma})_{\mu\nu} = \delta_{\mu\rho} \delta_{\nu\sigma} - \delta_{\mu\sigma} \delta_{\nu\rho} . \quad (7.49)$$

$T_{\mu\nu}$  given by (7.48a)–(7.48c) satisfies the commutation relation

$$[T_{\rho\sigma}, T_{\lambda\tau}] = \delta_{\sigma\lambda} T_{\rho\tau} - \delta_{\sigma\tau} T_{\rho\lambda} + \delta_{\rho\lambda} T_{\sigma\tau} - \delta_{\rho\tau} T_{\sigma\lambda} . \quad (7.50)$$

The Noether current corresponding to the above transformation is

$$\mathcal{M}_{\mu\rho\sigma} = x_\rho \mathcal{T}_{\mu\sigma} - x_\sigma \mathcal{T}_{\mu\rho} - \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} (T_{\rho\sigma})_\beta . \quad (7.51)$$

The conservation law is

$$\partial_\mu \mathcal{M}_{\mu\rho\sigma} = 0 , \quad (7.52)$$

where the generator  $M_{\rho\sigma}$  of the Lorentz transformation is defined by

$$M_{\rho\sigma} = \int d^3x \mathcal{M}_{0\rho\sigma} . \quad (7.53)$$

The same commutation relations hold among the  $M_{\rho\sigma}$  as among the  $-iT_{\rho\sigma}$ , i.e.,

$$[M_{\rho\sigma}, M_{\lambda\tau}] = -i(\delta_{\sigma\lambda} M_{\rho\tau} - \delta_{\sigma\tau} M_{\rho\lambda} + \delta_{\rho\lambda} M_{\sigma\tau} - \delta_{\rho\tau} M_{\sigma\lambda}) . \quad (7.54)$$

So far we have found the operators generating the Lorentz group. Since the generator of the parallel translation is  $P_\mu$ , we have to give the commutation relations involving  $P_\mu$ , i.e.,

$$[M_{\mu\nu}, P_\sigma] = i(\delta_{\mu\sigma} P_\nu - \delta_{\nu\mu} P_\sigma) , \quad (7.55)$$

$$[P_\mu, P_\nu] = 0 . \quad (7.56)$$

The generators of the Poincaré group form a closed algebra under the commutation relations. Note that even if the conservation law (7.52) holds, not all the  $M_{\rho\sigma}$  commute with  $P_\mu$ , e.g., with the Hamiltonian. This is because a coordinate is included explicitly in (7.51), so we need to use (7.38).

Next, we introduce the symmetric energy–momentum tensor. We write  $\mathcal{M}$  as

$$\mathcal{M}_{\mu\nu\sigma} = x_\rho \mathcal{T}_{\mu\sigma} - x_\sigma \mathcal{T}_{\mu\rho} + \mathcal{S}_{\mu\rho\sigma} \quad (7.57)$$

and apply  $\partial_\mu$  to both sides to obtain

$$\mathcal{J}_{\sigma\rho} - \mathcal{J}_{\rho\sigma} = -\partial_\mu \mathcal{S}_{\mu\rho\sigma} . \quad (7.58)$$

This implies that the canonical energy–momentum tensor is not generally symmetric. We thus define a new tensor  $\Theta_{\mu\nu}$  by

$$\Theta_{\rho\sigma} = \mathcal{T}_{\rho\sigma} + \partial_\mu \mathcal{G}_{\mu\rho\sigma} , \quad (7.59)$$

where

$$\mathcal{G}_{\mu\rho\sigma} = \frac{1}{2}(\mathcal{S}_{\mu\rho\sigma} + \mathcal{S}_{\rho\sigma\mu} + \mathcal{S}_{\sigma\rho\mu}) . \quad (7.60)$$

Since  $\mathcal{S}_{\mu\rho\sigma}$  is anti-symmetric with respect to  $\rho$  and  $\sigma$ ,

$$\begin{aligned} \mathcal{G}_{\rho\mu\sigma} &= \frac{1}{2}(\mathcal{S}_{\rho\mu\sigma} + \mathcal{S}_{\mu\sigma\rho} + \mathcal{S}_{\sigma\rho\mu}) \\ &= -\frac{1}{2}(\mathcal{S}_{\rho\sigma\mu} + \mathcal{S}_{\mu\rho\sigma} + \mathcal{S}_{\sigma\rho\mu}) - \mathcal{G}_{\mu\rho\sigma} . \end{aligned} \quad (7.61)$$

Therefore,  $\mathcal{G}$  is anti-symmetric with respect to the first two indices, i.e.,

$$\partial_\mu \mathcal{G}_{\mu\rho\sigma} = 0 , \quad (7.62)$$

$$\partial_\rho \Theta_{\rho\sigma} = \partial_\rho \mathcal{T}_{\rho\sigma} = 0 . \quad (7.63)$$

Moreover,

$$\begin{aligned} \partial_\mu (\mathcal{G}_{\mu\rho\sigma} - \mathcal{G}_{\mu\sigma\rho}) &= \frac{1}{2} \partial_\mu (\mathcal{S}_{\mu\rho\sigma} + \mathcal{S}_{\rho\sigma\mu} + \mathcal{S}_{\sigma\rho\mu} - \mathcal{S}_{\mu\sigma\rho} - \mathcal{S}_{\sigma\rho\mu} - \mathcal{S}_{\rho\sigma\mu}) \\ &= \partial_\mu \mathcal{S}_{\mu\rho\sigma} . \end{aligned} \quad (7.64)$$

Combining this with (7.58) and (7.59), it is straightforward to see that the tensor  $\Theta_{\rho\sigma}$  is symmetric in  $\rho$  and  $\sigma$ , i.e.,

$$\Theta_{\rho\sigma} = \Theta_{\sigma\rho} . \quad (7.65)$$

It is thus referred to as the *symmetric energy–momentum tensor*. Since this tensor was first introduced by Belinfante and Rosenfeld [92, 93], it is also called the

*Belinfante–Rosenfeld tensor.*<sup>1</sup> In addition, taking into account the fact that  $\mathcal{G}_{00v} = 0$ ,

$$P_v = \int d^3x \mathcal{J}_{0v}(x) = \int d^3x \Theta_{0v}(x) . \quad (7.66)$$

Similarly,

$$M_{\rho\sigma} = \int d^3x \mathcal{M}_{0\rho\sigma} = \int d^3x (x_\rho \Theta_{0\sigma} - x_\sigma \Theta_{0\rho}) . \quad (7.67)$$

To prove this, we have to show that

$$\int d^3x (x_\rho \partial_\mu \mathcal{G}_{\mu 0\sigma} - x_\sigma \partial_\mu \mathcal{G}_{\mu 0\rho}) = \int d^3x \mathcal{S}_{0\rho\sigma} .$$

Since  $\mu = 4$  does not produce any effects on the left-hand side,

$$\begin{aligned} (\text{left-hand side}) &= \int d^3x \partial_\mu (x_\rho \mathcal{G}_{\mu 0\sigma} - x_\sigma \mathcal{G}_{\mu 0\rho}) + \int d^3x (\mathcal{G}_{\sigma 0\rho} - \mathcal{G}_{\rho 0\sigma}) \\ &= \frac{1}{2} \int d^3x (\mathcal{S}_{\sigma 0\rho} + \mathcal{S}_{0\rho\sigma} + \mathcal{S}_{\rho 0\sigma} - \mathcal{S}_{\rho 0\sigma} - \mathcal{S}_{0\sigma\rho} - \mathcal{S}_{\sigma 0\rho}) \\ &= \int d^3x \mathcal{S}_{0\rho\sigma} = (\text{right-hand side}) , \end{aligned}$$

which proves (7.67). However, in the above proof we dropped the three-dimensional integral of the three-dimensional divergence. Moreover, it is not the canonical energy–momentum tensor but the symmetric energy–momentum tensor that becomes the source for the gravitational field. Using standard notation, the equation for the gravitational field is

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa \Theta_{\mu\nu} . \quad (7.68)$$

Next, we consider the spin of a particle. To do so, we introduce the Pauli–Liubanski operator

$$W_\alpha = \frac{1}{2i} \epsilon_{\alpha\beta\gamma\delta} M_{\beta\gamma} P_\delta , \quad (7.69)$$

where  $\epsilon_{\alpha\beta\gamma\delta}$  takes the values  $+1$  or  $-1$  when  $(\alpha, \beta, \gamma, \delta)$  is an even or odd permutation of  $(1, 2, 3, 4)$ , respectively, and zero otherwise. Although this axial

---

<sup>1</sup> In the original Japanese edition, the author referred to this as the Belinfante tensor instead of the Belinfante–Rosenfeld tensor.

vector commutes with  $P_\mu$ , it transforms as a vector under a proper Lorentz transformation:

$$[P_\mu, W_\alpha] = 0 . \quad (7.70)$$

The Poincaré algebra has two Casimir operators, i.e., operators which commute with all components of  $P$  and  $M$  :

$$P^2 = P_\mu P_\mu , \quad W^2 = W_\alpha W_\alpha . \quad (7.71)$$

We use these two quantities to define the mass and spin of a particle. First, note that  $W_\alpha$  has the properties

$$W_\alpha P_\alpha = 0 \quad (7.72)$$

and

$$[W_\alpha, W_\beta] = \epsilon_{\alpha\beta\gamma\delta} W_\gamma P_\delta . \quad (7.73)$$

We can diagonalize any set of operators which commute with each other, such as

$$P_\mu , \quad W_3 , \quad P^2 , \quad W^2 . \quad (7.74)$$

Assuming that the eigenvalue of  $-P^2$  is  $m^2$ ,  $m$  becomes the mass of this particle. When  $m \neq 0$ , we can pick out the rest frame of this particle, choosing the eigenvalues  $p_\mu$  of  $P_\mu$  to be

$$\mathbf{p} = 0 , \quad p_0 = m . \quad (7.75)$$

In this case, if we write  $\mathbf{W} = m\mathbf{S}$ , the three components of  $\mathbf{S}$  satisfy the same commutation relations as those of the angular momentum. This is clear from (7.73) and (7.75). Hence, the eigenvalue of  $W^2$  has the form,  $m^2 s(s+1)$ , where  $s$  stands for the spin of this particle. In fact, for the one-particle state in the rest frame,  $\mathbf{W}$  has the representation

$$W_i \rightarrow \frac{m}{2i} \epsilon_{ijk} T_{jk} . \quad (7.76)$$

When  $m = 0$ , things become more difficult. We will discuss this later.

## 7.5 Representations of the Lorentz Group

The commutation relations for the six creation operators of the Lorentz group are given by (7.54). In this section, we will discuss the issue of representations. To do so, we introduce the notation

$$\mathbf{M} = (M_{23}, M_{31}, M_{12}), \quad \mathbf{N} = (M_{14}, M_{24}, M_{34}). \quad (7.77)$$

Then (7.54) can be decomposed into the following three equations:

$$[M_i, M_j] = i\epsilon_{ijk}M_k, \quad [M_i, N_j] = i\epsilon_{ijk}N_k, \quad [N_i, N_j] = i\epsilon_{ijk}M_k. \quad (7.78)$$

Next, we introduce two linear combinations of  $\mathbf{M}$  and  $\mathbf{N}$ :

$$\mathbf{J} = \frac{1}{2}(\mathbf{M} + \mathbf{N}), \quad \mathbf{K} = \frac{1}{2}(\mathbf{M} - \mathbf{N}). \quad (7.79)$$

Note first that  $\mathbf{J}$  and  $\mathbf{K}$  commute, i.e.,

$$[J_i, K_j] = 0. \quad (7.80)$$

In addition, the commutation relations among the  $\mathbf{J}$  and those among the  $\mathbf{K}$  are isomorphic to those of the three components of the angular momentum:

$$[J_i, J_j] = i\epsilon_{ijk}J_k, \quad [K_i, K_j] = i\epsilon_{ijk}K_k. \quad (7.81)$$

Therefore, the proper Lorentz group can be decomposed into a direct product of two three-dimensional rotational groups:

$$SO(3) \times SO(3) \sim SU(2) \times SU(2). \quad (7.82)$$

The representation of the proper Lorentz group is characterized by the eigenvalues of the operators

$$\mathbf{J}^2, \quad J_3, \quad \mathbf{K}^2, \quad K_3, \quad (7.83)$$

i.e.,  $j_1(j_1 + 1)$ ,  $m_1$ ,  $j_2(j_2 + 1)$ , and  $m_2$ , respectively. We denote the corresponding eigenvector by

$$|j_1, m_1; j_2, m_2\rangle. \quad (7.84)$$

Here, as in the case of the angular momentum, the possible values of  $m$  are

$$m_1 = -j_1, -j_1 + 1, \dots, j_1, \quad m_2 = -j_2, -j_2 + 1, \dots, j_2. \quad (7.85)$$

When we use the basis in (7.84) to represent the proper Lorentz transformation  $\Lambda$ , we denote this representation by

$$\mathcal{D}^{(j_1, j_2)}(\Lambda) . \quad (7.86)$$

The relationship between this representation and the three-dimensional rotation group is given by

$$\mathbf{M} = \mathbf{J} + \mathbf{K} . \quad (7.87)$$

It turns out that, for three-dimensional rotations,  $j$  is given by the composition of  $j_1$  and  $j_2$ . The proper Lorentz transformation  $\Lambda$  is expressed by

$$\mathcal{D}(\Lambda) = \exp\left(\frac{1}{2}i\epsilon_{\mu\nu}M_{\mu\nu}\right) . \quad (7.88)$$

We use the following notation with six parameters  $\epsilon_{\mu\nu}$  :

$$\boldsymbol{\theta} = (\epsilon_{23}, \epsilon_{31}, \epsilon_{12}) , \quad i\boldsymbol{\omega} = (\epsilon_{14}, \epsilon_{24}, \epsilon_{34}) , \quad (7.89)$$

where  $\boldsymbol{\theta}$  and  $\boldsymbol{\omega}$  are real three-dimensional vectors. We can now rewrite (7.88) as

$$\mathcal{D}(\Lambda) = \exp(i\boldsymbol{\theta} \cdot \mathbf{M} - \boldsymbol{\omega} \cdot \mathbf{N}) . \quad (7.90)$$

In order to analyze the meaning of the two  $j$ 's, we compare  $\mathcal{D}^{(j,0)}$  with  $\mathcal{D}^{(0,j)}$ :

$$\mathcal{D}^{(j,0)} : \quad \mathbf{M} = \mathbf{N} = \mathbf{J} , \quad \mathbf{K} = 0 , \quad (7.91a)$$

$$\mathcal{D}^{(0,j)} : \quad \mathbf{M} = -\mathbf{N} = \mathbf{K} , \quad \mathbf{J} = 0 . \quad (7.91b)$$

Both representations can be written in terms of  $\mathbf{M}$  as follows:

$$\mathcal{D}^{(j,0)}(\Lambda) = \exp[(i\boldsymbol{\theta} - \boldsymbol{\omega}) \cdot \mathbf{M}] , \quad (7.92a)$$

$$\mathcal{D}^{(0,j)}(\Lambda) = \exp[(i\boldsymbol{\theta} + \boldsymbol{\omega}) \cdot \mathbf{M}] . \quad (7.92b)$$

These imply

$$\mathcal{D}^{(0,j)}(\Lambda) = \mathcal{D}^{(j,0)}(\Lambda^{-1})^\dagger . \quad (7.93)$$

To investigate the physical meaning of the two kinds of indices, we recall Weyl's equation. As already mentioned in Sect. 7.3,  $M_{\mu\nu}$  can be represented by  $-iT_{\mu\nu}$ . In the case of Weyl's equation, from (3.135), depending on the sign of helicity, we have

$$\mathbf{M} = \pm \mathbf{N} . \quad (7.94)$$



Therefore,  $j_1$  and  $j_2$  correspond to the positive and negative helicity states, respectively.

For space inversion, the transformation properties of these operators are

$$\mathbf{M} \rightarrow \mathbf{M}, \quad \mathbf{N} \rightarrow -\mathbf{N}, \quad \mathbf{J} \rightleftharpoons \mathbf{K}. \quad (7.95)$$

Space inversion does not belong to the proper Lorentz group, so  $\mathcal{D}^{(j_1, j_2)}$  is not enough here. It turns out that

$$\mathcal{D}^{(j_1, j_2)} \rightleftharpoons \mathcal{D}^{(j_2, j_1)}. \quad (7.96)$$

This is the reason why Weyl's equation is not invariant under space inversion. The solution of Dirac's equation preserving parity is

$$\mathcal{D}^{(1/2, 0)} \oplus \mathcal{D}^{(0, 1/2)}. \quad (7.97)$$

## 7.6 Spin of a Massless Particle

In Sect. 7.3, we discussed the spin of a particle with nonzero mass. In this section, we will study the spin of a massless particle and discuss its properties.

First, we express  $W_\mu$  using the notation in (7.77):

$$W_0 = \mathbf{M} \cdot \mathbf{P}, \quad \mathbf{W} = i\mathbf{N} \times \mathbf{P} + \mathbf{M}P_0. \quad (7.98)$$

Moreover, we know that  $W_\mu$  is orthogonal to  $P_\mu$ . Considering the irreducible representation  $\mathcal{D}^{(j, 0)}$  and noting that (7.91a) implies  $\mathbf{M} = \mathbf{N} = \mathbf{J}$  and  $\mathbf{K} = 0$ , (7.98) can be written

$$W_0 = \mathbf{J} \cdot \mathbf{P}, \quad \mathbf{W} = i\mathbf{J} \times \mathbf{P} + \mathbf{J}P_0. \quad (7.99)$$

As mentioned before, these imply

$$W^2 = m^2 j(j+1). \quad (7.100)$$

Considering a massless particle moving in the  $z$ -direction, we set

$$P_1 = P_2 = 0, \quad P_3 = P_0 = P, \quad (7.101)$$

where  $P$  is a c-number eigenvalue. Therefore,

$$W_0 = PJ_3, \quad W_1 = P(J_1 + iJ_2), \quad W_2 = -iW_1, \quad W_3 = W_0. \quad (7.102)$$

Since the mass is zero,

$$P_\mu^2 = W_\mu^2 = W_\mu P_\mu = 0 . \quad (7.103)$$

If  $W_\mu$  is a real vector, the equation above means that  $P_\mu$  is parallel to  $W_\mu$ . We call this a *real condition*. Denoting the simultaneous eigenstate of  $\mathbf{J}^2$  and  $j_3$  by  $|j, m\rangle$ , this condition implies

$$W_\mu |j, m\rangle = \lambda P_\mu |j, m\rangle , \quad (7.104)$$

where  $j$  is the value of the spin. However, since  $P_1 = P_2 = 0$ , the equation above implies

$$W_1 |j, m\rangle = W_2 |j, m\rangle = 0 . \quad (7.105)$$

Referring to (7.102),

$$(J_1 + iJ_2)|j, m\rangle = 0 . \quad (7.106)$$

This holds only for  $m = j$ . If parity is conserved,  $\mathcal{D}^{(0,j)}$  is allowed, and in this case  $m = -j$  is also possible. It is a property of the massless case that the only possible value of  $|m|$  is  $j$ . If a different value appears, it means that the state does not belong to an irreducible spin representation. Thus, the form of the equation itself must ensure  $|m| = j$ . Examples of such particles are the neutrino, the photon, and the graviton. For the neutrino,  $m$  is either  $1/2$  or  $-1/2$ , and it breaks parity. On the other hand, since the photon has interactions which preserve parity,  $m = \pm 1$  are both possible, while  $m = 0$  is not. For the photon, this is well-known: the photon is a transverse wave. We shall now check that both the Weyl equation and Maxwell equations ensure  $|m| = j$ .

### Weyl Equation

Weyl's equation corresponds to representations of the type  $\mathcal{D}^{(1/2,0)}$ , and (3.135) implies that

$$H = \boldsymbol{\sigma} \cdot \mathbf{P} = 2\mathbf{J} \cdot \mathbf{P} . \quad (7.107)$$

Therefore, for a neutrino going in the  $z$ -direction,

$$P = 2J_3 P , \quad \text{or} \quad J_3 = \frac{1}{2} . \quad (7.108)$$

Hence, we confirm that  $m = j$ .

### Maxwell Equations

These equations describe a spin-1 photon. The equations for a free electromagnetic field are

$$\text{rot } \mathbf{E} + \frac{\partial \mathbf{H}}{\partial t} = 0, \quad \text{rot } \mathbf{H} - \frac{\partial \mathbf{E}}{\partial t} = 0. \quad (7.109)$$

First, let us check the solution corresponding to  $\mathcal{D}^{(1,0)}$ . In order to obtain the right-handed circularly-polarized solution corresponding to this representation, we choose the following linear combinations of the components of the field strength  $F_{\mu\nu}$ :

$$F_1^{\text{R}} = F_{23} + F_{14}, \quad F_2^{\text{R}} = F_{31} + F_{24}, \quad F_3^{\text{R}} = F_{12} + F_{34}, \quad (7.110)$$

or in vector form,

$$\mathbf{F}^{\text{R}} = \mathbf{H} - i\mathbf{E}. \quad (7.111)$$

In the following, we shall drop the superscript R. The Maxwell equation for  $\mathbf{F}$  is

$$i \frac{\partial}{\partial t} \mathbf{F} = \text{rot } \mathbf{F}. \quad (7.112)$$

We introduce the following three components:

$$F_+ = -\frac{1}{\sqrt{2}}(F_1 - iF_2), \quad F_0 = F_3, \quad F_- = \frac{1}{\sqrt{2}}(F_1 + iF_2). \quad (7.113)$$

Denoting the wave function with these three components by  $\mathcal{F}$ , (7.112) transforms to

$$\frac{\partial}{\partial t} \mathcal{F} + \mathbf{S} \cdot \nabla \mathcal{F} = 0, \quad (7.114)$$

where  $\mathbf{S}$  has the three components

$$S_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (7.115)$$

These three components satisfy the same commutation relations as the angular momentum. For  $\mathcal{D}^{(1,0)}$ , since we have (7.91a),

$$\mathbf{M} = \mathbf{N} = \mathbf{J} = \mathbf{S}, \quad \mathbf{K} = 0. \quad (7.116)$$

Here we choose the  $x$ -dependence of  $\mathcal{F}$  corresponding to a traveling wave in the  $z$ -direction:

$$\mathcal{F}(x) = \mathcal{F} \exp[iP(x_3 - x_0)] . \quad (7.117)$$

Inserting this into (7.114), we obtain

$$P(1 - S_3)\mathcal{F} = 0 . \quad (7.118)$$

What we can understand from this is that, as expected, only  $S_3 = 1$  can be realized.

To obtain the left-handed solution corresponding to  $\mathcal{D}^{(0,1)}$ , instead of (7.111), we start with

$$\mathbf{F}^L = \mathbf{H} + i\mathbf{E} . \quad (7.119)$$

In this case, only  $S_3 = -1$  can be realized. There is no solution corresponding to  $S_3 = 0$ .

What we see from the above is that, if we start with the correct equation, only the solution corresponding to  $|m| = j$  can exist.

## 7.7 Pauli–Gürsey Group

In quantum mechanics, an absolute phase of the wave function cannot be observed. Two wave functions  $\psi_1$  and  $\psi_2$  related by

$$\psi_2 = e^{i\lambda}\psi_1 \quad (7.120)$$

are completely equivalent, in the sense that there can be no difference in their physical interpretation. Therefore, we can say that a set of wave functions differing by constant phases is an equivalence class as far as physics is concerned. Just the same thing happens in field theory.

For an interacting system including complex fields, the Lagrangian density is

$$\mathcal{L} = \mathcal{L}[\varphi_\alpha, \varphi_\alpha^\dagger; \varphi_{\alpha,\mu}, \varphi_{\alpha,\mu}^\dagger] . \quad (7.121)$$

This includes physical constants such as a mass  $m$  and a coupling constant  $g$ . If we compute observable quantities such as differential cross-sections for various reactions and the lifetimes of unstable particles, the results will be given as functions of the physical constants, momenta, and so on. Needless to say, in the final results, the fields themselves do not appear. Metaphorically speaking, the field is like a

Japanese *nabe*<sup>2</sup> used for cooking, and the physical constants are like the ingredients of the dish. Clearly, just as we could use a different *nabe* to cook, we can use a different symbol instead of  $\varphi_\alpha$  in (7.121), say  $\varphi'_\alpha$ , and the final observable result should be the same. Hence, for the complex field in (7.121), we introduce the constant phase transformation

$$\varphi_\alpha \rightarrow \varphi'_\alpha = e^{i\lambda_\alpha} \varphi_\alpha . \quad (7.122)$$

Denoting the group formed by such phase transformations by  $\mathcal{G}$ , we assume that the free part  $\mathcal{L}_f$  of the Lagrangian density is invariant under  $\mathcal{G}$ , i.e.,

$$\mathcal{L}_f[e^{i\lambda_\alpha} \varphi_\alpha, e^{-i\lambda_\alpha} \varphi_\alpha^\dagger; e^{i\lambda_\alpha} \varphi_{\alpha,\mu}, e^{-i\lambda_\alpha} \varphi_{\alpha,\mu}^\dagger] = \mathcal{L}_f[\varphi_\alpha, \varphi_\alpha^\dagger; \varphi_{\alpha,\mu}, \varphi_{\alpha,\mu}^\dagger] . \quad (7.123)$$

The problem is to see how the interaction part  $\mathcal{L}_{\text{int}}$  of the Lagrangian density behaves under  $\mathcal{G}$ . If  $\mathcal{L}_{\text{int}}$  can be kept invariant under  $\mathcal{G}$ , by choosing  $\mathcal{G}$  in a suitable way, a conservation law can be obtained as before. For example, denoting the charge associated with the field quantity  $\varphi_\alpha$  by  $e_\alpha$  and choosing the phase  $\lambda_\alpha$  to be

$$\lambda_\alpha = \lambda e_\alpha , \quad (7.124)$$

the interaction part does indeed remain invariant. We obtain the law of conservation of electric charge in this way.

But what if the interaction part is not invariant? In this case, the coupling constant transforms. As an example, consider the fields  $\psi_1$ ,  $\psi_2$ , and  $\varphi$ :

$$\mathcal{L}_{\text{int}} = g \bar{\psi}_1 \psi_2 \varphi + g^* \bar{\psi}_2 \psi_1 \varphi^\dagger , \quad (7.125)$$

and introduce the phase transformation

$$\psi_1 \rightarrow \psi_1 , \quad \psi_2 \rightarrow \psi_2 , \quad \varphi \rightarrow e^{i\lambda} \varphi . \quad (7.126)$$

Inserting (7.126) into (7.125),  $\mathcal{L}_{\text{int}}$  transforms to

$$\mathcal{L}_{\text{int}} \rightarrow g e^{i\lambda} \bar{\psi}_1 \psi_2 \varphi + g^* e^{-i\lambda} \bar{\psi}_2 \psi_1 \varphi^\dagger . \quad (7.127)$$

Comparing this with (7.125), it turns out that the coupling constant must transform according to

$$g \rightarrow g e^{i\lambda} . \quad (7.128)$$

---

<sup>2</sup> Hot pot.

Since the transformation above does not change the physical content of the theory, this means that any observable quantity depending on  $g$  must be invariant under this transformation, i.e.,

$$\mathcal{O}(ge^{i\lambda}) = \mathcal{O}(g) . \quad (7.129)$$

For a simple transformation like this,  $|g|^2 = g^*g$  is an invariant quantity and all observables become functions of  $|g|^2$  alone. This holds even if the phase transformation (7.126) is generalized much further.

Indeed, Pauli generalized the above group of phase transformations, which is commutative, to a non-commutative group [94]. His target was the neutrino field. Here we write down the Lagrangian density of the massless neutrino:

$$\mathcal{L}_f = -\bar{\psi} \gamma_\mu \partial_\mu \psi . \quad (7.130)$$

Strictly speaking, we ought to write  $\mathcal{L}_f$  in a symmetric form with respect to  $\psi$  and  $\bar{\psi}$ . Bearing this in mind, Pauli pointed out the invariance of (7.130) under the transformation

$$\psi \rightarrow \psi' = a\psi + b\gamma_5 C \bar{\psi}^T , \quad \bar{\psi} \rightarrow \bar{\psi}' = a^* \bar{\psi} - b^* C^{-1} \gamma_5 \psi . \quad (7.131)$$

The inverse transformation is

$$\psi' \rightarrow \psi = a^* \psi' - b\gamma_5 C \bar{\psi}'^T , \quad \bar{\psi}' \rightarrow \bar{\psi} = a \bar{\psi}' + b^* C^{-1} \gamma_5 \psi' , \quad (7.132)$$

where the complex numbers  $a$  and  $b$  satisfy the condition

$$|a|^2 + |b|^2 = 1 . \quad (7.133)$$

This transformation group is a generalization of the group of phase transformations, and we denote it by  $\mathcal{G}_1$ .

On the other hand, Gürsey noted that  $\mathcal{L}_f$  is also invariant under a transformation group  $\mathcal{G}_2$  containing transformations of the form [95]

$$\psi \rightarrow \psi' = e^{i\lambda\gamma_5} \psi , \quad \bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} e^{i\lambda\gamma_5} . \quad (7.134)$$

These two groups commute with each other. Using Noether's theorem to derive the conservation law for an infinitesimal transformation which leaves  $\mathcal{L}_f$  invariant, we

obtain the following conserved quantities. For  $\mathcal{G}_1$ ,

$$M_1 = \frac{1}{2} \int d^3x (\psi^\dagger \gamma_5 \gamma_4 C \psi^\dagger + \psi^T C^{-1} \gamma_4 \gamma_5 \psi) , \quad (7.135a)$$

$$M_2 = \frac{1}{2i} \int d^3x (\psi^\dagger \gamma_5 \gamma_4 C \psi^\dagger - \psi^T C^{-1} \gamma_4 \gamma_5 \psi) , \quad (7.135b)$$

$$M_3 = \int d^3x \psi^\dagger \psi , \quad (7.135c)$$

where we have summed over the spinor indices. In addition, for  $\mathcal{G}_2$ ,

$$N = \int d^3x \psi^\dagger \gamma_5 \psi . \quad (7.136)$$

The generators of this group satisfies the commutation relations

$$\left[ \frac{M_i}{2}, \frac{M_j}{2} \right] = i\epsilon_{ijk} \frac{M_k}{2} , \quad (7.137)$$

$$[M_j, N] = 0 . \quad (7.138)$$

This shows that the generalization  $\mathcal{G}_1 \times \mathcal{G}_2$  of the group of phase transformations is isomorphic to  $SU(2) \times U(1)$ . It was Gürsey who discovered this group structure.

It is entirely because the mass is zero that the Lagrangian density (7.130) permits such a large invariance group. If there is a mass term, only the constant phase transformation mentioned at the beginning of this section is allowed, whence  $M_3$  is the only conserved quantity. Conversely, if there are conserved quantities apart from  $M_3$  in an interacting case, this shows that the mass of the neutrino is zero. In the following, we shall seek such quantities.

Generally, this invariance group turns not to be an invariance group when interactions are introduced, since this induces the transformation among coupling constants, as explained above. Observable quantities must be invariant under this transformation among coupling constants. For beta decay, Pauli derived the quantity that remains invariant under the transformation of the coupling constants.

If the neutrino is massless, the symmetry implemented by the invariance group  $\mathcal{G}_1 \times \mathcal{G}_2$  is not completely broken. Some subgroup

$$\mathcal{G}_v \subset \mathcal{G}_1 \times \mathcal{G}_2 \quad (7.139)$$

may still be an invariance group. We shall now derive the group  $\mathcal{G}_v$  which leaves the beta decay interaction invariant.

We write the Lagrangian density of the beta decay interaction as  $\mathcal{L}_\beta$  and denote the generator of  $\mathcal{G}_\nu$  by  $M$ :

$$[\mathcal{L}_\beta, M] = 0 . \quad (7.140)$$

Here,  $M$  has the general form

$$M = cN + c_1M_1 + c_2M_2 + c_3M_3 . \quad (7.141)$$

By a suitable phase transformation, i.e., a transformation belonging to the invariance group  $\mathcal{G}_1 \times \mathcal{G}_2$ , we transform this  $M$  to the standard form

$$M = cN + c'_3M'_3 , \quad c'^2_3 = c^2_1 + c^2_2 + c^2_3 . \quad (7.142)$$

From now on, we shall write  $M'_3$  simply as  $M_3$ . Therefore, from (7.140),

$$[[\mathcal{L}_\beta, cN + c_3M_3], cN - c_3M_3] = 0 . \quad (7.143)$$

Since  $\mathcal{L}_\beta$  is linear in the neutrino fields  $\psi$  and  $\psi^\dagger$  and noting that

$$[[\psi, cN + c_3M_3], cN - c_3M_3] = (c^2 - c^2_3)\psi ,$$

$$[[\psi^\dagger, cN + c_3M_3], cN - c_3M_3] = (c^2 - c^2_3)\psi^\dagger ,$$

the invariance requirement (7.143) holds if we have

$$c^2 - c^2_3 = 0 . \quad (7.144)$$

Therefore, we must choose  $N - M_3$  or  $N + M_3$  as  $M$ . The experimental fact mentioned in Sect. 3.7 requires us to choose

$$M = N - M_3 . \quad (7.145)$$

The finite transformation generated by this creation operator is a kind of *chiral transformation*, and is given by

$$\psi \rightarrow \psi' = \exp[i(\mathbf{1} - \gamma_5)\lambda]\psi , \quad \bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} \exp[i(\mathbf{1} + \gamma_5)\lambda] . \quad (7.146)$$

To be invariant under this transformation,  $\psi$  and  $\bar{\psi}$  must appear in the interaction terms in the combinations

$$(\mathbf{1} + \gamma_5)\psi , \quad \bar{\psi}(\mathbf{1} - \gamma_5) . \quad (7.147)$$



This is nothing but  $\psi_L$ , introduced in Sect. 3.7. Thus, the neutrino can be described by the two-component Weyl spinor. It turns out that this ensures that the mass of the neutrino is also zero in the interacting case.

# Chapter 8

## S-Matrix



Reactions among elementary particles or atomic nuclei can be described by the  $S$ -matrix. In fact, all our knowledge of the various interactions of these particles is included in the  $S$ -matrix. It is the most fundamental quantity in elementary particle physics. Although the  $S$ -matrix can in principle be determined if the Lagrangian density is known, it is hard to compute in practice. Here we will discuss how to compute the  $S$ -matrix using perturbation theory, which is only valid for small coupling constants. This method was developed by Feynman and Dyson. Two different approaches to QED were greatly developed in the postwar era. One was based on the Tomonaga–Schwinger equation [87, 88], exploiting the traditional canonical formalism. The other started out with Feynman’s path integral [96]. This was a rather intuitive approach and its relationship with the traditional method was not clear. Nevertheless, the consequences of the two theories always coincided. Dyson developed a way to calculate the  $S$ -matrix, starting with the Tomonaga–Schwinger equation, and finally proved that it was equivalent to Feynman’s theory, i.e., he established the equivalence of the two theories [97]. The computational method discussed in the following is based on Dyson’s approach.

### 8.1 Definition of the $S$ -Matrix

The  $S$ -matrix has already been discussed in Sect. 6.1, where it was given by the two equations (6.32) and (6.33). In this section, we will see how to compute the expression

$$S = U(\infty, -\infty) . \quad (8.1)$$

The  $S$ -matrix has several important general properties. For example, the property of unitarity which arose as a consequence of (6.26) is related to the law of

conservation of probability:

$$S^\dagger S = SS^\dagger = 1 . \quad (8.2)$$

This discussion of unitarity was intuitive. A more rigorous discussion must be based on the asymptotic condition to be introduced later, which concerns the unitarity of the renormalized *S*-matrix.

Another property is the Poincaré invariance of the *S*-matrix, which is confirmed by computations. This invariance is expressed by

$$[S, P_\mu] = 0 , \quad [S, M_{\mu\nu}] = 0 . \quad (8.3)$$

Of course, if there are external (background) fields, this property is lost.

## 8.2 Dyson's Formula for the *S*-Matrix

Although it is possible to start from the Tomonaga–Schwinger equation, we shall use Schrödinger's equation in the interaction representation because the result is the same for both approaches. We rewrite (6.10) in the form

$$i \frac{\partial}{\partial t} U(t, t_0) = H_{\text{int}}(t) U(t, t_0) , \quad (8.4)$$

where

$$H_{\text{int}}(t) = \int_{x_0=t} d^3x \mathcal{H}_{\text{int}} . \quad (8.5)$$

With the initial condition  $U(t_0, t_0) = 1$ , we integrate (8.4) to obtain

$$U(t, t_0) = 1 - i \int_{t_0}^t dt' H_{\text{int}}(t') U(t', t_0) . \quad (8.6)$$

Then we insert this expression for  $U$  iteratively into the right-hand side of (8.6) to give

$$U(t, t_0) = 1 - i \int_{t_0}^t dt' H_{\text{int}}(t') + (-i)^2 \int_{t_0}^t dt' H_{\text{int}}(t') \int_{t_0}^{t'} dt'' H_{\text{int}}(t'') + \dots . \quad (8.7)$$

However, since this form is not very useful as it stands, we introduce Dyson's time-ordering operator  $T$ , where

$$T[H_{\text{int}}(t_1)H_{\text{int}}(t_2)\dots H_{\text{int}}(t_n)] = H_{\text{int}}(t'_1)H_{\text{int}}(t'_2)\dots H_{\text{int}}(t'_n), \quad (8.8)$$

and  $t'_1, t'_2, \dots, t'_n$  is a reordering of  $t_1, t_2, \dots, t_n$  chosen such that

$$t'_1 > t'_2 > \dots > t'_n. \quad (8.9)$$

From this definition, it is clear that the left-hand side of (8.8) becomes a symmetric function of  $t_1, t_2, \dots, t_n$ .

If  $f(x_1, x_2, \dots, x_n)$  is a symmetric function of  $x_1, x_2, \dots, x_n$ , we may write

$$\begin{aligned} \int_a^b dx_1 \int_a^{x_1} dx_2 \dots \int_a^{x_{n-1}} dx_n f(x_1, x_2, \dots, x_n) \\ = \frac{1}{n!} \int_a^b dx_1 \int_a^b dx_2 \dots \int_a^b dx_n f(x_1, x_2, \dots, x_n). \end{aligned} \quad (8.10)$$

Using this formula, we have

$$\begin{aligned} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n H_{\text{int}}(t_1)H_{\text{int}}(t_2)\dots H_{\text{int}}(t_n) \\ = \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n T[H_{\text{int}}(t_1)H_{\text{int}}(t_2)\dots H_{\text{int}}(t_n)] \\ = \frac{1}{n!} \int_{t_0}^t dt_1 \int_{t_0}^t dt_2 \dots \int_{t_0}^t dt_n T[H_{\text{int}}(t_1)H_{\text{int}}(t_2)\dots H_{\text{int}}(t_n)]. \end{aligned} \quad (8.11)$$

Using the above expression, we can rewrite the general term in the expansion (8.7). Hence, we obtain Dyson's formula for  $U(t, t_0)$  [98]:

$$U(t, t_0) = 1 + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_{t_0}^t dt_1 \dots \int_{t_0}^t dt_n T[H_{\text{int}}(t_1)\dots H_{\text{int}}(t_n)]. \quad (8.12)$$

Taking the limits  $t \rightarrow \infty$  and  $t_0 \rightarrow -\infty$ , and using the fact that  $H_{\text{int}}(t)$  is a spatial integral of  $\mathcal{H}_{\text{int}}(x)$ , we obtain Dyson's formula for the  $S$ -matrix [98]:

$$S = 1 + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \dots \int d^4x_n [\mathcal{H}_{\text{int}}(x_1)\dots \mathcal{H}_{\text{int}}(x_n)], \quad (8.13)$$

where the integral is taken over the whole of spacetime. It is obvious now that  $S$  commutes with  $P_\mu$ .

The next problem is the Lorentz invariance of the *S*-matrix. We note that  $d^4x$  is Lorentz invariant and  $\mathcal{H}_{\text{int}}(x)$  is a Lorentz scalar. Here, a problem is the time-ordering operator *T*. Apparently, the definition of *T* depends on how we choose the time axis. So, does the *S*-matrix itself depend on the choice of time axis?

If the separation between two spacetime points  $x_1$  and  $x_2$  is time-like, i.e.,

$$(x_1 - x_2)^2 = (\mathbf{x}_1 - \mathbf{x}_2)^2 - (t_1 - t_2)^2 < 0 ,$$

then  $t_1 - t_2$  has the same sign for any Lorentz system, and in this case the time order does not depend on the choice of time axis. However, if the separation is space-like, i.e.,

$$(x_1 - x_2)^2 = (\mathbf{x}_1 - \mathbf{x}_2)^2 - (t_1 - t_2)^2 > 0 ,$$

then we may have  $t_1 > t_2$  or  $t_1 < t_2$ , depending on how we choose the time axis. On the other hand, in the case where  $x_1 - x_2$  is space-like, if

$$\mathcal{H}_{\text{int}}(x_1)\mathcal{H}_{\text{int}}(x_2) = \mathcal{H}_{\text{int}}(x_2)\mathcal{H}_{\text{int}}(x_1) , \quad (x_1 - x_2)^2 > 0 , \quad (8.14)$$

then the ordering is not important. But (8.14) is nothing other than the integrability condition (6.20) for the Tomonaga–Schwinger equation. Therefore, as long as this condition is satisfied, *S* commutes with  $M_{\mu\nu}$ , and hence is Lorentz invariant. This proves (8.3).

In fact, the Hamiltonian density may depend on a hypersurface  $\sigma$ , in which case the discussion becomes more complicated. This happens, for example, when the interactions involve derivatives of the fields.

### 8.3 Wick's Theorem

The Feynman–Dyson theory tells us how to compute the *S*-matrix, starting with Dyson's formula derived in the last section. It is Wick's theorem that provides the foundation for this method.

In the interaction picture, a field operator satisfies a free field equation, i.e., a linear equation, as mentioned in Sect. 5.4, so it can be decomposed into a positive-frequency part and a negative-frequency part:

$$\varphi_\alpha(x) = \varphi_\alpha^{(+)}(x) + \varphi_\alpha^{(-)}(x) , \quad (8.15)$$

where  $\varphi_\alpha^{(+)}(x)$  is an annihilation operator and  $\varphi_\alpha^{(-)}(x)$  is a creation operator. It is clear that

$$\varphi_\alpha^{(+)}(x)|0\rangle = 0 , \quad \langle 0|\varphi_\alpha^{(-)}(x) = 0 . \quad (8.16)$$

In the following, we will apply the decomposition (8.15) to field operators in the interaction picture, and express the product of operators using (8.16). First, for simplicity, we consider a neutral scalar field:

$$\begin{aligned} \varphi(x_1)\varphi(x_2)\dots\varphi(x_n) \\ = [\varphi^{(+)}(x_1) + \varphi^{(-)}(x_1)][\varphi^{(+)}(x_2) + \varphi^{(-)}(x_2)]\dots[\varphi^{(+)}(x_n) + \varphi^{(-)}(x_n)]. \end{aligned}$$

Expanding this product, we would like to change the order of the terms, shifting  $(-)$  to the left and  $(+)$  to the right, to give something of the form

$$\varphi^{(-)}\varphi^{(-)}\dots\varphi^{(+)}\varphi^{(+)}\dots. \quad (8.17)$$

However, since  $\varphi^{(-)}$  and  $\varphi^{(+)}$  do not commute, each time we change the order of two operators, a commutator will show up. For example,

$$\begin{aligned} \varphi(x)\varphi(y) &= [\varphi^{(+)}(x) + \varphi^{(-)}(x)][\varphi^{(+)}(y) + \varphi^{(-)}(y)] \\ &= \varphi^{(+)}(x)\varphi^{(+)}(y) + \varphi^{(-)}(x)\varphi^{(-)}(y) + \varphi^{(-)}(x)\varphi^{(+)}(y) + \varphi^{(+)}(x)\varphi^{(-)}(y). \end{aligned}$$

Except for the last term, all operators are in the order indicated in (8.17). Here, using

$$\varphi^{(+)}(x)\varphi^{(-)}(y) = \varphi^{(-)}(y)\varphi^{(+)}(x) + [\varphi^{(+)}(x), \varphi^{(-)}(y)],$$

all terms can be set in the standard form, i.e.,

$$\begin{aligned} \varphi(x)\varphi(y) &= \varphi^{(+)}(x)\varphi^{(+)}(y) + \varphi^{(-)}(x)\varphi^{(-)}(y) + \varphi^{(-)}(x)\varphi^{(+)}(y) + \varphi^{(-)}(y)\varphi^{(+)}(x) \\ &\quad + [\varphi^{(+)}(x), \varphi^{(-)}(y)]. \end{aligned} \quad (8.18)$$

The sum of the first four terms, which are in the standard form (8.17), is called the *normal product*, expressed briefly as  $:\varphi(x)\varphi(y):$ . The last commutator is a  $c$ -number. It can be calculated as follows (see Sect. 4.1):

$$\begin{aligned} [\varphi^{(+)}(x), \varphi^{(-)}(y)] &= \left[ \sum_{\mathbf{p}} \frac{1}{\sqrt{2p_0V}} e^{ip \cdot x} a(\mathbf{p}), \sum_{\mathbf{q}} \frac{1}{\sqrt{2q_0V}} e^{-iq \cdot y} a^\dagger(\mathbf{q}) \right] \\ &= \sum_{\mathbf{p}} \frac{1}{2p_0V} e^{ip \cdot (x-y)} \\ &= i\Delta^{(+)}(x-y). \end{aligned} \quad (8.19)$$

Therefore,

$$\varphi(x)\varphi(y) = :\varphi(x)\varphi(y): + i\Delta^{(+)}(x-y). \quad (8.20)$$

The normal product is clearly symmetric, i.e.,

$$:\varphi(x)\varphi(y): = :\varphi(y)\varphi(x): . \quad (8.21)$$

This symmetry also holds when there are more field operators in the product. However, if the field obeys Fermi statistics, the normal product becomes anti-symmetric. For a Dirac field  $\psi$ , when  $x'_1, x'_2, \dots, x'_n$  is a permutation of  $x_1, x_2, \dots, x_n$ ,

$$:\psi(x'_1)\psi(x'_2)\dots\psi(x'_n): = (-1)^P :\psi(x_1)\psi(x_2)\dots\psi(x_n): , \quad (8.22)$$

where  $(-1)^P$  is equal to  $+1$  or  $-1$  for an even permutation or an odd permutation, respectively. For a field obeying Bose statistics, there is no such change of sign.

From the definition of the normal product and (8.16), any vacuum expectation value of a normal product vanishes. For example,

$$\langle 0 | :\varphi(x)\varphi(y): | 0 \rangle = 0 . \quad (8.23)$$

Therefore, from (8.20),

$$\langle 0 | \varphi(x)\varphi(y) | 0 \rangle = i\Delta^{(+)}(x - y) . \quad (8.24)$$

Generalizing this result, we obtain

$$\langle 0 | \varphi(x_1)\varphi(x_2)\dots\varphi(x_n) | 0 \rangle = \sum_{k_1 < k_2, k_3 < k_4, \dots} i\Delta^{(+)}(x_{k_1} - x_{k_2})\dots i\Delta^{(+)}(x_{k_{n-1}} - x_{k_n}) , \quad (8.25)$$

where  $k_1, k_2, \dots, k_n$  is a permutation of  $1, 2, \dots, n$  and  $n$  is an even number. Moreover, under the restriction of the inequalities  $k_1 < k_2, k_3 < k_4, \dots$ , we sum over all possible combinations. If  $n$  is an odd number, then  $\varphi(x_1)\varphi(x_2)\dots\varphi(x_n)|0\rangle$  is an odd number of quantum states, so it is orthogonal to  $|0\rangle$  and the expectation value (8.25) vanishes.

Using the normal product, the current density considered in (6.56) can be written as

$$j_\mu = \frac{1}{2}ie[\bar{\psi}, \gamma_\mu \psi] = ie : \bar{\psi} \gamma_\mu \psi : . \quad (8.26)$$

Therefore, we can deduce the following equation, which says that the vacuum has no charge:

$$\langle 0 | j_\mu(x) | 0 \rangle = 0 . \quad (8.27)$$

We now move on to the problem of evaluating  $S$ -matrix elements. From Dyson's formula, the matrix element

$$S_{ba} = \langle b|S|a\rangle \quad (8.28)$$

means

$$S_{ba} = \langle b|a\rangle + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \dots \int d^4x_n \langle b|T[\mathcal{H}_{\text{int}}(x_1) \dots \mathcal{H}_{\text{int}}(x_n)]|a\rangle. \quad (8.29)$$

We shall consider here the simplest possible interaction, viz., the Yukawa interaction between a Fermi particle and a Bose particle:

$$\mathcal{H}_{\text{int}}(x) = j(x)\varphi(x), \quad j(x) = g : \bar{\psi}(x)\psi(x) : , \quad (8.30)$$

where  $\psi$  and  $\varphi$  stand for a Dirac field and a neutral scalar field, respectively. We then have the following decomposition:

$$T[\mathcal{H}_{\text{int}}(x_1) \dots \mathcal{H}_{\text{int}}(x_n)] = T[j(x_1) \dots j(x_n)] \dots T[\varphi(x_1) \dots \varphi(x_n)]. \quad (8.31)$$

The states  $|a\rangle$  and  $|b\rangle$  can also be decomposed into the direct product of the states for the scalar particle and the Dirac particle:

$$|a\rangle = |a^S\rangle \otimes |a^D\rangle, \quad |b\rangle = |b^S\rangle \otimes |b^D\rangle. \quad (8.32)$$

Hence,

$$\begin{aligned} \langle b|T[\mathcal{H}_{\text{int}}(x_1) \dots \mathcal{H}_{\text{int}}(x_n)]|a\rangle & \\ &= \langle b^D|T[j(x_1) \dots j(x_n)]|a^D\rangle \langle b^S|T[\varphi(x_1) \dots \varphi(x_n)]|a^S\rangle. \end{aligned} \quad (8.33)$$

We begin with the second factor, which is easier to evaluate. The scalar particle states are expressed in the form

$$|a^S\rangle = a^\dagger(k_1) \dots a^\dagger(k_i)|0^S\rangle, \quad |b^S\rangle = a^\dagger(k'_1) \dots a^\dagger(k'_f)|0^S\rangle. \quad (8.34)$$

In the following, we simply denote  $|0^S\rangle$  by  $|0\rangle$ . Then, extracting  $i$  annihilation operators and  $f$  creation operators, we obtain

$$\begin{aligned} \langle b^S|T[\varphi(x_1) \dots \varphi(x_n)]|a^S\rangle & \\ &= \langle 0|a(k'_1) \dots a(k'_f)T[\varphi(x_1) \dots \varphi(x_n)]a^\dagger(k_1) \dots a^\dagger(k_i)|0\rangle. \end{aligned}$$



In this evaluation, we assume that none of the momenta  $k_1, \dots, k_i$  is equal to any of the momenta  $k'_1, \dots, k'_f$ . If there are equal momenta, this corresponds to a situation where particles go straight through without interaction. We begin by eliminating  $a^\dagger(k_i)$ . To do this, we use (8.16) and bring in the commutator with  $a^\dagger(k_i)$ :

$$\begin{aligned} \langle 0|AB \dots a^\dagger(k_i)|0\rangle &= \langle 0|[AB \dots, a^\dagger(k_i)]|0\rangle + \langle 0|a^\dagger(k_i)AB \dots |0\rangle \\ &= \langle 0|[AB \dots, a^\dagger(k_i)]|0\rangle. \end{aligned} \quad (8.35)$$

To evaluate the commutator, we use

$$[a(k'), a^\dagger(k)] = \delta_{k',k} \quad (8.36)$$

and

$$\langle \varphi(x), a^\dagger(k) \rangle = \frac{1}{\sqrt{2k_0V}} e^{ik \cdot x} = \langle 0|\varphi(x)|k \rangle. \quad (8.37)$$

Given the assumption of no equal momenta,  $a$  and  $a^\dagger$  commute, whence

$$\begin{aligned} \langle 0|a(k'_1) \dots a(k'_f) T[\varphi(x_1) \dots \varphi(x_n)] a^\dagger(k_1) \dots a^\dagger(k_i) |0\rangle \\ = \sum_{j=1}^n \langle 0|a(k'_1) \dots a(k'_f) T[\varphi(x_1) \dots \varphi(x_{j-1}) \varphi(x_{j+1}) \dots \varphi(x_n)] \\ \times a^\dagger(k_1) \dots a^\dagger(k_{i-1}) |0\rangle \langle 0|\varphi(x_j)|0 \rangle. \end{aligned}$$

Iterating, we eventually obtain

$$\begin{aligned} \langle 0|a(k'_1) \dots a(k'_f) T[\varphi(x_1) \dots \varphi(x_n)] a^\dagger(k_1) \dots a^\dagger(k_i) |0\rangle \\ = \sum_{\text{comb}} \langle k'_1 | \varphi(x'_1) | 0 \rangle \dots \langle k'_f | \varphi(x'_f) | 0 \rangle \langle 0 | \varphi(x'_{f+1}) | k_1 \rangle \dots \langle 0 | \varphi(x'_{f+i}) | k_i \rangle \\ \times \langle 0 | T[\varphi(x'_{f+i+1}) \dots \varphi(x'_n)] | 0 \rangle, \end{aligned} \quad (8.38)$$

where  $x'_1, \dots, x'_n$  is a permutation of  $x_1, \dots, x_n$ , and selecting  $(f+i)$  variables from  $n$  variables, we have summed over all the possible combinations that can be written in the form above. We can also write the above equation in the form

$$\begin{aligned} \langle 0|a(k'_1) \dots a(k'_f) T[\varphi(x_1) \dots \varphi(x_n)] a^\dagger(k_1) \dots a^\dagger(k_i) |0\rangle \\ = \sum_{\text{comb}} \langle b^S | : \varphi(x'_1) \dots \varphi(x'_{f+i}) : | a^S \rangle \langle 0 | T[\varphi(x'_{f+i+1}) \dots \varphi(x'_n)] | 0 \rangle, \end{aligned} \quad (8.39)$$

where we have summed over all combinations which separate the  $n$  variables into two groups of  $(f + i)$  and  $(n - f - i)$  variables. Note also that there is a similar equation for operators, viz.,

$$T[\varphi(x_1) \dots \varphi(x_n)] = \sum_{\text{comb}} : \varphi(x'_1) \dots \varphi(x'_j) : \langle 0 | T[\varphi(x'_{j+1}) \dots \varphi(x'_n)] | 0 \rangle, \quad (8.40)$$

where the sum is over all combinations which separate the  $n$  variables into two groups, with one belonging to the normal product and the other to the expectation value.

The next problem is to evaluate the expectation value. To begin with,

$$\begin{aligned} \langle 0 | T[\varphi(x)\varphi(y)] | 0 \rangle &= \theta(x_0 - y_0) \langle 0 | \varphi(x)\varphi(y) | 0 \rangle + \theta(y_0 - x_0) \langle 0 | \varphi(y)\varphi(x) | 0 \rangle \\ &= i\theta(x_0 - y_0) \Delta^{(+)}(x - y) + i\theta(y_0 - x_0) \Delta^{(+)}(y - x) \\ &= \Delta_F(x - y), \end{aligned} \quad (8.41)$$

where we have used (4.29). Generalizing the equation above,

$$\langle 0 | T[\varphi(x_1) \dots \varphi(x_n)] | 0 \rangle = \sum_{\text{comb}} \Delta_F(x'_1 - x'_2) \dots \Delta_F(x'_{n-1} - x'_n), \quad (8.42)$$

where if  $n$  is even we have summed over all combinations which separate  $n$  variables into  $n/2$  pairs and if  $n$  is odd the expression vanishes. To prove this, assuming the time order

$$t'_1 > t'_2 > \dots > t'_n, \quad (t = x_0), \quad (8.43)$$

we obtain

$$\langle 0 | T[\varphi(x_1) \dots \varphi(x_n)] | 0 \rangle = \langle 0 | \varphi(x'_1) \dots \varphi(x'_n) | 0 \rangle. \quad (8.44)$$

Since  $k_1 < k_2$  means that  $t'_{k_1} > t'_{k_2}$ , the definition of  $\Delta_F$  implies in this case that

$$\Delta_F(x'_{k_1} - x'_{k_2}) = i\Delta^{(+)}(x'_{k_1} - x'_{k_2}). \quad (8.45)$$

Although the right-hand side of (8.44) can be written in the form (8.25), taking into account (8.45), it becomes equivalent to (8.42). Since the time order (8.43) is arbitrary, it turns out that (8.42) holds for any time order.

To summarize so far,

$$T[\varphi(x_1) \dots \varphi(x_n)] = \sum_{\text{comb}} : \varphi(x'_1) \dots \varphi(x'_j) : \Delta_F(x'_{j+1} - x'_{j+2}) \dots \Delta_F(x'_{n-1} - x'_n). \quad (8.46)$$

Wick introduced a contraction symbol for  $\Delta_F$  :

$$\overline{\varphi(x)\varphi(y)} = \Delta_F(x - y) . \quad (8.47)$$

This specifies a pair which appears in a vacuum expectation value. Using this symbol, (8.46) can be viewed as an expansion in terms of different contractions:

$$\begin{aligned} T[\varphi(x_1) \dots \varphi(x_n)] &= : \varphi(x_1) \dots \varphi(x_n) : \\ &+ \sum_{j \neq k} : \overline{\varphi(x_j) \dots \varphi(x_k)} \dots \varphi(x_n) : \\ &+ \sum_{j \neq k} \sum_{l \neq m} : \overline{\varphi(x_j) \dots \varphi(x_k)} \dots \overline{\varphi(x_l) \dots \varphi(x_m)} \dots \varphi(x_n) : \\ &+ \dots . \end{aligned} \quad (8.48)$$

As discussed in Sect. 8.2, the T-product is symmetric for boson fields, i.e.,

$$T[\varphi(x'_1) \dots \varphi(x'_n)] = T[\varphi(x_1) \dots \varphi(x_n)] . \quad (8.49)$$

For fermionic fields, it is anti-symmetric in the sense of (8.22):

$$T[\psi(x'_1) \dots \psi(x'_n)] = (-1)^P T[\psi(x_1) \dots \psi(x_n)] . \quad (8.50)$$

Of course, the above equation also holds if  $\psi$  is replaced by  $\bar{\psi}$ . Moreover, the contraction for fermionic fields is

$$\overline{\psi_\alpha(x)\bar{\psi}_\beta(y)} = \langle 0 | T[\psi_\alpha(x)\bar{\psi}_\beta(y)] | 0 \rangle = S_{F\alpha\beta}(x - y) , \quad (8.51)$$

where  $S_F$  is given by

$$S_F(x) = -(\gamma_\mu \partial_\mu - m) \Delta_F(x) . \quad (8.52)$$

Generally, the contraction functions, also known as Feynman propagators, satisfy the following equations:

$$D_{\alpha\beta}(\partial) \Delta_{F\beta\gamma}(x) = i \delta_{\alpha\gamma} \delta^4(x) , \quad (8.53)$$

$$\Delta_{F\alpha\beta}(x) = C_{\alpha\beta}(\partial) \Delta_F(x) . \quad (8.54)$$

Regarding the contraction for fermionic fields, the following equation is trivial:

$$\overline{\psi(x)\psi(y)} = \overline{\bar{\psi}(x)\bar{\psi}(y)} = 0 . \quad (8.55)$$

The general form of Wick's theorem is

$$T[AB \dots Z] = :AB \dots Z: + \sum_{\text{one cont}} :\overline{AB} \dots Z: + \sum_{\text{two cont}} :\overline{ABCD} \dots Z: + \dots, \quad (8.56)$$

where ' $n$  cont' instructs to sum over all possible terms with  $n$  contractions.

Combining Wick's theorem with the above property of the contraction function, we obtain

$$D_{\alpha\beta}(\partial_x)T[\varphi_\beta(x) \dots] = i \frac{\delta}{\delta\varphi_\alpha^\dagger} T[\dots], \quad (8.57)$$

for a complex field. For instance,

$$D_{\alpha\beta}(\partial_x)T[\varphi_\beta(x), \varphi_\gamma^\dagger(y)] = i \frac{\delta\varphi_\gamma^\dagger(y)}{\delta\varphi_\alpha^\dagger(x)} = i\delta_{\alpha\gamma}\delta^4(x-y). \quad (8.58)$$

For a Dirac field,

$$-\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m\right)T[\psi(x) \dots] = i \frac{\delta}{\delta\bar{\psi}(x)} T[\dots]. \quad (8.59)$$

In order to carry out a functional derivative with respect to  $\bar{\psi}(x)$ , we determine the sign by first moving the  $\bar{\psi}$  to be differentiated in the T-product to the extreme left and then carrying out the functional derivative. Since this formula plays an important role, e.g., it can be used to transform the  $S$ -matrix, it is referred to as a *reduction formula*. A similar formula in the Heisenberg picture will be mentioned later.

## 8.4 Feynman Diagrams

To understand the structure of the  $S$ -matrix intuitively, it is useful to introduce the Feynman diagrams. As an example, we explain the idea for QED in the Fermi gauge.

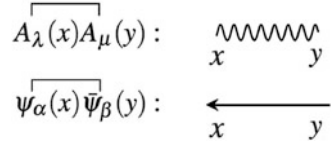
When we expand the  $S$ -matrix in perturbation theory as a power series in the coupling constant, the  $n$ th order term is given by

$$S^{(n)} = \frac{(-i)^n}{n!} \int d^4x_1 \dots \int d^4x_n T[\mathcal{H}_{\text{int}}(x_1) \dots \mathcal{H}(x_n)]. \quad (8.60)$$

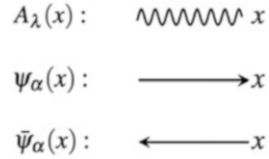
In QED,

$$\mathcal{H}_{\text{int}}(x) = -j_\mu(x)A_\mu(x) = -ie : \bar{\psi}(x)\gamma_\mu\psi(x) : A_\mu(x). \quad (8.61)$$

**Fig. 8.1** Representation of contractions in a Feynman diagram



**Fig. 8.2** Representation of operators in the normal product



We expand the T-product in (8.60) as a sum of normal products using Wick's theorem. In this case, the operators in the normal product create and annihilate free particles. The contraction functions describe the propagation of virtual particles which cannot be observed. In a Feynman diagram, these operators are represented as lines, providing a way to understand their role visually.

When two operators are contracted, this is represented by a line segment (see Fig. 8.1). In this case, the propagation of a virtual photon is expressed by an undirected wavy line and the propagation of an electron by a directed line, in such a way that the direction coincides with the direction of motion of the electric charge  $e$ . If  $x_0 > y_0$ , this means that a virtual electron created at a point  $y$  propagates to a point  $x$  and is annihilated there. If  $x_0 < y_0$ , then a virtual positron created at a point  $x$  propagates to a point  $y$  and is annihilated there. Regarding the undirected photon, a virtual photon is created at a past point, either  $x$  or  $y$ , and annihilated at a future point. This interpretation is obvious from (8.41) (Fig. 8.2).

In contrast, operators in the normal product are expressed as undirected and directed half lines for a photon and an electron, respectively. These lines are called *external lines*. In the case of  $A_\lambda$ , the line corresponds to a photon which is either created or absorbed at a point  $x$ . In the case of  $\psi$ , the line corresponds to either a positron which is created or an electron which is absorbed at a point  $x$ . In the case of  $\bar{\psi}$ , the line corresponds to either an electron which is created or a positron which is absorbed at a point  $x$ .

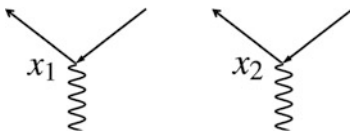
Each term in the expansion of the S-matrix as a sum of normal products corresponds to one Feynman diagram. In the following, we consider the case  $n = 2$ :

$$S^{(2)} = \frac{(-e)^2}{2} \int d^4x_1 \int d^4x_2 T[:\bar{\psi}(x_1)\gamma_\mu\psi(x_1):, : \bar{\psi}(x_2)\gamma_\nu\psi(x_2):]T[A_\mu(x_1), A_\nu(x_2)]. \tag{8.62}$$

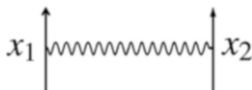
We use Wick's theorem to expand this T-product in terms with different numbers of contractions:

1. When there is no contraction,

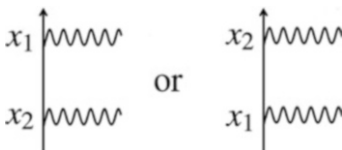
$$:\bar{\psi}(x_1)\gamma_\mu\psi(x_1)\bar{\psi}(x_2)\gamma_\nu\psi(x_2): : A_\mu(x_1)A_\nu(x_2) : . \tag{8.63}$$



**Fig. 8.3** Feynman diagrams for terms in the expansion of the T-product with no contractions



**Fig. 8.4** Feynman diagram for the case where only a pair of A's is contracted in the expansion of the T-product



**Fig. 8.5** Feynman diagrams for the case where a  $\psi$  is contracted with a  $\bar{\psi}$  in the expansion of the T-product

Feynman diagrams corresponding to these terms consist of two disconnected parts (Fig. 8.3). They do not correspond to physically realizable processes since they cannot satisfy the energy–momentum conservation law.

2. When only a pair of A's is contracted,

$$:\bar{\psi}(x_1)\gamma_\mu\psi(x_1)\gamma_\nu\bar{\psi}(x_2)\gamma_\nu\psi(x_2): \delta_{\mu\nu}D_F(x_1 - x_2) . \tag{8.64}$$

This term corresponds to electron–electron scattering, proton–proton scattering, electron–proton scattering, and so on, and the corresponding Feynman diagram is shown in Fig. 8.4.

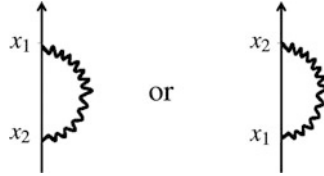
3. When a  $\psi$  is contracted with a  $\bar{\psi}$ ,

$$[:\bar{\psi}(x_1)\gamma_\mu S_F(x_1 - x_2)\gamma_\nu\psi(x_2): + : \bar{\psi}(x_2)\gamma_\nu S_F(x_2 - x_1)\gamma_\mu\psi(x_1):]: A_\mu(x_1)A_\nu(x_2): . \tag{8.65}$$

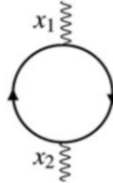
This term, illustrated in Fig. 8.5, expresses Compton scattering, electron–positron annihilation into two photons or creation from two photons, and so on.

4. When a  $\psi$  is contracted with a  $\bar{\psi}$  and a pair of A's are contracted,

$$[:\bar{\psi}(x_1)\gamma_\mu S_F(x_1 - x_2)\gamma_\nu\psi(x_2): + : \bar{\psi}(x_2)\gamma_\nu S_F(x_2 - x_1)\gamma_\mu\psi(x_1):]: \delta_{\mu\nu}D_F(x_1 - x_2) . \tag{8.66}$$



**Fig. 8.6** Electron self-energy. Feynman diagrams for the case where a  $\psi$  is contracted with a  $\bar{\psi}$  and a pair of  $A$ 's is contracted in the expansion of the T-product



**Fig. 8.7** Vacuum polarization. Feynman diagram for the case where two pairs of  $\psi$  and  $\bar{\psi}$  are contracted in the expansion of the T-product

These terms correspond to the self-energy of the electron. Corresponding Feynman diagrams are shown in Fig. 8.6.

5. When two pairs of  $\psi$  and  $\bar{\psi}$  are contracted,

$$- \text{Tr}[\gamma_\mu S_F(x_1 - x_2)\gamma_\nu S_F(x_2 - x_1)] : A_\mu(x_1)A_\nu(x_2) : . \tag{8.67}$$

This corresponds to the phenomenon called *vacuum polarization*, which gives the photon self-energy. The Feynman diagram is shown in Fig. 8.7. Let us see why the minus sign and the trace have appeared:

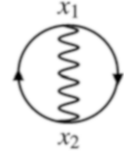
$$\begin{aligned} \overbrace{\bar{\psi}(x_1)\gamma_\mu\psi(x_1)\psi(x_2)\gamma_\nu\psi(x_2)} &= \overbrace{\bar{\psi}(x_1)\gamma_\mu S_F(x_1 - x_2)\gamma_\nu\psi(x_2)} \\ &= \overbrace{\bar{\psi}_\alpha(x_1)[\gamma_\mu S_F(x_1 - x_2)\gamma_\nu]_{\alpha\beta}\psi_\beta(x_2)} \\ &= -[\gamma_\mu S_F(x_1 - x_2)\gamma_\nu]_{\alpha\beta} \overbrace{\psi_\beta(x_2)\bar{\psi}_\alpha(x_1)} \\ &= -[\gamma_\mu S_F(x_1 - x_2)\gamma_\nu]_{\alpha\beta} [S_F(x_2 - x_1)]_{\beta\alpha} \\ &= -\text{Tr}[\gamma_\mu S_F(x_1 - x_2)\gamma_\nu S_F(x_2 - x_1)] . \end{aligned}$$

This shows that the minus sign necessarily comes in when there is a fermion loop, and this will turn out to play an important role in the gauge theory.

6. When all operators are contracted,

$$- \text{Tr}[\gamma_\mu S_F(x_1 - x_2)\gamma_\nu S_F(x_1 - x_2)]\delta_{\mu\nu} D_F(x_1 - x_2) . \tag{8.68}$$

**Fig. 8.8** Vacuum bubble.  
Feynman diagram for the case where all operators are contracted in the expansion of the T-product



This is a process that starts with the vacuum and ends with the vacuum. It is represented by the bubble graph shown in Fig. 8.8. The transition probability amplitude for going from the vacuum to the vacuum, viz.,

$$\langle 0|S|0\rangle \quad (8.69)$$

is obtained using Dyson's formula by summing over all contributions from bubble graphs. By the probability conservation law, the absolute value of this expression should be unity. However, the phase diverges. Adjusting the unobservable phase, we thus redefine the  $S$ -matrix by

$$S' = S/\langle 0|S|0\rangle . \quad (8.70)$$

To compute  $S'$ , it turns out that, dropping all bubble graphs without external lines at the outset, we can compute the  $S$ -matrix using only the remaining graphs.

## 8.5 Examples of $S$ -Matrix Elements

Let us compute the  $S$ -matrix elements for some examples using the method discussed in the last section. In a real situations, we need to calculate cross-sections and decay amplitudes, but we shall discuss these in the next chapter.

We begin by writing the  $S$ -matrix element in the form

$$S_{fi} = \delta_{fi} - i(2\pi)^4 \delta^4(P_f - P_i) T_{fi} . \quad (8.71)$$

The first term corresponds to the case where the incoming particles go past each other without interaction. The second term corresponds to the situation where there are interactions. Here we have used the representation which diagonalizes the total energy-momentum tensor, and this gives rise to the four-dimensional  $\delta$ -function which expresses the conservation law.



### 8.5.1 Compton Scattering

The normal product corresponding to this process was given in the last section. We consider a transition from the state  $|p, k\rangle$  to the state  $|p', k'\rangle$ , where  $p$  and  $p'$  express the four-momenta of the electrons and  $k$  and  $k'$  the four-momenta of the photons. There are two terms in the normal product (8.65), but  $x_1$  and  $x_2$  and  $\mu$  and  $\nu$  are all dummy variables, and since the whole expression is symmetric under the replacements  $x_1 \leftrightarrow x_2$  and  $\mu \leftrightarrow \nu$ , we keep only one term and drop the factor of  $1/2!$ . This yields

$$\begin{aligned} \langle p', k' | S^{(2)} | p, k \rangle &= e^2 \int d^4 x_1 \int d^4 x_2 \langle p' | : \bar{\psi}(x_1) \gamma_\mu S_F(x_1 - x_2) \gamma_\nu \psi(x_2) : | p \rangle \\ &\quad \times \langle k' | : A_\mu(x_1) A_\nu(x_2) : | k \rangle \\ &= e^2 \int d^4 x_1 \int d^4 x_2 \langle p' | \bar{\psi}(x_1) | 0 \rangle \langle 0 | A_\nu(x_2) | k \rangle \\ &\quad \times [\langle k' | A_\mu(x_1) | 0 \rangle \langle 0 | A_\nu(x_2) | k \rangle + \langle k' | A_\nu(x_2) | 0 \rangle \langle 0 | A_\mu(x_1) | k \rangle] . \end{aligned} \quad (8.72)$$

Feynman diagrams in the momentum representation are shown in Fig. 8.9.

In order to evaluate this expression, we use the following matrix elements and the Fourier representation:

$$\langle 0 | A_\mu(x) | k \rangle = \frac{1}{\sqrt{2k_0 V}} e_\mu^{(\lambda)} e^{ik \cdot x} , \quad \langle k' | A_\mu(x) | 0 \rangle = \frac{1}{\sqrt{2k'_0}} e_\mu^{(\lambda')} e^{-ik' \cdot x} ,$$

$$\langle 0 | \psi(x) | p \rangle = \frac{1}{\sqrt{V}} u(p) e^{ip \cdot x} , \quad \langle p' | \bar{\psi}(x) | 0 \rangle = \frac{1}{\sqrt{V}} \bar{u}(p') e^{-ip' \cdot x} ,$$

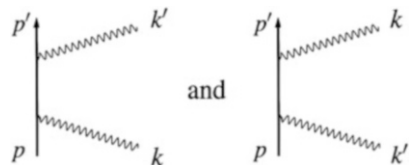
$$S_F(x) = \frac{-i}{(2\pi)^4} \int d^4 p \frac{-i \not{p} \cdot \gamma + m}{p^2 + m^2 - i\epsilon} e^{ip \cdot x} .$$

Inserting these expressions into (8.72), we obtain  $T_{fi}^{(2)}$  as defined in (8.71):

$$T_{fi}^{(2)} = \frac{1}{\sqrt{2k_0 V}} \frac{1}{\sqrt{2k'_0 V}} \frac{1}{\sqrt{V}} \frac{1}{\sqrt{V}} \quad (8.73)$$

$$\times e^2 \bar{u}(p') \left[ (e' \cdot \gamma') \frac{-i(p+k) \cdot \gamma + m}{(p+k)^2 + m^2} (e \cdot \gamma) + (e \cdot \gamma) \frac{-i(p-k') \cdot \gamma + m}{(p-k')^2 + m^2} (e' \cdot \gamma) \right] u(p) ,$$

**Fig. 8.9** Feynman diagrams for Compton scattering



where  $e$  stands for  $e^{(\lambda)}$  and  $e'$  for  $e^{(\lambda')}$ . Using the relation for external momenta on the mass shell, i.e., the Einstein relation (3.35), we have

$$\begin{aligned}(p+k)^2 + m^2 &= p^2 + m^2 + k^2 + 2p \cdot k = 2p \cdot k, \\ (p' - k')^2 + m^2 &= p'^2 + m^2 + k'^2 - 2p' \cdot k' = -2p' \cdot k'.\end{aligned}$$

Using the Dirac equation,

$$(ip \cdot \gamma + m)u(p) = \bar{u}(p')(ip' \cdot \gamma + m) = 0.$$

We can now define the invariant amplitude  $\mathcal{T}_{fi}^{(2)}$  by

$$\mathcal{T}_{fi}^{(2)} = \sqrt{2k_0 V} \sqrt{2k'_0 V} \sqrt{2p_0 V} \sqrt{2p'_0 V} \sqrt{2p'_0 V} T_{fi}^{(2)}. \quad (8.74)$$

Equation (8.73) now gives

$$\begin{aligned}\mathcal{T}_{fi}^{(2)} &= e^2 \sqrt{2p'_0} \bar{u}(p') \left[ (e' \cdot \gamma) \frac{-i(p+k) \cdot \gamma + m}{2p \cdot k} (e \cdot \gamma) \right. \\ &\quad \left. - (e \cdot \gamma) \frac{-i(p-k') \cdot \gamma + m}{2p \cdot k'} (e' \cdot \gamma) \right] u(p) \sqrt{2p_0}.\end{aligned} \quad (8.75)$$

It is clear from (3.195) that the Dirac spinor multiplied by  $\sqrt{2p_0}$  has the Casimir operator

$$\sum_n \sqrt{2p_0} u_r(p, h) \bar{u}_s(p, h) \sqrt{2p_0} = (-ip \cdot \gamma + m)_{rs}, \quad (8.76)$$

where we sum over helicities. Therefore, this provides a normalization that does not depend on how we choose the Lorentz system. We will compute the cross-section from (8.75) in the next chapter.

### 8.5.2 Pion Decay to Muons

The  $\pi^+$ -meson usually decays according to

$$\pi^+ \rightarrow \mu^+ + \nu. \quad (8.77)$$

We can take the following phenomenological expression for the interaction describing this decay:

$$\mathcal{H}_{\text{int}} = \frac{ig}{m_\pi} \bar{\psi}_\nu \gamma_\lambda (1 + \gamma_5) \psi_\mu \partial_\lambda \varphi + \text{h.c.}, \quad (8.78)$$

where  $m_\pi$  is the mass of  $\pi^+$ ,  $\psi_\nu$  and  $\psi_\mu$  are field operators for the neutrino and the  $\mu$ -meson, respectively, and h.c. indicates the Hermitian conjugate. If we take  $|\mu^+, \nu\rangle = a^\dagger(\mu^+)a^\dagger(\nu)|0\rangle$  as the final state, the lowest order of the decay amplitude has the form

$$T_{fi} = \langle \mu^+, \nu | \mathcal{H}_{\text{int}}(0) | \pi^+ \rangle. \quad (8.79)$$

Hence, denoting the four-momenta of  $\pi^+$ ,  $\mu^+$ , and  $\nu$  by  $P$ ,  $p$ , and  $q$ , respectively, we have  $P = p + q$  and

$$\begin{aligned} T_{fi} &= -\frac{ig}{m_\pi} \langle q | \bar{\psi}_\nu(0) | 0 \rangle \gamma_\lambda (1 + \gamma_5) \langle p | \psi_\mu(0) | 0 \rangle \langle 0 | \partial_\lambda \varphi(0) | \pi^+ \rangle \\ &= \frac{1}{\sqrt{2P_0V} \sqrt{V} \sqrt{V}} \frac{g}{m_\pi} \bar{u}(q) \gamma_\lambda (1 + \gamma_5) v(p) P_\lambda. \end{aligned} \quad (8.80)$$

In order to simplify this equation, we use

$$\bar{u}(q) i \not{q} \cdot \gamma = (i p \cdot \gamma - m_\mu) v(p) = 0.$$

Therefore, the invariant decay amplitude is

$$\begin{aligned} \mathcal{T}_{fi} &= \sqrt{2P_0V} \sqrt{2q_0V} \sqrt{2p_0V} T_{fi} \\ &= -ig \frac{m_\mu}{m_\pi} \sqrt{2q_0} \bar{u}(q) (1 - \gamma_5) v(p) \sqrt{2p_0}. \end{aligned} \quad (8.81)$$

## Two-Photon Decay of $\pi^0$

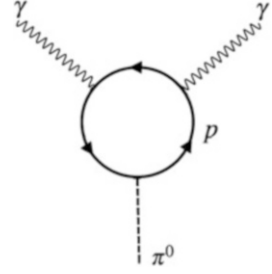
Although the two examples mentioned above are rather simple, they are complicated enough to provide a good illustration of this kind of computation. Another good example is the two-photon decay of the  $\pi^0$ -meson, which includes a fermion loop and no divergences, i.e.,

$$\pi^0 \rightarrow 2\gamma. \quad (8.82)$$

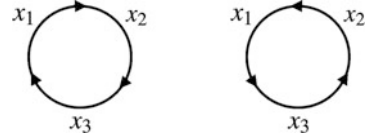
The interactions involved in this decay process are the Yukawa-type interaction between the  $\pi^0$  and the proton and the electromagnetic interaction with the proton. Assuming that the former has the form given in (6.97) and using  $\psi$  and  $\varphi$  to denote the field operators of the proton and  $\pi^0$ , respectively, we have

$$\mathcal{H}_{\text{int}} = iG \bar{\psi} \gamma_5 \psi \varphi - ie \bar{\psi} \gamma_\mu \psi A_\mu \equiv \mathcal{H}_\pi + \mathcal{H}_{\text{em}}. \quad (8.83)$$

**Fig. 8.10** Feynman diagram for the two-photon decay of  $\pi^0$



**Fig. 8.11** Feynman diagrams for the contraction of fermion operators



The above interactions are sufficient for the lowest order computation. The Feynman diagram is shown in Fig. 8.10.

Denoting the four-momenta of the  $\pi^0$  and the two photons by  $q$ ,  $k'$ , and  $q''$ , respectively, the  $S$ -matrix element corresponding to the above graph is given by

$$\begin{aligned} \langle k', k'' | S | q \rangle &= \frac{(-i)^3}{2!} \int d^4x_1 \int d^4x_2 \int d^4x_3 \langle k', k'' | T' [ \mathcal{H}_{em}(x_1) \mathcal{H}_{em}(x_2) \mathcal{H}_\pi(x_3) ] | q \rangle \\ &= \frac{1}{2} e^2 G \int d^4x_1 \int d^4x_2 \int d^4x_3 \\ &\quad \times \langle 0 | T [ \bar{\psi}(x_1) \gamma_\mu \psi(x_1), \bar{\psi}(x_2) \gamma_\nu \psi(x_2), \bar{\psi}(x_3) \gamma_5 \psi(x_3) ] | 0 \rangle \\ &\quad \times \langle k', k'' | : A_\mu(x_1) A_\nu(x_2) : | 0 \rangle \langle 0 | \varphi(x_3) | q \rangle . \end{aligned} \tag{8.84}$$

The reason why the coefficient  $1/3!$  has been replaced by  $1/2!$  is that there are three ways of choosing the coordinate of  $\mathcal{H}_\pi$ , i.e.,  $x_1$ ,  $x_2$ , or  $x_3$ , and each of them gives the same contribution, while we have kept only one, i.e.,  $x_3$ , and multiplied the whole expression by three. Moreover, similarly to the case with Compton scattering, the external photon lines give two terms, but since each gives the same contribution, we keep only one and drop the factor of  $1/2$ . As a consequence,

$$\begin{aligned} \langle k', k'' | S | q \rangle &= e^2 G \int d^4x_1 \int d^4x_2 \int d^4x_3 \\ &\quad \times \langle 0 | T [ \bar{\psi}(x_1) \gamma_\mu \psi(x_1), \bar{\psi}(x_2) \gamma_\nu \psi(x_2), \bar{\psi}(x_3) \gamma_5 \psi(x_3) ] | 0 \rangle \\ &\quad \times \langle k' | A_\mu(x_1) | 0 \rangle \langle k'' | A_\nu(x_2) | 0 \rangle \langle 0 | \varphi(x_3) | q \rangle . \end{aligned} \tag{8.85}$$

Note also that there are two ways of contracting fermion operators, as shown in Fig. 8.11. The difference between these two diagrams is that the directions of the

fermion loops are opposite:

$$\begin{aligned} \langle 0|T[\dots]|0\rangle &= -\text{Tr}[\gamma_5 S_F(x_3 - x_2)\gamma_\nu S_F(x_2 - x_1)\gamma_\mu S_F(x_1 - x_3)] \\ &= -\text{Tr}[\gamma_5 S_F(x_3 - x_1)\gamma_\mu S_F(x_1 - x_2)\gamma_\nu S_F(x_2 - x_3)]. \end{aligned} \quad (8.86)$$

These two terms are in fact equivalent to each other. In order to show this, we make use of the relation

$$C^{-1}S_F(x - y)C = S_F(y - x)^T. \quad (8.87)$$

This follows because

$$\begin{aligned} C^{-1}S_F(x - y)C &= C^{-1}\left(-\gamma_\mu \frac{\partial}{\partial x_\mu} + m\right)C \Delta_F(x - y) \\ &= \left(\gamma_\mu^T \frac{\partial}{\partial x_\mu} + m\right)\Delta_F(y - x) \\ &= \left(-\gamma_\mu^T \frac{\partial}{\partial y_\mu} + m\right)\Delta_F(y - x) = S_F(y - x)^T. \end{aligned}$$

We also use the following relations:

$$C^{-1}\gamma_\mu C = -\gamma_\mu^T, \quad C^{-1}\gamma_5 C = \gamma_5^T, \quad \text{Tr} A^T = \text{Tr} A.$$

Therefore,

$$\begin{aligned} &\text{Tr}[\gamma_5 S_F(x_3 - x_2)\gamma_\nu S_F(x_2 - x_1)\gamma_\mu S_F(x_1 - x_3)]^T \\ &= \text{Tr}[S_F(x_1 - x_3)^T \gamma_\mu^T S_F(x_2 - x_1)^T \gamma_\nu^T S_F(x_3 - x_2)^T \gamma_5^T] \\ &= \text{Tr}[C^{-1}S_F(x_3 - x_1)C \cdot C^{-1}(-\gamma_\mu)C \cdot C^{-1}S_F(x_1 - x_2)C \\ &\quad \times C^{-1}(-\gamma_\nu)C \cdot C^{-1}S_F(x_2 - x_3)C \cdot C^{-1}\gamma_5 C] \\ &= \text{Tr}[\gamma_5 S_F(x_3 - x_1)\gamma_\mu S_F(x_1 - x_2)\gamma_\nu S_F(x_2 - x_3)]. \end{aligned} \quad (8.88)$$

i.e., the first trace is equivalent to the second trace. It is clear from the above example that, if there are an odd number of photons, they will cancel each other by flipping signs. Now,

$$\begin{aligned} \langle k', k''|S|q\rangle &= -2e^2 G \int d^4x_1 \int d^4x_2 \int d^4x_3 \langle k'|A_\mu(x_1)|0\rangle \langle k''|A_\nu(x_2)|0\rangle \langle 0|\varphi(x_3)|q\rangle \\ &\quad \times \text{Tr}[\gamma_5 S_F(x_3 - x_1)\gamma_\mu S_F(x_1 - x_2)\gamma_\nu S_F(x_2 - x_3)]. \end{aligned} \quad (8.89)$$

In the following, we write the polarization vectors of the two photons as  $e'$  and  $e''$ . Then, inserting the wave function of one particle and the Fourier representation

of  $S_F$ , and integrating over  $x_1$ ,  $x_2$ , and  $x_3$ , a factor  $(2\pi)^{4 \times 3}$  appears. However, this factor cancels the one arising from the Fourier representation of  $S_F$ . At the same time, three  $\delta$ -functions of the four-momenta appear, along with three four-dimensional momentum integrals arising from the Fourier representation of  $S_F$ . Two of the momentum integrals cancel two of the  $\delta$ -functions, leaving just one four-dimensional momentum integral and one  $\delta$ -function expressing the conservation of the total energy–momentum. Hence, computing  $T_{fi}$ , except for this  $\delta$ -function, and extracting the invariant decay amplitude  $\mathcal{T}_{fi}$ , we obtain

$$\mathcal{T}_{fi} = \frac{2e^2 G}{(2\pi)^4} e'_\mu e''_\nu \int d^4 p \quad (8.90)$$

$$\text{Tr} \left[ \gamma_5 \frac{-i(p-k') \cdot \gamma + m}{(p-k')^2 + m^2 - i\epsilon} \gamma_\mu \frac{-ip \cdot \gamma + m}{p^2 + m^2 - i\epsilon} \gamma_\nu \frac{-i(p+k'') \cdot \gamma + m}{(p+k'')^2 + m^2 - i\epsilon} \right].$$

First we compute the trace in the numerator. The details of this calculation can be found in my previous book *Relativistic Quantum Mechanics* [78]. Here we give only the relevant formulas:

$$\text{Tr}(\mathbf{1}) = 4, \quad \text{Tr}(\gamma_\mu \gamma_\nu) = 4\delta_{\mu\nu}. \quad (8.91)$$

The trace of a product of an odd number of  $\gamma$ -matrices is zero. To find the trace of an even number of  $\gamma$ -matrices, we just do the same calculation as when we contracted all the fermion fields in the Wick contraction. For example,

$$\begin{aligned} \text{Tr}(\gamma_\sigma \gamma_\lambda \gamma_\mu \gamma_\nu) &\longrightarrow \overbrace{(\gamma_\sigma \gamma_\lambda \gamma_\mu \gamma_\nu)} + \overbrace{(\gamma_\sigma \gamma_\lambda \gamma_\mu \gamma_\nu)} + \overbrace{(\gamma_\sigma \gamma_\lambda \gamma_\mu \gamma_\nu)} \\ &\longrightarrow 4(\delta_{\sigma\lambda} \delta_{\mu\nu} + \delta_{\sigma\nu} \delta_{\lambda\mu} - \delta_{\sigma\mu} \delta_{\lambda\nu}). \end{aligned}$$

Clearly, we replace contracted pairs by Kronecker deltas. Signs are determined as though the pairs of  $\gamma$ -matrices to be contracted are next to each other, and when reordering them, treating them as anti-commutative. Since  $\gamma_5 = \gamma_1 \gamma_2 \gamma_3 \gamma_4$ , we have

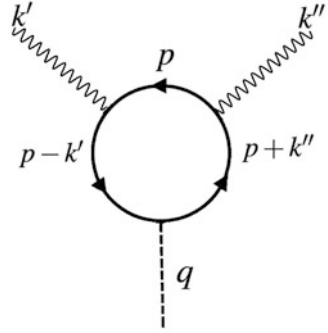
$$\text{Tr}(\gamma_5 \gamma_\lambda \gamma_\mu \gamma_\nu \gamma_\sigma) = 4\epsilon_{\lambda\mu\nu\sigma}, \quad (8.92)$$

which implies

$$\begin{aligned} \text{Tr}[\gamma_5 (-i(p-k') \cdot \gamma + m) \gamma_\mu (-ip \cdot \gamma + m) \gamma_\nu (-i(p+k'') \cdot \gamma + m)] \\ = 4m\epsilon_{\mu\nu\alpha\beta} [p_\alpha (p+k'')_\beta - (p-k')_\alpha (p+k'')_\beta + (p-k')_\alpha p_\beta] \\ = 4m\epsilon_{\mu\nu\alpha\beta} k'_\alpha k''_\beta. \end{aligned} \quad (8.93)$$

The Feynman diagram corresponding to this process is shown in Fig. 8.12 for the momentum representation. As is clear from this diagram, the conservation of four-momentum holds at each vertex, just as the conservation of electric current holds in

**Fig. 8.12** Momentum space Feynman diagram for the two-photon decay of  $\pi^0$



an electric circuit. This is because, corresponding to each vertex,  $\delta$ -functions of the momenta appear as a consequence of the four-dimensional coordinate integrals.

Since the trace above does not depend on  $p$ , we take it outside the integral:

$$\mathcal{T}_{fi} = \frac{8me^2G}{(2\pi)^4} \epsilon_{\mu\nu\alpha\beta} e'_\mu e''_\nu k'_\alpha k''_\beta \quad (8.94)$$

$$\int \frac{d^4p}{[(p-k')^2 + m^2][p^2 + m^2 - i\epsilon][(p+k'')^2 + m^2 - i\epsilon]}.$$

To carry out the integral over  $p$ , we use Feynman's formula:

$$\int \frac{d^4p}{(p^2 + \Lambda - i\epsilon)^3} = \frac{i\pi^2}{2(\Lambda - i\epsilon)}. \quad (8.95)$$

This is a fundamental formula in the Feynman–Dyson theory. We give a proof of this. First, note that

$$\int_0^\infty dx e^{-i\alpha x} = \lim_{\epsilon \rightarrow +0} \int_0^\infty dx e^{-i\alpha x - \epsilon x} = \lim_{\epsilon \rightarrow +0} \frac{-i}{\alpha - i\epsilon}.$$

In the following, we omit reference to the limit. From this,

$$\int_0^\infty dx \exp[-i(p^2 + \Lambda)x] = \frac{-i}{p^2 + \Lambda - i\epsilon}. \quad (8.96)$$

Differentiating this equation twice with respect to  $\Lambda$  yields

$$\int_0^\infty dx x^2 \exp[-i(p^2 + \Lambda)x] = \frac{2i}{(p^2 + \Lambda - i\epsilon)^3}. \quad (8.97)$$

Moreover, from the formula for the Gaussian integral, if  $x > 0$ , then

$$\begin{aligned} \int d^4 p \exp(-ip^2 x) &= \left[ \prod_{j=1}^3 \int dp_j \exp(-ip_j^2 x) \right] \int dp_0 \exp(ip_0^2 x) \\ &= -\frac{i\pi^2}{x^2}. \end{aligned}$$

Thus,

$$\begin{aligned} \int d^4 p \int_0^\infty dx x^2 \exp[-i(p^2 + \Lambda)x] &= -i\pi^2 \int_0^\infty dx \exp(-i\Lambda x) \\ &= -\frac{\pi^2}{\Lambda - i\epsilon}. \end{aligned}$$

However, since the left-hand side has been obtained by integrating the left-hand side of (8.97) over  $p$ , we can use the latter equation to deduce (8.95).

To apply this formula to (8.94), we need another formula, viz.,

$$\frac{1}{A_1 A_2 \dots A_n} = (n-1)! \int_0^1 dx_1 \dots \int_0^1 dx_n \delta(1 - \sum x_i) \frac{1}{(x_1 A_1 + x_2 A_2 + \dots + x_n A_n)^n}. \tag{8.98}$$

A condition for this formula to hold is that the denominator on the right-hand side should not vanish anywhere within the domain of integration. The following are variants of the above formula:

$$\frac{1}{AB} = \int_0^1 \frac{dx}{[xA + (1-x)B]^2}, \quad \frac{1}{ABC} = 2 \int_0^1 dx \int_0^1 dy \frac{x}{[A(1-x) + Bxy + Cx(1-y)]^3}.$$

We shall use the last of these with

$$\begin{aligned} A &= (p - k')^2 + m^2 - i\epsilon = p^2 - 2p \cdot k' + m^2 - i\epsilon, \\ B &= p^2 + m^2 - i\epsilon, \\ C &= (p + k'')^2 + m^2 - i\epsilon = p^2 + 2p \cdot k'' + m^2 - i\epsilon, \end{aligned}$$

where we have used  $k'^2 = k''^2 = 0$ . Therefore, the  $p$ -integral is

$$\begin{aligned} I &= \int d^4 p \int_0^1 2x dx \int_0^1 dy \frac{1}{[p^2 + m^2 + (1-x)(-2p \cdot k') + x(1-y)(2p \cdot k'') - i\epsilon]^3} \\ &= \int d^4 p' \int_0^1 2x dx \int_0^1 dy \frac{1}{[p'^2 + m^2 + 2x(1-x)(1-y)k' \cdot k'' - i\epsilon]^3}, \end{aligned} \tag{8.99}$$



where we have eliminated the linear terms in  $p$  by making the replacement

$$p' = p - (1-x)k' + x(1-y)k'' . \quad (8.100)$$

We carry out the  $p'$ -integral in (8.99) using Feynman's formula (8.101):

$$I = i\pi^2 \int_0^1 x \, dx \int_0^1 dy \frac{1}{m^2 + 2x(1-x)(1-y)k'k''} . \quad (8.101)$$

Since  $q^2 = (k' + k'')^2 = -\mu^2$ , it follows that  $2k' \cdot k'' = -\mu^2$ , whence

$$I = i\pi^2 \int_0^1 x \, dx \int_0^1 dy \frac{1}{m^2 - \mu^2 x(1-x)(1-y)} .$$

Here  $m$  is the proton mass and  $\mu$  is the mass of  $\pi^0$ . Given that the value of  $4m^2$  is about 200 times the value of  $\mu^2$ ,

$$\frac{\mu^2}{m^2} x(1-x)(1-y) \leq \frac{\mu^2}{4m^2} \ll 1 .$$

Therefore,

$$I = \frac{i\pi^2}{m^2} \int_0^1 x \, dx \int_0^1 dy \left[ 1 + \mathcal{O}\left(\frac{\mu^2}{4m^2}\right) \right] \approx \frac{i\pi^2}{2m^2} . \quad (8.102)$$

Inserting this result into (8.94),

$$\mathcal{T}_{fi} = \frac{1}{\pi} \left( \frac{e^2}{4\pi} \right) \frac{G}{m} (i\epsilon_{\mu\nu\alpha\beta} e'_\mu e''_\nu k'_\alpha k''_\beta) . \quad (8.103)$$

Note that, since the last factor includes  $i$ , this is in fact real. In the next chapter, we will use it to compute the lifetime of  $\pi^0$ .

## 8.6 Furry's Theorem

As shown in the loop calculation for  $\pi^0 \rightarrow 2\gamma$ , if a contribution from a reversed loop has the opposite sign, the two loops cancel each other out, which tells us that the process is forbidden. Let us consider how we could derive this rule without using perturbation theory. QED is invariant under charge conjugation:

$$\mathcal{C}^{-1} \mathcal{H}_{\text{int}}(x) \mathcal{C} = \mathcal{H}_{\text{int}}(x) . \quad (8.104)$$

According to the discussion in Sect. 3.9, since the current  $j_\mu$  changes sign under charge conjugation, (8.104) implies that  $A_\mu$  also changes its sign:

$$\mathcal{C}^{-1} A_\mu(x) \mathcal{C} = -A_\mu(x) . \quad (8.105)$$

Now, if (8.104) holds, the  $S$ -matrix is also invariant under charge conjugation. Dyson's formula for the  $S$ -matrix can be written in the form

$$S = T \exp \left[ -i \int d^4x \mathcal{H}_{\text{int}}(x) \right] . \quad (8.106)$$

This tells us to time-order the Hamiltonian density produced by expanding the exponential in a power series, and then carry out the space-time integral. Thus,

$$\mathcal{C}^{-1} S \mathcal{C} = T \exp \left[ -i \int d^4x \mathcal{C}^{-1} \mathcal{H}_{\text{int}}(x) \mathcal{C} \right] = S . \quad (8.107)$$

Hence, considering  $|a\rangle$  and  $|b\rangle$  as eigenstates of  $\mathcal{C}$  and setting

$$\mathcal{C}|a\rangle = \epsilon_a|a\rangle , \quad \mathcal{C}|b\rangle = \epsilon_b|b\rangle , \quad (8.108)$$

the transition  $a \rightarrow b$  will be forbidden whenever  $\epsilon_a \neq \epsilon_b$ . In the  $n$ -photon state,

$$\mathcal{C}|n\gamma\rangle = (-1)^n |n\gamma\rangle . \quad (8.109)$$

Therefore, the transition from odd numbers of photon states to even numbers of photon states is prohibited.

Furthermore, from the charge-conjugation invariance of the Hamiltonian density (8.83), for  $\pi^0$ ,

$$\mathcal{C}^{-1} \varphi(x) \mathcal{C} = \varphi(x) . \quad (8.110)$$

Thus, any process whereby  $\pi^0$  decays into odd numbers of photons is forbidden. This series of selection rules based on charge conjugation is called *Furry's theorem* [98–100].

## 8.7 Two-Photon Decays of Neutral Mesons

In the last section, we discussed a selection rule based on charge conjugation. Let us consider other grounds for such rules, e.g., a selection rule based on the conservation of angular momentum. In the following, we discuss the process whereby a neutral meson  $M^0$  decays into two photons:

$$M^0 \rightarrow 2\gamma . \quad (8.111)$$

It is clear from the discussion in the last section that the invariant decay amplitude  $\mathcal{T}$  can be written in the following ways depending on whether  $M^0$  has spin zero or one:

$$\mathcal{T}(M^0 \rightarrow 2\gamma) \propto \begin{cases} c_{\mu\nu}(k, k') : A_\mu(k) A_\nu(k') \varphi(k + k') : & \text{spin } 0 , \\ c_{\lambda\mu\nu}(k, k') : A_\mu(k) A_\nu(k') \varphi_\lambda(k + k') : & \text{spin } 1 . \end{cases} \quad (8.112)$$

Normally, the right-hand side consists of operators, but we have taken suitable matrix elements. Note also that Fourier-transformed field operators are included in the normal product. By (6.66), the normal product above should commute with  $\partial_\mu A_\mu(x)$ . Looking back to the discussion in Sect. 5.4, this implies that, in the normal product in  $\mathcal{T}$ , electromagnetic fields only appear in the following combinations:

$$F_{\mu\nu}(k) , \quad k_\mu A_\mu(k) . \quad (8.113)$$

Taking into account the Lorenz condition, we can set the latter to zero. If  $\mathcal{T}$  includes only terms of the form in (8.113), then  $\mathcal{T}$  should be invariant under the replacement

$$A_\mu(k) \rightarrow A_\mu(k) + \lambda k_\mu . \quad (8.114)$$

This is the requirement of gauge invariance. It turns out to be related to the spin property of massless particles discussed in depth in Sect. 8.5.

#### Case with $0^+$

In this case,  $c_{\mu\nu}$  must be a tensor built up from  $k$  and  $k'$ , i.e., it must be a linear combination of the following:

$$\delta_{\mu\nu} , \quad k_\mu k_\nu , \quad k'_\mu k'_\nu , \quad k_\mu k'_\nu , \quad k'_\mu k_\nu .$$

Using the Lorenz condition in (8.112), any term involving  $k_\mu$  or  $k'_\nu$  in  $c_{\mu\nu}$  has to vanish, so the only possible form is

$$c_{\mu\nu} = a\delta_{\mu\nu} + bk'_\mu k_\nu . \quad (8.115)$$

Hence, taking into account the invariance under (8.114),

$$c_{\mu\nu} k_\mu = c_{\mu\nu} k'_\nu = 0 , \quad \text{i.e., } a + b k \cdot k' = 0 . \quad (8.116)$$

Therefore, the only possible form of  $c_{\mu\nu}$  is

$$c_{\mu\nu} = -b[(k \cdot k')\delta_{\mu\nu} - k'_\mu k_\nu] , \quad (8.117)$$

and the corresponding form of  $\mathcal{T}$  is

$$\begin{aligned}\mathcal{T}(M^0 \rightarrow 2\gamma) &\propto [(k \cdot k')\delta_{\mu\nu} - k'_\mu k_\nu] : A_\mu(k)A_\nu(k')\varphi(k+k') : \\ &\propto : F_{\mu\nu}(k)F_{\mu\nu}(k')\varphi(k+k') :\end{aligned}\quad (8.118)$$

This means that the effective interaction of the two-photon decay of the neutral scalar meson is given by

$$F_{\mu\nu}F_{\mu\nu}\varphi. \quad (8.119)$$

### Case with $0^-$

In this case, since  $c_{\mu\nu}$  is a pseudo-tensor built up from  $k$  and  $k'$ ,

$$c_{\mu\nu} = b\epsilon_{\mu\nu\alpha\beta}k'_\alpha k_\beta. \quad (8.120)$$

Therefore, the only possible form of  $\mathcal{T}$  is

$$\mathcal{T}(M^0 \rightarrow 2\gamma) \propto : F_{\mu\nu}(k)\tilde{F}_{\mu\nu}(k')\varphi(k+k') : , \quad (8.121)$$

and the effective interaction is given by

$$F_{\mu\nu}\tilde{F}_{\mu\nu}\varphi, \quad (8.122)$$

where

$$\tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F_{\alpha\beta}. \quad (8.123)$$

### Case with $1^-$

In this case, there is also an irreducibility condition for the spin-1 field, viz.,

$$(k+k')_\lambda\varphi_\lambda(k+k') = 0. \quad (8.124)$$

Thus,  $k_\lambda$  and  $k'_\lambda$  are not independent, and  $k_\mu$  and  $k'_\nu$  cannot be used, so we must take  $c_{\lambda\mu\nu}$  to be a linear combination of the following:

$$(k_\lambda - k'_\lambda)\delta_{\mu\nu}, \quad (k_\lambda - k'_\lambda)k'_\mu k_\nu, \quad k'_\mu\delta_{\lambda\nu}, \quad k_\nu\delta_{\lambda\mu}.$$

However, the first two contradict the boson symmetry, and the last two are excluded by gauge invariance:

$$k_\mu c_{\lambda\mu\nu} = k'_\nu c_{\lambda\mu\nu} = 0. \quad (8.125)$$

Therefore, the  $1^-$ -neutral meson cannot decay into two photons. Likewise, the  $1^+$ -neutral meson cannot decay into two photons.

**Theorem 8.1** *The two-photon decay of a spin-1 neutral meson is forbidden.*

By a similar argument, we can also prove the following:

**Theorem 8.2** *The decay of a spin-0 particle into another spin-0 particle and one photon is forbidden.*

### Two-Photon Polarization

For simplicity, if we consider the Coulomb gauge,

$$\mathbf{A} \parallel \mathbf{e}, \quad \mathbf{E} \parallel \mathbf{e}, \quad \mathbf{H} \parallel \mathbf{k} \times \mathbf{e}, \quad (8.126)$$

$$\text{decay amplitude of } 0^+ \propto \mathbf{e}_1 \cdot \mathbf{e}_2, \quad (8.127)$$

$$\text{decay amplitude of } 0^- \propto (\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{k}. \quad (8.128)$$

Hence, choosing the direction of  $\mathbf{k}_1$  to be the  $z$ -axis, since  $\mathbf{k}_2 = -\mathbf{k}_1$ , the above amplitudes can be expressed in terms of the right polarization  $R$  and the left polarization  $L$ :

$$\begin{aligned} \mathbf{e}_1 \cdot \mathbf{e}_2 &= \frac{e_{1x} + ie_{1y}}{\sqrt{2}} \cdot \frac{e_{2x} - ie_{2y}}{\sqrt{2}} + \frac{e_{1x} - ie_{1y}}{\sqrt{2}} \cdot \frac{e_{2x} + ie_{2y}}{\sqrt{2}} \\ &\propto \Phi^{\text{RR}} + \Phi^{\text{LL}}, \end{aligned} \quad (8.129)$$

$$(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{k}_1 \propto \Phi^{\text{RR}} - \Phi^{\text{LL}}. \quad (8.130)$$

From these, the transformation property under parity  $\mathcal{P}$  is

$$\mathcal{P}(\Phi^{\text{RR}} \pm \Phi^{\text{LL}}) = \Phi^{\text{LL}} \pm \Phi^{\text{RR}} = \pm(\Phi^{\text{RR}} \pm \Phi^{\text{LL}}). \quad (8.131)$$

### Case with Higher Spin

We consider the case with spin greater than 2. We write a field operator whose spin is greater than 2 as follows:

$$\varphi_{\lambda\mu\dots} \quad (8.132)$$

This should satisfy the following irreducibility conditions:

- Symmetry. The above  $\varphi$  is symmetric under the interchange of any two indices:

$$\varphi_{\dots\lambda\dots\mu\dots} = \varphi_{\dots\mu\dots\lambda\dots} \quad (8.133)$$

- Rule for contractions. The contraction over any pair of indices vanishes:

$$\varphi_{\dots\lambda\dots\lambda\dots} = 0 . \quad (8.134)$$

- Rule for divergence. The divergence with respect to an arbitrary index vanishes:

$$\partial_\lambda \varphi_{\dots\lambda\dots} = 0 . \quad (8.135)$$

We thus construct scalar effective interactions using two  $F_{\mu\nu}$ 's and  $\varphi_{\mu\nu\dots}$ . In this case, we take into account the following properties:

- $F_{\mu\nu}$  and  $\tilde{F}_{\mu\nu}$  are anti-symmetric under the interchange of  $\mu$  and  $\nu$ .
- $\partial_\mu F_{\mu\nu} = \partial_\mu \tilde{F}_{\mu\nu} = 0$ .

Considering first the spin-2 case, there are two possibilities for forming scalars:

$$\text{even parity } (\partial_\alpha \partial_\beta F_{\mu\nu}) F_{\mu\nu} \varphi_{\alpha\beta} , \quad F_{\mu\alpha} F_{\mu\beta} \varphi_{\alpha\beta} , \quad (8.136)$$

$$\text{odd parity } (\partial_\alpha \partial_\beta F_{\mu\nu}) \tilde{F}_{\mu\nu} \varphi_{\alpha\beta} , \quad F_{\mu\alpha} \tilde{F}_{\mu\beta} \varphi_{\alpha\beta} . \quad (8.137)$$

These can be generalized:

$$(\partial_\alpha \partial_\beta \dots F_{\mu\nu}) F_{\mu\nu} \varphi_{\alpha\beta\dots} , \quad (8.138)$$

$$(\partial_\alpha \partial_\beta \dots F_{\mu\rho}) F_{\nu\rho} \varphi_{\mu\nu\alpha\beta\dots} . \quad (8.139)$$

Here, both spin and parity are even or odd. However, when both are odd, carrying out the integration by parts to construct the  $S$ -matrix, this contradicts the boson symmetry.<sup>1</sup> We thus obtain the following:

**Theorem 8.3** *The two-photon decay of a neutral meson whose spin and parity are both odd is forbidden.*

If the spin is even and the parity is odd,

$$(\partial_{\alpha\beta} \dots F_{\mu\nu}) \tilde{F}_{\mu\nu} \varphi_{\alpha\beta\dots} . \quad (8.142)$$

---

<sup>1</sup> Due to (8.118) and (8.121), using the momentum representation, (8.138) and (8.139) can be written in the form

$$\mathcal{T}(M^0 \rightarrow 2\gamma) \propto (k - k')_\alpha (k - k')_\beta \dots : F_{\mu\nu}(k) F_{\mu\nu}(k') \varphi_{\alpha\beta\dots}(k + k') : , \quad (8.140)$$

$$\mathcal{T}(M^0 \rightarrow 2\gamma) \propto (k - k')_\alpha (k - k')_\beta \dots : F_{\mu\rho}(k) F_{\nu\rho}(k') \varphi_{\alpha\beta\dots}(k + k') : . \quad (8.141)$$

It is now straightforward to check the boson symmetry.

If the spin is odd and the parity is even,

$$(\partial_\alpha \partial_\beta \dots F_{\mu\rho}) \tilde{F}_{\nu\rho} \varphi_{\mu\nu\alpha\beta\dots} . \quad (8.143)$$

If both the spin and the parity are even, then general two-photon states are

$$\phi^{\text{RR}} + \phi^{\text{LL}} , \quad \phi^{\text{RL}} , \quad \phi^{\text{LR}} . \quad (8.144)$$

If the spin is even and the parity is odd, the system decays into the polarization state

$$\phi^{\text{RR}} - \phi^{\text{LL}} . \quad (8.145)$$

Finally, if the spin is odd and the parity is even, it decays into the polarization states

$$\phi^{\text{RL}} , \quad \phi^{\text{LR}} . \quad (8.146)$$

The above results were derived by Landau and Yang [[101](#), [102](#)] on the basis of angular momentum considerations.<sup>2</sup>

---

<sup>2</sup> Nowadays, this selection rule is called the Landau–Yang theorem.

# Chapter 9

## Cross-Sections and Decay Widths



In the last chapter, we discussed ways to evaluate  $S$ -matrix elements. However, in order to compare theory with experiment, we must also compute cross-sections and decay widths. We shall now focus on this aspect.

### 9.1 Møller's Formulas

We start with the transition probability per unit time:

$$w_{fi} = \lim_{\substack{t_1 \rightarrow -\infty \\ t_2 \rightarrow \infty}} \frac{|\langle f|U(t_2, t_1) - 1|i\rangle|^2}{t_2 - t_1} . \tag{9.1}$$

Although this ratio is in general a complicated function of  $t_1$  and  $t_2$ , it becomes a simple one in the above limit. Since  $U(t_2, t_1)$  becomes the  $S$ -matrix in this limit, using

$$\langle f|S|i\rangle = \delta_{fi} - i(2\pi)^4 \delta^4(P_f - P_i) \langle f|T|i\rangle , \tag{9.2}$$

we have

$$|\langle f|S - 1|i\rangle|^2 = (2\pi)^4 \delta^4(P_f - P_i) |\langle f|T|i\rangle|^2 \times (2\pi)^4 \delta^4(0) . \tag{9.3}$$

If we take  $V$  and  $t$  to be a quantisation volume and a macroscopic time, respectively, then

$$(2\pi)^4 \delta^4(0) = \int d^4x e^{i0 \cdot x} = \int d^4x \sim Vt ,$$

$$t_2 - t_1 \sim t .$$



Hence, Eq. (9.1) becomes

$$w_{fi} = (2\pi)^4 \delta^4(P_f - P_i) |\langle f|T|i \rangle|^2 V . \quad (9.4)$$

In addition, the expression for the cross-section is

$$\sigma_{fi} = \frac{V}{v_{\text{rel}}} w_{fi} = \frac{(2\pi)^4}{v_{\text{rel}}} \delta^4(P_f - P_i) |\langle f|T|i \rangle|^2 V^2 . \quad (9.5)$$

As for a true cross-section, we must sum over the final states. For instance, if there are  $n$  particles in the final state, summing over momenta is equivalent to introducing the integral

$$\sum_f \rightarrow \frac{V}{(2\pi)^3} \int d^3 p_1 \dots \frac{V}{(2\pi)^3} \int d^3 p_n . \quad (9.6)$$

If there is only one particle in the initial state, the width (probability) of its decay to  $n$  particles is

$$\begin{aligned} w(1 \rightarrow n) &= (2\pi)^4 \frac{V}{(2\pi)^3} \int d^3 p_1 \dots \frac{V}{(2\pi)^3} \int d^3 p_n |\langle f|T|i \rangle|^2 V \delta^4(P_f - P_i) \\ &= \frac{(2\pi)^{4-3n}}{2P_0} \int \frac{d^3 p_1}{2p_{10}} \dots \int \frac{d^3 p_n}{2p_{n0}} |\langle p_1, \dots, p_n | \mathcal{T} | P \rangle|^2 \delta^4(P_f - P_i) , \end{aligned} \quad (9.7)$$

where the invariant decay width  $\mathcal{T}$  has been defined by

$$\langle p_1, \dots, p_n | \mathcal{T} | P \rangle = \sqrt{2p_{10}V} \dots \sqrt{2p_{n0}V} \langle p_1, \dots, p_n | T | P \rangle \sqrt{2P_0V} . \quad (9.8)$$

This is just the same definition as (8.74).

The integral in (9.7) is Lorentz invariant, as is clear from the discussion in Chap. 4. This can be clearly seen from the relation

$$\int \frac{d^3 p}{2p_0} \dots = \int d^4 p \theta(p_0) \delta(p^2 + m^2) \dots \quad (9.9)$$

Thus, the decay width in (9.7) is Lorentz invariant except for the appearance of  $P_0$  in the denominator. The average lifetime  $\tau$  of the decaying particle is given by the reciprocal of the total decay width, i.e.,

$$\frac{1}{\tau} = \sum_n w(1 \rightarrow n) , \quad (9.10)$$

and  $\tau$  transforms as  $P_0$ . Thus, assuming that the lifetime in the rest frame is  $\tau_0$ , the lifetime in a moving frame is given by

$$\tau = \frac{\tau_0}{\sqrt{1 - \beta^2}}, \quad (9.11)$$

where  $\beta$  is the speed of the particle. Although  $w$  is the decay probability per unit time, it is equivalent to the decay width in natural units. In order to make this point clear, it can be denoted by  $\Gamma (1 \rightarrow n)$ .

If there are two particles in the initial state, we can speak of a scattering cross-section. Combining (9.5) with (9.6), the total cross-section for a reaction with  $n$  particles in the final state can be calculated from

$$\sigma(2 \rightarrow n) = \frac{(2\pi)^{4-3n}}{4v_{\text{rel}}P_{10}P_{20}} \int \frac{d^3p_1}{2p_{10}} \dots \int \frac{d^3p_n}{2p_{n0}} |\langle p_1, \dots, p_n | \mathcal{J} | P_1, P_2 \rangle|^2 \delta^4(P_f - P_i). \quad (9.12)$$

Then in the center-of-mass frame or the laboratory frame, the relative velocity  $v_{\text{rel}}$  can be written as

$$v_{\text{rel}} = \left| \frac{\mathbf{P}_1}{P_{10}} - \frac{\mathbf{P}_2}{P_{20}} \right|. \quad (9.13)$$

We can generalize  $v_{\text{rel}}P_{10}P_{20}$  to the Lorentz-invariant form

$$\begin{aligned} v_{\text{rel}}P_{10}P_{20} &= |\mathbf{P}_1 P_{20} - \mathbf{P}_2 P_{10}| \\ &\rightarrow \sqrt{(\mathbf{P}_1 P_{20} - \mathbf{P}_2 P_{10})^2 - (\mathbf{P}_1 \times \mathbf{P}_2)^2} \\ &= \sqrt{(P_1 \cdot P_2)^2 - M_1^2 M_2^2} \equiv B, \end{aligned} \quad (9.14)$$

where  $M_1$  and  $M_2$  are the masses of the two incident particles. This  $B$  is certainly a Lorentz-invariant quantity, and it coincides with  $v_{\text{rel}}P_{10}P_{20}$  in the center-of-mass frame or the laboratory frame. Therefore, the following cross-section is a Lorentz-invariant quantity:

$$\sigma(2 \rightarrow n) = \frac{(2\pi)^{4-3n}}{4B} \int \frac{d^3p_1}{2p_{10}} \dots \int \frac{d^3p_n}{2p_{n0}} |\langle p_1, \dots, p_n | \mathcal{J} | P_1, P_2 \rangle|^2 \delta^4(P_f - P_i). \quad (9.15)$$

We call this *Møller's formula* [103]. If the particles in the final state are of the same kind, then we must of course divide this equation by a suitable factor so as not to count the same state several times.

Moreover, if the initial wave is not polarized and we do not observe spins in the final states, then we have to take an average of the spin states of the initial particles,

and sum over the spin states of the particles in the final states. The same thing also holds true for the decay width:

$$|\mathcal{T}_{fi}|^2 \rightarrow \frac{1}{2S_i + 1} \sum_{S_i} \sum_{S_f} |\mathcal{T}_{fi}|^2 \quad (\text{decay width}) \quad (9.16)$$

$$|\mathcal{T}_{fi}|^2 \rightarrow \frac{1}{(2S_1 + 1)(2S_2 + 1)} \sum_{S_i} \sum_{S_f} |\mathcal{T}_{fi}|^2 \quad (\text{cross-section}), \quad (9.17)$$

where  $S_i$  stands for the spin of the decaying particle in (9.16),  $S_1$  and  $S_2$  stand for the spins of the two incident particles in (9.17), and the summations over  $S_i$  and  $S_f$  are taken over the spin states in the initial and final states, respectively.

We often have a situation where there are two particles in the final state. Let us consider the integral in that case:

$$\begin{aligned} I &= \int \frac{d^3 p_1}{2p_{10}} \int \frac{d^3 p_2}{2p_{20}} \delta^4(P_f - P_i) \dots \\ &= \int d^4 p_1 \int d^4 p_2 \theta(p_{10}) \delta(p_1^2 + m_1^2) \theta(p_{20}) \delta(p_2^2 + m_2^2) \delta^4(P_f - P_i) \dots \end{aligned} \quad (9.18)$$

To carry out this integral, we introduce the change of variables

$$p_1 = \frac{P_f}{2} + \Delta, \quad p_2 = \frac{P_f}{2} - \Delta, \quad d^4 p_1 d^4 p_2 = d^4 P_f d^4 \Delta. \quad (9.19)$$

Hence,

$$\begin{aligned} I &= \int d^4 \Delta \delta \left[ \left( \frac{P_i}{2} + \Delta \right)^2 + m_1^2 \right] \delta \left[ \left( \frac{P_i}{2} - \Delta \right)^2 + m_2^2 \right] \dots \\ &= \frac{1}{2} \int d^4 \Delta \delta \left( \Delta^2 + \frac{P_i^2}{4} + \frac{m_1^2 + m_2^2}{2} \right) \delta \left( P_i \cdot \Delta + \frac{m_1^2 - m_2^2}{2} \right) \dots \end{aligned}$$

To evaluate this, we first go to the center-of-mass frame, with

$$P_{i0} = W = \sqrt{s} \quad (\text{total energy of the system}). \quad (9.20)$$

Therefore, since  $\mathbf{P}_i = 0$ ,  $\Delta_0$  can be obtained as

$$\Delta_0 = \frac{m_1^2 - m_2^2}{2W} = \frac{m_1^2 - m_2^2}{2\sqrt{s}}. \quad (9.21)$$

From the first  $\delta$ -function, we obtain

$$\vec{\Delta}^2 = \frac{(m_1^2 - m_2^2)^2}{4s} + \frac{s}{4} - \frac{m_1^2 + m_2^2}{2} = \frac{[s - (m_1 - m_2)^2][s - (m_1 + m_2)^2]}{4s}. \quad (9.22)$$

Thus, the integral itself is given by

$$\begin{aligned} I &= \frac{1}{2W} \int d^3\Delta \delta[\Delta^2 + \dots] \dots \\ &= \frac{1}{2\sqrt{s}} \int \Delta^2 d\Delta d\Omega \delta[\Delta^2 + \dots] \dots \\ &= \frac{\Delta_f}{4\sqrt{s}} \int d\Omega \dots, \end{aligned} \quad (9.23)$$

where  $\Delta_f$  is the absolute value of the momentum of the final state given by (9.22). Thus,

$$\int \frac{d^3 p_1}{2p_{10}} \int \frac{d^3 p_2}{2p_{20}} \delta^4(P_f - P_i) \dots = \frac{\sqrt{[s - (m_1 - m_2)^2][s - (m_1 + m_2)^2]}}{8s} \int d\Omega \dots \quad (9.24)$$

This yields the differential cross-section in the center-of-mass system:

$$\frac{d\sigma}{d\Omega} = \frac{1}{(4\pi)^2 B} \frac{\Delta_f}{4\sqrt{s}} |\langle p_1, p_2 | \mathcal{J} | P_1, P_2 \rangle|^2. \quad (9.25)$$

Incidentally, taking  $\Delta_i$  to be the absolute value of the relativistic momentum in the initial state,

$$\begin{aligned} B^2 &= \frac{1}{4}(s - M_1^2 - M_2^2)^2 - M_1^2 M_2^2 \\ &= \frac{1}{4}[s - (M_1 - M_2)^2][s - (M_1 + M_2)^2] = s\Delta_i^2. \end{aligned} \quad (9.26)$$

Therefore, the differential cross-section in the center-of-mass system becomes

$$\frac{d\sigma}{d\Omega} = \frac{1}{(8\pi W)^2} \frac{\Delta_f}{\Delta_i} |\langle p_1, p_2 | \mathcal{J} | P_1, P_2 \rangle|^2. \quad (9.27)$$

### Compton Scattering Cross-Section

Next, we consider Compton scattering in the *laboratory frame*. Letting  $P_1$  and  $P_2$  be the four-momenta of the incident photon and the stationary electron, respectively,

and  $p_1$  and  $p_2$  the four-momenta of the scattered photon and electron, we denote the space and time components by

$$P_1 = (\mathbf{P}_1, E_1^0), \quad P_2 = (\mathbf{0}, m), \quad p_1 = (\mathbf{p}_1, E_1), \quad p_2 = (\mathbf{p}_2, E_2). \quad (9.28)$$

Thus, the integral in phase space is

$$I = \int \frac{d^3 p_1}{2E_1} \int \frac{d^3 p_2}{2E_2} \delta^3(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{P}_1) \delta(E_1 + E_2 - E_1^0 - m) \dots \quad (9.29)$$

Integrating over  $p_2$  and taking  $p_2 = P_1 - p_1$ , we have

$$\begin{aligned} I &= \int \frac{d^3 p_1}{4E_1 E_2} \delta(E_1 + E_2 - E_1^0 - m) \dots \\ &= \int \frac{p_1^2 d p_1 d\Omega_1}{4E_1 E_2} \delta(E_1 + E_2 - E_1^0 - m) \dots, \end{aligned} \quad (9.30)$$

where  $p_1$  is the absolute value of  $\mathbf{p}_1$  and  $d\Omega_1$  is the differential solid angle of the momentum of the scattered photon. Since

$$E_2^2 = m^2 + (\mathbf{P}_1 - \mathbf{p}_1)^2 = m^2 + P_1^2 + p_1^2 - 2p_1 P_1 \cos \theta, \quad (9.31)$$

we obtain immediately

$$\frac{\partial E_1}{\partial p_1} = \frac{p_1}{E_1}, \quad \frac{\partial E_2}{\partial p_1} = \frac{p_1 - P_1 \cos \theta}{E_2}. \quad (9.32)$$

Carrying out the integration over  $p_1$ ,

$$\begin{aligned} I &= \int d\Omega_1 \frac{p_1^2}{4E_1 E_2 \left( \frac{p_1}{E_1} + \frac{p_1}{E_2} - \frac{P_1}{E_2} \cos \theta \right)} \dots \\ &= \frac{p_1^2}{4} \int d\Omega_1 \frac{1}{p_1(E_1 + E_2) - P_1 E_1 \cos \theta} \dots \end{aligned} \quad (9.33)$$

Introducing the new notation  $p_1 = E_1 = \omega'$ ,  $P_1 = E_1^0 = \omega$  and  $E_1 + E_2 = \omega + m$ , we can rewrite the denominator of the equation above using the Compton formula, to give

$$\omega'(\omega + m) - \omega\omega' \cos \theta = \omega'[m + \omega(1 - \cos \theta)] = \omega m. \quad (9.34)$$

Therefore, the above integral becomes

$$I = \frac{\omega'^2}{4\omega m} \int d\Omega_1 \dots \dots \tag{9.35}$$

Since in this problem  $B = \omega m$ , the differential cross-section for Compton scattering in the laboratory system is

$$\frac{d\sigma}{d\Omega_1} = \frac{1}{(8\pi m)^2} \left(\frac{\omega'}{\omega}\right)^2 |\langle f | \mathcal{J} | i \rangle|^2 . \tag{9.36}$$

**Two-Particle Decay Width**

Finally, applying (9.24) to the case where a particle with mass  $M$  decays into two particles with masses  $m_1$  and  $m_2$ , the decay width is given by

$$\Gamma(1 \rightarrow 2) = \frac{1}{16\pi M^3} \sqrt{[M^2 - (m_1 - m_2)^2][M^2 - (m_1 + m_2)^2]} |\mathcal{F}(i \rightarrow 2)|^2 . \tag{9.37}$$

In this case, of course, we must make the replacement indicated in (9.17). If the final state is a two-particle system, the integral in phase space can be carried out as above.

**9.2 Examples of Cross-Sections and Decay Widths**

Since we now know how to compute  $S$ -matrix elements and the formula for the cross-section, we can combine the two and derive some concrete consequences.

**Compton Scattering**

First, we derive an explicit expression for the cross-section in Compton scattering. We denote the four-momenta of the incident photon and the electron by  $k$  and  $p$ , respectively, and denote the scattered momenta by  $k'$  and  $p'$ . In the laboratory system,

$$\mathbf{p} = 0 , \quad p_0 = m . \tag{9.38}$$

Writing the polarization vectors of the photons before and after the scattering as  $e$  and  $e'$ , we can assume without loss of generality that

$$\mathbf{p} \cdot \mathbf{e} = \mathbf{p} \cdot \mathbf{e}' = 0 , \tag{9.39}$$

because it has been already shown in (5.90) that, for a transverse wave, we can choose  $e$  and  $e'$  such that they have only spatial components. Moreover, since we

can impose the Lorenz condition on the observed photons, we can assume that

$$k \cdot e = k' \cdot e' = 0. \quad (9.40)$$

Using these conditions, both  $(k \cdot \gamma)$  and  $(p \cdot \gamma)$  anticommute with  $(e \cdot \gamma)$  and  $(e' \cdot \gamma)$ . Now we can simplify (8.75) using

$$\begin{aligned} \bar{u}(p')(e' \cdot \gamma) [-i(p+k) \cdot \gamma + m] (e \cdot \gamma) u(p) \\ = \bar{u}(p')(e' \cdot \gamma) (e \cdot \gamma) [i(p+k) \cdot \gamma + m] u(p) \\ = \bar{u}(p')(e' \cdot \gamma) (e \cdot \gamma) (ik \cdot \gamma) u(p) \end{aligned} \quad (9.41)$$

and

$$\bar{u}(p')(e \cdot \gamma) [-i(p-k') \cdot \gamma + m] (e' \cdot \gamma) U(p) = -\bar{u}(p)(e \cdot \gamma) (e' \cdot \gamma) (ik' \cdot \gamma) u(p). \quad (9.42)$$

Therefore, the invariant scattering amplitude can be written as

$$\mathcal{T}_{fi} = e^2 \sqrt{2p'_0} \bar{u}(p') \left[ (e' \cdot \gamma) (e \cdot \gamma) \frac{ik \cdot \gamma}{2p \cdot k} + (e \cdot \gamma) (e' \cdot \gamma) \frac{ik' \cdot \gamma}{2p \cdot k'} \right] u(p) \sqrt{2p_0}. \quad (9.43)$$

Using (3.195), we take the average over the electron spins in the initial state and sum over the electron spins in the final states. Then, by (9.17), we find

$$\begin{aligned} |\mathcal{T}_{fi}|^2 \rightarrow \frac{1}{2} e^4 \text{Tr} \left\{ \left[ (e' \cdot \gamma) (e \cdot \gamma) \frac{ik \cdot \gamma}{2p \cdot k} + (e \cdot \gamma) (e' \cdot \gamma) \frac{ik' \cdot \gamma}{2p \cdot k'} \right] (-ip \cdot \gamma + m) \right. \\ \left. \times \left[ \frac{ik' \cdot \gamma}{2p \cdot k'} (e' \cdot \gamma) (e \cdot \gamma) + \frac{ik \cdot \gamma}{2p \cdot k} (e \cdot \gamma) (e' \cdot \gamma) \right] (-ip' \cdot \gamma + m) \right\}. \end{aligned} \quad (9.44)$$

Inserting this into (9.36), the differential cross-section in the laboratory system is

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{lab}} = \frac{e^4}{128\pi^2} \frac{\omega'^2}{m^2\omega^2} \text{Tr}[\dots]. \quad (9.45)$$

When computing the trace, it is useful to introduce the vector

$$a = \frac{k'}{2p \cdot k'} - \frac{k}{2p \cdot k}, \quad (9.46)$$

which is orthogonal to  $p$ . Then,

$$(e' \cdot \gamma)(e \cdot \gamma) \frac{ik \cdot \gamma}{2p \cdot k'} + (e \cdot \gamma)(e' \cdot \gamma) \frac{ik' \cdot \gamma}{2p \cdot k'} = (e \cdot \gamma)(e' \cdot \gamma)(ia \cdot \gamma) + 2(e \cdot e') \frac{ik \cdot \gamma}{2p \cdot k}.$$

The trace now reduces to

$$\text{Tr}[\dots] = 8(e \cdot e')^2 + 2 \frac{(\omega - \omega')^2}{\omega\omega'}. \quad (9.47)$$

Hence,

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{lab}} = \left( \frac{e^2}{4\pi} \right)^2 \frac{1}{4m^2} \left( \frac{\omega'}{\omega} \right)^2 \left[ 4(e \cdot e')^2 + \frac{(\omega - \omega')^2}{\omega\omega'} \right]. \quad (9.48)$$

This is valid when the polarizations of the incident photon and the scattered photon are both specified. Then averaging the polarizations of the incident wave and summing over those of the scattered wave, we obtain

$$(e \cdot e')^2 \rightarrow \frac{1}{2} \sum_{\lambda} \sum_{\lambda'} (e, e')^2 = \frac{1}{2} (1 + \cos^2 \theta). \quad (9.49)$$

The part without polarization vectors is simply doubled.

To derive (9.49), we argue as follows. If we choose the direction of  $k$  to be the  $z$ -direction, then we can choose the  $x$ - and the  $y$ -directions along the polarization vectors:

$$\sum_{\lambda} (\mathbf{a} \cdot \mathbf{e}^{(\lambda)})^2 = a_x^2 + a_y^2 = \mathbf{a}^2 - a_z^2 = \mathbf{a}^2 - \frac{(\mathbf{a} \cdot \mathbf{k})^2}{k^2}.$$

Using this equation twice, we obtain (9.49). Here,  $\theta$  is the scattering angle of the photon. The resulting cross-section is given finally by the Klein–Nishina formula [104]

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{lab}} = \frac{1}{2} \left( \frac{e^2}{4\pi} \right)^2 \frac{1}{[m + \omega(1 - \cos \theta)]^2} \left\{ \frac{\omega^2(1 - \cos \theta)^2}{m[m + \omega(1 - \cos \theta)]} + 1 + \cos^2 \theta \right\}, \quad (9.50)$$

where we have used (9.34) to express  $\omega'$  in terms of  $\omega$  and  $\cos \theta$ . In particular, if  $\omega \ll m$ , this reduces to the classical Thomson formula

$$\left. \frac{d\sigma}{d\Omega} \right|_{\text{lab}} = \left( \frac{e^2}{4\pi m} \right)^2 \frac{1 + \cos^2 \theta}{2} = r_0^2 \frac{1 + \cos^2 \theta}{2}, \quad (9.51)$$



where  $r_0$  is the classical electron radius given by

$$r_0 = \frac{e^2}{4\pi mc^2} = \frac{e^2}{4\pi \hbar c} \frac{\hbar}{mc} = 2.818 \times 10^{-13} \text{cm} \quad (\text{CGS units}). \quad (9.52)$$

The total cross-section in this case is called the *Thomson cross-section*, given by

$$\sigma = \frac{8\pi}{3} r_0^2 = \sigma_{\text{Thomson}} = 6.65 \times 10^{-25} \text{cm}^2. \quad (9.53)$$

Integrating (9.50) over the solid angle, we obtain the total cross-section as

$$\sigma = \frac{3}{4} \left\{ \frac{1+\gamma}{\gamma^3} \left[ \frac{2\gamma(1+\gamma)}{1+2\gamma} - \log(1+2\gamma) \right] + \frac{1}{2\gamma} \log(1+2\gamma) - \frac{1+3\gamma}{(1+2\gamma)^2} \right\} \sigma_{\text{Thomson}}, \quad (9.54)$$

where  $\gamma = \omega/m$ . Two approximations can be given:

$$\sigma = \begin{cases} \left( 1 - 2\gamma + \frac{26}{5}\gamma^2 + \dots \right) \sigma_{\text{Thomson}}, & \text{for } \gamma \ll 1, \\ \frac{3}{8\gamma} \left[ \log(2\gamma) + \frac{1}{2} \right] \sigma_{\text{Thomson}}, & \text{for } \gamma \gg 1. \end{cases} \quad (9.55)$$

### Pion Decay

Next, let us apply (9.37) to the decay of pions  $\pi$  into muons  $\mu$ . We use (8.81) to compute

$$|\mathcal{T}_{fi}|^2 \rightarrow \sum_{S_f} |\mathcal{T}_{fi}|^2, \quad (9.56)$$

where we have summed over spins in the final states:

$$\sum_{S_f} |\mathcal{T}_{fi}|^2 = g^2 \left( \frac{m_\mu}{m_\pi} \right)^2 \text{Tr}[(1 - \gamma_5)(-i\not{p} \cdot \gamma - m_\mu)(1 + \gamma_5)(-i\not{q} \cdot \gamma)]. \quad (9.57)$$

The trace here becomes

$$\text{Tr}[\dots] = -8\not{p} \cdot \not{q} = 4(m_\pi^2 - m_\mu^2). \quad (9.58)$$

Therefore, using (9.37), we obtain the decay width

$$\Gamma(\pi \rightarrow \mu + \nu) = \left( \frac{g^2}{4\pi} \right) m_\pi \left( \frac{m_\mu}{m_\pi} \right)^2 \left[ 1 - \left( \frac{m_\mu}{m_\pi} \right)^2 \right]. \quad (9.59)$$

Since its reciprocal is the lifetime  $\tau = 2.56 \times 10^{-8}$  s of  $\pi^+$ ,  $g^2$  must be given by

$$\frac{g^2}{4\pi} = 1.76 \times 10^{-15}. \quad (9.60)$$

We also assume the interaction (8.78) for  $\pi^+ \rightarrow e^+ + \nu$ , with the same value of  $g^2$ . This is called the *universality assumption*. We obtain the branching ratio

$$\frac{\Gamma(\pi^+ \rightarrow e^+ + \nu)}{\Gamma(\pi \rightarrow \mu^+ + \nu)} = \left(\frac{m_e}{m_\mu}\right)^2 \frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2} = 1.29 \times 10^{-4}. \quad (9.61)$$

This ratio is confirmed by experiment, and will later provide an important insight when investigating the structure of the interaction.

Lastly, we shall consider the decay  $\pi^0 \rightarrow 2\gamma$ .  $\mathcal{T}_{fi}$  has been given in (8.103). To find its absolute value squared, we must sum over the polarization states of the two photons. Regarding  $e^{(\lambda')}$  and  $e^{(\lambda')}$ , we use (5.91) to sum over  $\lambda'$  and  $\lambda''$ :

$$\begin{aligned} \sum_{\lambda'} \sum_{\lambda''} (i\epsilon_{\mu\nu\alpha\beta} e'_\mu e''_\nu k'_\alpha k''_\beta)^2 &= -\epsilon_{\mu\nu\alpha\beta} \epsilon_{\mu\nu\alpha'\beta'} k'_\alpha k''_{\alpha'} k'_\beta k''_{\beta'} \\ &= 2(\delta_{\alpha\beta'} \delta_{\alpha'\beta} - \delta_{\alpha\alpha'} \delta_{\beta\beta'}) k'_\alpha k''_{\alpha'} k'_\beta k''_{\beta'} \\ &= 2[(k' \cdot k'')^2 - k'^2 k''^2] \\ &= \frac{1}{2} \mu^4. \end{aligned} \quad (9.62)$$

Here we denote the mass of  $\pi^0$  and the mass of the proton by  $\mu$  and  $m$ , respectively. We use (9.37) and divide by 2 so as not to count the same state twice, since there are two identical particles in the final state. Then,

$$\begin{aligned} \Gamma(\pi^0 \rightarrow 2\gamma) &= \frac{1}{2} \frac{1}{16\pi\mu} \left(\frac{1}{\pi}\right)^2 \left(\frac{e^2}{4\pi}\right)^2 \frac{G^2 \mu^4}{m^2} \\ &= \frac{\alpha^2}{64\pi^3} G^2 \left(\frac{\mu}{m}\right)^2 \mu, \quad \text{where } \alpha = \frac{e^2}{4\pi} \approx \frac{1}{137}. \end{aligned} \quad (9.63)$$

Taking the reciprocal,

$$\tau = 0.7 \times 10^{-15} \left(\frac{G^2}{4\pi}\right)^{-1} \text{ s}. \quad (9.64)$$

If we take  $G^2/4\pi = 15$ , we find  $\tau \approx 0.5 \times 10^{-16}$  s. However, experiment yields  $2 \times 10^{-16}$  s, so the agreement with theory is not very good. We will discuss this issue later in relation to the quark model.

Naturally, when there is no good agreement between theory and experiment, this leaves us with a problem, i.e., to find out whether perturbation theory is not a good approximation or the model is not good.

### 9.3 Inclusive Reactions

The cross-section considered up to now is the one for the exclusive reaction, which corresponds to observing all particles in the final state. On the other hand, we can also consider the case where certain particles in the final state are observed and others are not. We then refer to an *inclusive reaction*. In this section, we shall discuss how to compute the cross-section for an inclusive reaction.

We thus consider the inelastic scattering of an electron and a nucleon. As an example, we study the reaction

$$e^- + p \rightarrow e^- + \dots, \quad (9.65)$$

where we only observe the electron in the final state (see Fig. 9.1). We assume that an electron with four-momentum  $q$  exchanges a virtual photon with the proton and observe the electron whose four-momentum changes to  $q'$ . We assume that the proton absorbing the virtual photon goes into a state comprising  $n$  hadrons. We denote the four-momentum of the proton in the initial state by  $p$ , and denote the  $n$ -hadron state by  $|n\rangle$ . We take the electromagnetic interaction into account up to second order, splitting the Hamiltonian density as follows:

$$\mathcal{H}_{\text{int}}(x) = -j_\mu(x)A_\mu(x), \quad (9.66)$$

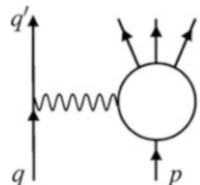
$$j_\mu = j_\mu^e(x) + j_\mu^h(x), \quad (9.67)$$

where  $j_\mu^e$  and  $j_\mu^h$  are the four-current densities of the electron and the hadrons, respectively. Therefore, to lowest order, the relevant  $S$ -matrix element becomes

$$S_{fi} = (-i)^2 \int d^4x \int d^4y \langle q' | j_\mu^e(x) | q \rangle D_F(x-y) \langle n | j_\mu^h(y) | p \rangle, \quad (9.68)$$

which describes elastic scattering. However, for the strong interaction, we must include higher order corrections. Denoting the Hamiltonian density of the strong

**Fig. 9.1** Inclusive reaction.  
An electron scatters on a  
proton



interaction by  $\mathcal{H}_s$ , higher order corrections are taken into account by making the substitution

$$\langle n | j_\mu^h(x) | p \rangle \longrightarrow \frac{\langle n | T \{ j_\mu^h(x) \exp[-i \int d^4 y \mathcal{H}_s(y)] \} | p \rangle}{\langle 0 | T \exp[-i \int d^4 y \mathcal{H}_s(y)] | 0 \rangle},$$

where we have dropped the bubble graphs from (8.70) and we have used the notation in (8.106). The right-hand side of this is equivalent to

$$\langle \mathbf{n} | \mathbf{j}_\mu^h(x) | \mathbf{p} \rangle, \quad (9.69)$$

as will be shown in Chap. 11. Here the bold face indicates the Heisenberg picture, but for the time being we will write (9.69) using the usual typeface, as there is no risk of confusion. We reinterpret the matrix elements of the hadron parts in (9.68) as standing for the Heisenberg picture. We rewrite (9.68) in the momentum representation and set  $V = 1$ :

$$\begin{aligned} S_{fi} &= -ie \int d^4 x \int d^4 y \bar{u}(q') \gamma_\mu u(q) e^{i(q-q') \cdot x} \left[ \frac{-i}{(2\pi)^4} \int d^4 k \frac{e^{ik \cdot (x-y)}}{k^2 - i\epsilon} \right] \langle n | j_\mu^h(0) | p \rangle e^{i(p-n) \cdot y} \\ &= (-i)^2 (2\pi)^4 e \int d^4 k \delta^4(q - q' + k) \delta^4(p - n - k) \frac{1}{k^2 - i\epsilon} \bar{u}(q') \gamma_\mu u(q) \langle n | j_\mu^h(0) | p \rangle \\ &= ie (2\pi)^4 \delta^4(p - n - q' + q) \frac{1}{(q' - q)^2} \bar{u}(q') i \gamma_\mu u(q) \langle n | j_\mu^h(0) | p \rangle, \end{aligned} \quad (9.70)$$

where  $e$  stands for the electric charge of the electron, including its sign. The matrix  $T$  is now given by

$$T_{fi} = -e \frac{1}{(q' - q)^2} \bar{u}(q') i \gamma_\mu u(q) \langle n | j_\mu^h(0) | p \rangle. \quad (9.71)$$

Thus, the total cross-section is

$$\begin{aligned} \sigma &= \frac{(2\pi)^4}{v_{\text{rel}}} \int \frac{d^3 q'}{(2\pi)^3} \sum_n \delta^4(n + q' - q - p) |T_{fi}|^2 \\ &= \frac{(2\pi)^4}{v_{\text{rel}}} \int \frac{d^3 q'}{(2\pi)^3} \sum_n \frac{e^2}{[(q' - q)^2]^2} \delta^4(n + q' - q - p) |\bar{u}(q') i \gamma_\mu u(q) \langle n | j_\mu^h(0) | p \rangle|^2. \end{aligned} \quad (9.72)$$

We introduce the four-vector  $Q = q - q'$  and decompose the right-hand side of (9.72) into the electron part and the hadronic part:

$$\sigma = \frac{M}{2p_0q_0v_{\text{rel}}} \frac{1}{(2\pi)^3} \int \frac{d^3q'}{2q'_0} \frac{e^2}{Q^4} 4q_0q'_0 \bar{u}(q') i\gamma_\nu u(q) \cdot \bar{u}(q) i\gamma_\mu u(q') \quad (9.73)$$

$$\times \sum_n \frac{p_0}{M} \langle p | j_\mu^h(0) | n \rangle \langle n | j_\nu^h(0) | p \rangle (2\pi)^4 \delta^4(n - p - Q) ,$$

where  $M$  is the proton mass. We also set  $B = p_0q_0v_{\text{rel}}$ . Averaging the electron spins over the initial states and summing them over the final states, we obtain

$$4q_0q'_0 \bar{u}(q') i\gamma_\nu u(q) \bar{u}(q) i\gamma_\mu u(q') \rightarrow \frac{1}{2} \text{Tr} [i\gamma_\nu (-iq \cdot \gamma + m) i\gamma_\mu (-iq' \cdot \gamma + m)] . \quad (9.74)$$

The value of the trace is

$$\frac{1}{2} \text{Tr} [\dots] = -2 [\delta_{\mu\nu} (q \cdot q' + m^2) - q_\mu q'_\nu - q'_\mu q_\nu] . \quad (9.75)$$

Furthermore, the hadron part transforms as a second-rank tensor:

$$\sum_n \frac{p_0}{M} \langle p | j_\mu^h(0) | n \rangle \langle n | j_\nu^h(0) | p \rangle (2\pi)^4 \delta^4(n - p - Q) \quad (9.76)$$

$$= \int d^4x \frac{p_0}{M} \langle p | j_\mu^h(x) j_\nu^h(0) | p \rangle e^{-iQ \cdot x} .$$

The total cross-section is now

$$\sigma = \frac{M}{2B} \frac{1}{(2\pi)^3} \int \frac{d^3q'}{2q'_0} \frac{e^2}{Q^4} \left( \frac{1}{2} \text{Tr} \right)_{\mu\nu} \int d^4x \frac{p_0}{M} \langle p | j_\mu^h(x) j_\nu^h(0) | p \rangle e^{-iQ \cdot x} . \quad (9.77)$$

In the following, we consider the laboratory system:

$$B = qM , \quad q = |\mathbf{q}| . \quad (9.78)$$

Then considering a very high energy and setting  $q \approx q_0 = E$  and  $q' \approx q'_0 = E'$ ,

$$\sigma = \frac{1}{2E} \frac{1}{(2\pi)^3} \int \frac{E'^2 dE' d\Omega'}{2E'} \frac{e^2}{Q^4} \left( \frac{1}{2} \text{Tr} \right)_{\mu\nu} \left( \int \right)_{\mu\nu} . \quad (9.79)$$

Inserting (9.75),

$$\frac{d^2\sigma}{d\Omega'dE'} = \frac{1}{(2\pi)^3} \frac{E'}{2E} \frac{e^2}{Q^4} [q_\mu q'_\nu + q'_\mu q_\nu - \delta_{\mu\nu}(q \cdot q' + m^2)] \left( \int \right)_{\mu\nu} . \quad (9.80)$$

The last factor on the right-hand side is a second-rank tensor proportional to  $e^2$ , whence we may write

$$\int d^4x \frac{p_0}{M} \langle p | j_\mu^h(x) j_\nu^h(0) | p \rangle e^{-iQ \cdot x} = 2\pi e^2 W_{\mu\nu} . \quad (9.81)$$

Here,  $W_{\mu\nu}$  is a function of the two four-momenta  $p$  and  $Q$  with the general form

$$W_{\mu\nu} = W_1 \left( \delta_{\mu\nu} - \frac{Q_\mu Q_\nu}{Q^2} \right) + \frac{1}{M^2} W_2 \left( p_\mu - \frac{p \cdot Q}{Q^2} Q_\mu \right) \left( p_\nu - \frac{p \cdot Q}{Q^2} Q_\nu \right) . \quad (9.82)$$

Equation (9.82) has been chosen in such a way that the following relation holds, corresponding to the conservation of electric current:

$$Q_\mu W_{\mu\nu} = Q_\nu W_{\mu\nu} = 0 . \quad (9.83)$$

Note also the following relations, which hold at high energies:

$$[q_\mu q'_\nu + q'_\mu q_\nu - \delta_{\mu\nu}(q \cdot q' + m^2)] \left( \delta_{\mu\nu} - \frac{Q_\mu Q_\nu}{Q^2} \right) \approx Q^2 \approx 4EE' \sin^2 \frac{\theta}{2} , \quad (9.84)$$

$$[q_\mu q'_\nu + q'_\mu q_\nu - \delta_{\mu\nu}(q \cdot q' + m^2)] \left( p_\mu - \frac{p \cdot Q}{Q^2} Q_\mu \right) \left( p_\nu - \frac{p \cdot Q}{Q^2} Q_\nu \right) \approx 2M^2 EE' \cos^2 \frac{\theta}{2} , \quad (9.85)$$

Substituting this into (9.80),

$$\frac{d^2\sigma}{d\Omega'dE'} = 4\alpha^2 \frac{E'^2}{Q^4} \left( 2W_1 \sin^2 \frac{\theta}{2} + W_2 \cos^2 \frac{\theta}{2} \right) . \quad (9.86)$$

The quantities  $W_1$  and  $W_2$  are called *structure functions*. Both are scalar functions constructed from  $p$  and  $Q$ . Taking into account the fact that  $p^2 = -M^2$ ,

$$Q^2 , \quad p \cdot Q \equiv M\nu , \quad (9.87)$$

are independent scalar quantities, where  $\nu$  is equal to the energy loss of the electron in the laboratory system:

$$\nu = E - E' . \quad (9.88)$$

According to Bjorken, if  $Q^2/2M\nu = x$  is finite in the limit  $Q^2 \rightarrow \infty$  and  $M\nu \rightarrow \infty$ , then both become functions of the ratio  $x$  alone, i.e.,

$$MW_1 \rightarrow F_1(x) , \quad \nu W_2 \rightarrow F_2(x) . \quad (9.89)$$

This is *Bjorken's scaling law* [105], which would later provide a basis for the gauge theory of strong interactions.

## 9.4 Optical Theorem

In quantum mechanics, the scattering amplitude for spin-0 particles passing through a central force field can be expressed by the following partial wave expansion:

$$f(\theta) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l} \sin \delta_l P_l(\cos \theta) . \quad (9.90)$$

Here  $k$  is the modulus of the momentum and  $P_l$  is the  $l$ th Legendre polynomial. If the scattering is purely elastic, the phase shift  $\delta_l$  is a real number, so the total cross-section  $\sigma$  satisfies

$$\sigma = \int d\Omega |f(\theta)|^2 = \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l = \frac{4\pi}{k} \text{Im} f(0) . \quad (9.91)$$

This is called the *optical theorem*. We shall now consider its generalization to field theory.

Inserting (8.71) into (8.2), we express the unitarity of the  $S$ -matrix in terms of  $T$ :

$$\text{Im} \mathcal{T}_{fi} = -\frac{(2\pi)^4}{2} \sum_j T_{fj}^\dagger \delta^4(P_f - P_j) T_{ji} . \quad (9.92)$$

Setting  $f = i$ ,

$$\text{Im} \mathcal{T}_{ii} = -\frac{(2\pi)^4}{2} \sum_j \delta^4(P_i - P_j) |T_{ji}|^2 . \quad (9.93)$$

We rewrite this relation in terms of the invariant amplitude  $\mathcal{T}$ :

$$\begin{aligned} \text{Im}\mathcal{T}_{ii} &= -\sum_n \frac{(2\pi)^{4-3n}}{2} \int \frac{d^3 p_1}{2p_{10}} \dots \int \frac{d^3 p_n}{2p_{n0}} \delta^4(P_i - P_n) |\mathcal{T}_{ni}|^2 \\ &= -2B \sum_{n=2}^{\infty} \sigma(i \rightarrow n), \end{aligned} \quad (9.94)$$

where we have assumed that the initial state  $|i\rangle$  is a two-particle state. Denoting the total cross-section by  $\sigma$ ,

$$\text{Im}\mathcal{T}_{ii} = -2B\sigma \quad (|i\rangle \text{ a two-particle state}). \quad (9.95)$$

This is the optical theorem in field theory. Similarly, for the total decay width  $\Gamma$ , we obtain

$$\text{Im}\mathcal{T}_{ii} = -P_0\Gamma \quad (|i\rangle \text{ a one-particle state}). \quad (9.96)$$

Note that the optical theorem in field theory holds for inelastic scattering as well as elastic scattering.

As an example of the optical theorem, let us consider the process where an electron-positron pair annihilates to produce hadrons:

$$e^- + e^+ \longrightarrow \text{hadrons}. \quad (9.97)$$

We consider the propagators, including the lowest order for the electromagnetic interaction and higher order corrections for the strong interaction:

$$D'_{F_{\mu\nu}} = \frac{\langle 0|T[A_\mu(x)A_\nu(y)\exp[-i\int d^4z\mathcal{H}_s(z)]]|0\rangle}{\langle 0|T\exp[-i\int d^4z\mathcal{H}_s(z)]]|0\rangle}. \quad (9.98)$$

Now corresponding to  $\mathcal{T}_{ii}$ , we consider the process

$$e^-(p) + e^+(\bar{p}) \longrightarrow e^-(p') + e^+(\bar{p}'), \quad (9.99)$$

where the four-momenta are indicated in brackets. If only hadrons exist between the in and out states, the  $S$ -matrix is

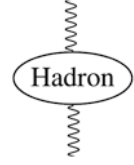
$$S^{(2)} = -\frac{1}{2} \int d^4x \int d^4y :j_\mu^e(x)D'_{F_{\mu\nu}}(x-y)j_\nu^e(y):. \quad (9.100)$$

Setting  $V = 1$ , the T-matrix element for this process can be written as

$$T_{fi} = -e^2 \bar{u}(p') i\gamma_\mu v(\bar{p}') D'_{F_{\mu\nu}}(p + \bar{p}) \bar{v}(\bar{p}) i\gamma_\nu u(p). \quad (9.101)$$



**Fig. 9.2** Hadrons produced by electron-positron annihilation



Thus, the invariant scattering amplitude for forward scattering is

$$\mathcal{T}_{ii} = -e^2 D'_{F\mu\nu}(p + \bar{p})(2p_0)(2\bar{p}_0)\bar{u}(p)i\gamma_\mu v(\bar{p})\bar{v}(\bar{p})i\gamma_\nu u(p). \quad (9.102)$$

Averaging over the spins in the initial state,

$$\begin{aligned} (2p_0)(2\bar{p}_0)\bar{u}(p)i\gamma_\mu v(\bar{p})\bar{v}(\bar{p})i\gamma_\nu u(p) &\rightarrow \frac{1}{4}\text{Tr}[i\gamma_\mu(-i\bar{p}\cdot\gamma - m)i\gamma_\nu(-ip\cdot\gamma + m)] \\ &= p_\mu\bar{p}_\nu + \bar{p}_\mu p_\nu + \delta_{\mu\nu}(m^2 - p\cdot\bar{p}). \end{aligned} \quad (9.103)$$

Setting

$$s = -(p + \bar{p})^2 = 2(m^2 - p\cdot\bar{p}), \quad (9.104)$$

we have

$$B = \sqrt{(p\cdot\bar{p} - m^4)^2} = \frac{1}{2}\sqrt{s(s - 4m^2)}. \quad (9.105)$$

To consider the imaginary part of  $\mathcal{T}_{ii}$ , we introduce only the contribution of  $D'_{F\mu\nu}$  for hadrons between the in and out states (see the Feynman diagram in Fig. 9.2):

$$D'_{F\mu\nu} \rightarrow \frac{1}{k^4}(k_\mu k_\nu - \delta_{\mu\nu}k^2)\sigma^h(-k^2) = \frac{1}{k^4}\Pi_{\mu\nu}^h(-k^2), \quad k = p + \bar{p}. \quad (9.106)$$

This can be written in this form for basically the same reasons as were given to justify (9.82). As can be read off from the Feynman diagram,  $\Pi_{\mu\nu}^h(k^2)$  can be written in the form

$$\langle 0|T[J_\mu^h(x)J_\nu^h(y)]|0\rangle = \frac{-i}{(2\pi)^4}\int d^4k e^{ik\cdot(x-y)}\Pi_{\mu\nu}^h(-k^2). \quad (9.107)$$

In this regard, however, it should be remembered that the left-hand side has been defined in the Heisenberg picture. Writing

$$\sigma^h(-k^2) = \Pi^h(s), \quad s = -k^2, \quad (9.108)$$

we obtain

$$\begin{aligned}
 -\text{Im}\mathcal{T}_{ii} &= e^2 \left( p_\mu \bar{p}_\nu + \bar{p}_\mu p_\nu + \frac{1}{2} \delta_{\mu\nu} s \right) \frac{1}{s^2} (k_\mu k_\nu - \delta_{\mu\nu} k^2) \text{Im}\Gamma^h(s) \\
 &= e^2 \frac{s + 2m^2}{s} \text{Im}\Gamma^h(s) .
 \end{aligned}
 \tag{9.109}$$

Therefore, from the optical theorem the total cross-section of (9.97) is

$$\sigma(e^- + e^+ \rightarrow \text{hadrons}) = \frac{1}{2B} (-\text{Im}\mathcal{T}_{ii}) = \frac{e^2}{s} \frac{s + 2m^2}{\sqrt{s(s - 4m^2)}} \text{Im}\Gamma^h(s) .
 \tag{9.110}$$

In particular, assuming  $s \gg m^2$  at high energies,

$$\sigma(e^- + e^+ \rightarrow \text{hadrons}) = \frac{e^2}{s} \text{Im}\Gamma^h(s) .
 \tag{9.111}$$

We shall see later that measurements of this cross-section can tell us how many types of quarks (flavours) there are.

## 9.5 Three-Body Decays

In Sect. 9.1, we discussed the integral in phase space when the final state is a two-particle state. In this section, we consider the integral for the three-particle state. Although this sort of integral is complicated in many cases, we will discuss several tractable examples here.

To begin with, we consider the decay of the  $\mu$ -meson:

$$\mu^+ \rightarrow e^+ + \nu + \bar{\nu} .
 \tag{9.112}$$

We shall not be concerned whether or not  $\bar{\nu}$  appearing here is an antiparticle of  $\nu$  because it is not relevant to the following discussion. In this process, the only quantity observed in the final state is the electron. We will discuss the form of the electron energy spectrum. The reaction (9.112) can be described by what is known as the *Fermi interaction*. The invariant decay width  $\mathcal{T}_{fi}$  can be expressed in terms of four Dirac spinors and suitable  $\gamma$ -matrices. If we average or sum spins, then the following can be expressed in terms of four Casimir operators and traces of products of suitable  $\gamma$ -matrices:

$$\sum_{S_i} \sum_{S_f} |\mathcal{T}_{fi}|^2 .
 \tag{9.113}$$

Therefore, the result must be linear scalar functions of  $p_\nu$ ,  $p_{\bar{\nu}}$ , and  $p_e$ , or  $m_e$ ,  $p_\mu$ , and  $m_\mu$ . The possible combinations are:

$$m_\mu m_e (p_\nu \cdot p_{\bar{\nu}}), \quad (p_\mu \cdot p_e)(p_\nu \cdot p_{\bar{\nu}}), \quad (p_\mu \cdot p_\nu)(p_e \cdot p_{\bar{\nu}}), \quad (p_\mu \cdot p_{\bar{\nu}})(p_e \cdot p_\nu). \quad (9.114)$$

Inserting these into the formula for the decay width, we carry out the integration. Since the  $\mu$ -meson is about 200 times more massive than the electron, the speed of the emerging electron is close to the speed of light, so  $m_e$  can be neglected. Then we do not need to consider the first term in (9.114). The third and fourth terms will yield the same results if we integrate over  $p_\nu$  and  $p_{\bar{\nu}}$  without observing the neutrino, so we cannot tell the difference between them as long as we discuss possible shapes of the energy spectrum. Hence, in order to determine the form of the energy spectrum, it is enough to consider only the two terms

$$(p_\mu \cdot p_e)(p_\nu \cdot p_{\bar{\nu}}), \quad (p_\mu \cdot p_\nu)(p_e \cdot p_{\bar{\nu}}). \quad (9.115)$$

The parameter which determines the form of the spectrum is simply a ratio of their coefficients, and chosen in a suitable way, it is known as the *Michel parameter* [106].

Up to a multiplicative constant, the decay rate is given by

$$\begin{aligned} \Gamma &\propto \int \frac{d^3 p_e}{2p_{e0}} \int \frac{d^3 p_\nu}{2p_{\nu 0}} \int \frac{d^3 p_{\bar{\nu}}}{2p_{\bar{\nu} 0}} \delta^4(P - p_e - p_\nu - p_{\bar{\nu}}) \langle |\mathcal{T}_{fi}|^2 \rangle \\ &= \int \frac{d^3 p_\nu}{2p_{\nu 0}} \int d^4 Q \delta^4(P - p_e - Q) \int \frac{d^3 p_\nu}{2p_{\nu 0}} \frac{d^3 p_{\bar{\nu}}}{2p_{\bar{\nu} 0}} \delta^4(Q - p_\nu - p_{\bar{\nu}}) \langle |\mathcal{T}_{fi}|^2 \rangle, \end{aligned} \quad (9.116)$$

where  $\langle |\mathcal{T}_{fi}|^2 \rangle$  is averaged over the spins. It will be a linear combination of the two terms in (9.115). In order to integrate over the neutrino momenta, we must evaluate

$$\begin{aligned} I_{\alpha\beta} &= \int \frac{d^3 p_\nu}{2p_{\nu 0}} \int \frac{d^3 p_{\bar{\nu}}}{2p_{\bar{\nu} 0}} \delta^4(Q - p_\nu - p_{\bar{\nu}}) (p_\nu)_\alpha (p_{\bar{\nu}})_\beta \\ &= \int d^4 p_\nu \int d^4 p_{\bar{\nu}} \delta_+(p_\nu^2) \delta_+(p_{\bar{\nu}}^2) \delta^4(Q - p_\nu - p_{\bar{\nu}}) (p_\nu)_\alpha (p_{\bar{\nu}})_\beta, \end{aligned} \quad (9.117)$$

where we have used the notation  $\delta_+(p^2) = \theta(p_0) \delta(p^2)$ . We make the change of variables

$$p_\nu = \frac{Q'}{2} + \Delta, \quad p_{\bar{\nu}} = \frac{Q'}{2} - \Delta. \quad (9.118)$$

Therefore, the integral is now over  $d^4 Q' d^4 \Delta$ . Now integrating over  $Q'$ , we can take  $Q' = Q$  in the remaining integral,

$$\begin{aligned} I_{\alpha\beta} &= \int d^4 \Delta \delta \left[ \left( \frac{Q}{2} + \Delta \right)^2 \right] \delta \left[ \left( \frac{Q}{2} - \Delta \right)^2 \right] \left( \frac{Q}{2} + \Delta \right)_\alpha \left( \frac{Q}{2} - \Delta \right)_\beta \\ &= \int d^4 \Delta \delta \left( \frac{Q^2}{4} + \Delta^2 \right) \delta(2Q \cdot \Delta) \left( \frac{Q}{2} + \Delta \right)_\alpha \left( \frac{Q}{2} - \Delta \right)_\beta, \end{aligned} \quad (9.119)$$

where we have assumed that the integration is carried out only in the region  $Q_0 > 2|\Delta_0|$ . Since  $Q$  is a time-like vector, we choose the direction of  $Q$  to be the time axis. Then integrating  $\delta(Q \cdot \Delta)$  with respect to  $\Delta_0$ , we can take

$$\Delta_0 = 0 \quad (9.120)$$

in the remaining integral and only terms of even order in  $\vec{\Delta}$  survive:

$$I_{\alpha\beta} = \frac{1}{2Q_0} \int \frac{\vec{\Delta}^2 d|\vec{\Delta}^2| d\Omega}{2|\vec{\Delta}|} \delta \left( \frac{Q^2}{4} + \vec{\Delta}^2 \right) \left( \frac{1}{4} Q_\alpha Q_\beta - \Delta_\alpha \Delta_\beta \right). \quad (9.121)$$

By symmetry, the integral in the second term survives only when  $\alpha = \beta \neq 4$ , so we can make the following replacement:

$$\begin{aligned} \Delta_\alpha \Delta_\beta &\rightarrow \frac{1}{3} \delta_{\alpha\beta} (1 - \delta_{\alpha 4}) \vec{\Delta}^2 = \frac{1}{3} \left( \delta_{\alpha\beta} - \frac{Q_\alpha Q_\beta}{Q^2} \right) \vec{\Delta}^2 \\ &= -\frac{1}{12} (\delta_{\alpha\beta} Q^2 - Q_\alpha Q_\beta). \end{aligned} \quad (9.122)$$

Integrating the first term and the remaining one,

$$I_{\alpha\beta} = \frac{\pi}{24} (\delta_{\alpha\beta} Q^2 + 2Q_\alpha Q_\beta). \quad (9.123)$$

Inserting this result into (9.116), we are left with two integrals:

$$\int \frac{d^3 p}{2p_0} \int d^4 Q \delta^4(P - p - Q)(P \cdot p)(4Q^2 + 2Q^2), \quad (9.124)$$

$$\int \frac{d^3 p}{2p_0} \int d^4 Q \delta^4(P - p - Q)[(P \cdot p)Q^2 + 2(P \cdot Q)(p \cdot Q)], \quad (9.125)$$

where we have taken  $p = p_e$  and  $P = p_\mu$ . We consider the rest frame of the  $\mu$ -meson, i.e.,

$$P_0 = m_\mu, \quad \mathbf{P} = 0. \quad (9.126)$$

The energy of the electron attains a maximum value of  $m_\mu/2$  when the neutrinos are emitted in opposite directions, and since we are assuming  $m_e = 0$ , we can set

$$p_0 = |\mathbf{p}| = \frac{1}{2}m_\mu x, \quad (9.127)$$

where the possible values of  $x$  lie between 0 and 1, and

$$\frac{d^3 p}{2p_0} \propto x dx. \quad (9.128)$$

Then the  $x$ -dependences of the scalar products in the integral are

$$\begin{aligned} P \cdot p &= -\frac{1}{2}m_\mu^2 x, & -Q^2 &= -P^2 + 2p \cdot P = m_\mu^2(1-x), \\ P \cdot Q &= P^2 - p \cdot P = m_\mu^2 \left(-1 + \frac{x}{2}\right), & p \cdot Q &= p \cdot P = -\frac{1}{2}m_\mu^2 x. \end{aligned}$$

After a trivial integration over the solid angle, we now obtain

$$\Gamma \propto \int_0^1 dx x^2(a + bx), \quad (9.129)$$

whence the spectrum is a linear combination of  $x^2$  and  $x^3$ . Normalizing the integral to unity and leaving the ratio of two suitable linear combinations of  $x^2$  and  $x^3$  as a parameter, the energy spectrum of the electron becomes

$$dN = 4x^2 \left[ 3(1-x) + \frac{2}{3}\rho(4x-3) \right] dx. \quad (9.130)$$

This has been chosen so that the first term vanishes when  $x = 1$  and the integral over  $x$  from 0 to 1 is unity. Carrying out the integral over  $x$ , the second term vanishes, and its coefficient  $\rho$  is called the *Michel parameter*. The value of  $\rho$  can be determined from the form of the spectrum when  $x \sim 1$ . The value of  $\rho$  depends on the choice of the  $\gamma$ -matrices in the Fermi four-fermion interaction. For the vector–axial vector (VA) type,  $\rho = 0.75$  and for the scalar–pseudoscalar (SP) type,  $\rho = 0$ . This will be discussed later.

In association with the three-body decay, another important idea is the *Dalitz plot* [43, 107]. Dalitz proposed this to determine the parity and spin of the  $K^+$ -meson from the energy distribution of the  $\pi$ -meson in the decay process

$$K^+ \rightarrow \pi^+ + \pi^+ + \pi^-. \quad (9.131)$$

The decay width is given by

$$\begin{aligned} \Gamma(K^+ \rightarrow \pi^+ + \pi^+ + \pi^-) & \quad (9.132) \\ &= \frac{(2\pi)^{-5}}{2M} \int \frac{d^3 p_1}{2p_{10}} \int \frac{d^3 p_2}{2p_{20}} \int \frac{d^3 p_3}{2p_{30}} \delta^4(P_f - P_i) |\mathcal{T}(K^+ \rightarrow \pi^+ + \pi^+ + \pi^-)|^2, \end{aligned}$$

where  $P_i$  is the four-momentum of the  $K^+$  meson,  $p_1$  and  $p_2$  are the four-momenta of the two  $\pi^+$  mesons, and  $p_3$  is the four-momentum of the  $\pi^-$  meson. These satisfy the relation  $P_f = p_1 + p_2 + p_3$ . In order not to count the same state twice, we restrict the integration to the region

$$p_{10} < p_{20}. \quad (9.133)$$

$|\mathcal{T}|^2$  is a function of the scalar products  $(p_i \cdot p_j)$ . Only two of them are independent so we choose

$$p_1 \cdot P_i, \quad p_3 \cdot P_i, \quad (9.134)$$

as independent variables. In the rest frame of the  $K$ -meson, these scalar products can be written as  $-M\omega_1$  and  $-M\omega_3$ , respectively. Here,  $M$  is the mass of the  $K$ -meson and  $\omega_1$  and  $\omega_3$  are the energies of the  $\pi$ -mesons in the rest frame of the  $K$ -meson. From the law of conservation of energy,

$$\omega_1 + \omega_2 + \omega_3 = M. \quad (9.135)$$

Since  $|\mathcal{T}|^2$  can be viewed as a function of  $\omega_1$  and  $\omega_3$ , the integrals in (9.132) can be written in the form

$$\begin{aligned} I &= \int \frac{d^3 p_1}{\omega_1} \int \frac{d^3 p_2}{\omega_2} \int \frac{d^3 p_3}{\omega_3} \delta^3(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \delta(\omega_1 + \omega_2 + \omega_3 - M) f(\omega_1, \omega_3) \\ &= \int d\omega'_1 \int d\omega'_3 K(\omega'_1, \omega'_3) f(\omega'_1, \omega'_3), \quad (9.136) \end{aligned}$$

where

$$K(\omega'_1, \omega'_3) = \int \frac{p_1^2 d\Omega_1}{\omega_1} \frac{p_3^2 d\Omega_3}{\omega_3} \frac{\delta(\omega_1 + \omega_2 + \omega_3 - M)}{\omega_2} \delta(\omega_1 - \omega'_1) \delta(\omega_3 - \omega'_3). \quad (9.137)$$

Here  $p_1 = |\mathbf{p}_1|$  and  $p_3 = |\mathbf{p}_3|$ . We begin by integrating with respect to  $\omega_3$ . Choosing the direction of  $\mathbf{p}_1$  as the  $z$ -direction and taking the angle between  $\mathbf{p}_1$

and  $\mathbf{p}_3$  as  $\theta$ ,

$$\int d\Omega_1 \int d\Omega_3 \dots = 8\pi^2 \int_{-1}^1 d(\cos \theta) \dots \quad (9.138)$$

We have left  $\cos \theta$  in the above because, from the law of conservation of momentum,  $\omega_2$  is a function of  $\cos \theta$ :

$$\begin{aligned} \omega_1^2 &= p_1^2 + \mu^2, & \omega_3^2 &= p_3^2 + \mu^2, \\ \omega_2^2 &= (\mathbf{p}_1 + \mathbf{p}_3)^2 + \mu^2 = p_1^2 + p_3^2 + 2p_1 p_3 \cos \theta + \mu^2. \end{aligned} \quad (9.139)$$

In order to change the integration variables from  $p_1$ ,  $p_3$ , and  $\cos \theta$  to  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$ , we use the Jacobian

$$\frac{\partial(\omega_1, \omega_2, \omega_3)}{\partial(p_1, p_3, \cos \theta)} = \frac{p_1^2 p_3^2}{\omega_1 \omega_2 \omega_1}. \quad (9.140)$$

Hence,

$$\begin{aligned} K(\omega'_1, \omega'_3) &= 8\pi^2 \int dp_1 dp_2 d(\cos \theta) \frac{\partial(\omega_1, \omega_2, \omega_3)}{\partial(p_1, p_3, \cos \theta)} \\ &\quad \times \delta(\omega_1 + \omega_2 + \omega_3 - M) \delta(\omega_1 - \omega'_1) \delta(\omega_3 - \omega'_3) \\ &= 8\pi^2 \int d\omega_1 d\omega_2 d\omega_3 \delta(\omega_1 + \omega_2 + \omega_3 - M) \delta(\omega_1 - \omega'_1) \delta(\omega_3 - \omega'_3). \end{aligned} \quad (9.141)$$

If the values of  $\omega'_1$  and  $\omega'_3$  are physically realizable, then this integral yields  $8\pi^2$ . If not, it yields zero. All unnecessary terms can thus be completely dropped. This is an amazing simplification. The equation (9.132) for the decay width becomes

$$\Gamma(\mathbf{K}^+ \rightarrow \pi^+ + \pi^+ + \pi^-) = \frac{(4\pi)^{-3}}{M} \int_{\omega_1 < \omega_2} d\omega_1 d\omega_3 |\mathcal{T}(\mathbf{K}^+ \rightarrow \pi^+ + \pi^+ + \pi^-)|^2. \quad (9.142)$$

What is important here is the proportionality relation

$$\frac{d^2\Gamma}{d\omega_1 d\omega_3} \propto |\mathcal{T}|^2. \quad (9.143)$$

Thus, plotting examples of each decay as points in the  $(\omega_1, \omega_3)$  plane, the density of the plot turns out to be proportional to the modulus squared of the decay width, so that  $|\mathcal{T}|^2$  can be found from experimental results as a function of  $\omega_1$  and  $\omega_3$ .

Dalitz showed how to produce this plot inside an equilateral triangle as follows. Defining the kinetic energy of each  $\pi$ -meson as  $T_i = \omega_i - \mu$ , energy conservation implies

$$T_1 + T_2 + T_3 = M - 3\mu \equiv Q . \quad (9.144)$$

Then, picking one point inside the triangle and drawing perpendiculars from that point to the three edges, the sum of the perpendiculars will be equal to the height of the triangle. We choose the height of the triangle as  $Q$  in suitable units. Then we express an example decay observed by experiment as an internal point so that the lengths of each perpendicular connecting the point with each edge are equal to  $T_1$ ,  $T_2$ , and  $T_3$ , respectively. This gives what is known as the Dalitz plot.

Since the value of  $Q$  is about 75 MeV, the pions emitted in the decay process can be treated as non-relativistic to a good approximation. Let us investigate what kind of plot we obtain in this case. The laws of conservation of energy and momentum are

$$p_1^2 + p_2^2 + p_3^2 = 2\mu Q , \quad (9.145)$$

$$\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 = 0 . \quad (9.146)$$

This means that  $p_1$ ,  $p_2$ , and  $p_3$  must satisfy the triangle inequality

$$(p_1 + p_3)^2 \geq p_2^2 \geq (p_1 - p_3)^2 , \quad (9.147)$$

or

$$2p_1 p_2 \geq |p_2^2 - p_1^2 - p_3^2| .$$

Writing this in terms of  $T$ ,

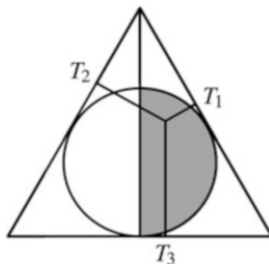
$$4T_1 T_3 \geq (T_2 - T_1 - T_3)^2 = (Q - 2T_1 - 2T_3)^2 . \quad (9.148)$$

This describes the interior of an ellipse in an oblique coordinate system. However, since this diagram is symmetric in  $T_1$ ,  $T_2$ , and  $T_3$ , it must remain invariant under a rotation of 120 degrees around the center of mass of the regular triangle. This means it has to be a circle. Hence, this is the interior of the circle inscribed inside an equilateral triangle. In Fig. 9.3, the lengths of the perpendiculars connecting the point in the interior of the circle with the three edges of the triangle  $T_1$ ,  $T_2$  and  $T_3$ , respectively. Due to the condition  $T_1 < T_2$ , points are restricted to lie within the semicircle.

Dalitz's plot was originally proposed to determine the spin and parity of  $K^+$ . To do so, we must determine the spin and parity of the final state. We assume that the relative orbital angular momentum of the two  $\pi^+$  particles is  $l$ , and that the relative



**Fig. 9.3** Dalitz plot



**Table 9.1** Smallest combinations of  $(l, l')$  for combinations of spins and parities in several kinds of  $3\pi$  system

Spins and parities in $3\pi$ -system	Smallest value of $(l, l')$
$0^+$	Nothing
$0^-$	(0,0)
$1^+$	(0,1)
$1^-$	(2,2)
$2^+$	(2,1)
$2^-$	(0,2), (2,0)
$\vdots$	$\vdots$

orbital angular momentum of the center of the two  $\pi^+$  particles and  $\pi^-$  is  $l'$ . Since the spin of  $\pi$  is zero, denoting the spin of  $K^+$  by  $J$ , we have

$$J = l + l' . \tag{9.149}$$

- Spinless K Meson.** In this case, since  $J = 0$ , we obviously have  $l = l'$ . Then the parity of the system with three  $\pi$ -mesons is  $(-1)^{l+l'} = 1$  from the orbital part. The intrinsic parity of  $\pi$  itself is  $-1$ , so it will be  $(-1)^3 = -1$  for three pions. The product of these two is thus  $-1$ . Therefore, if the  $K$  meson is spinless, the parity is odd and it can be denoted by  $0^-$ .
- General Case.** From the Bose symmetry for particles of the same kind,  $l$  is even. Table 9.1 gives the smallest combinations of  $(l, l')$  for combinations of spins and parities in several kinds of  $3\pi$  system. It should be obvious that the parity is given by  $(-1)^{l'+1}$ .

Since the value of  $Q$  for this decay is small, viz., 75 MeV, the large decay width corresponding to  $(l, l')$  becomes small due to the centrifugal force. Denoting the relative momenta corresponding to  $l$  and  $l'$  by  $p$  and  $q$ ,

$$\mathcal{F} \sim (pR)^l (qR)^{l'} , \tag{9.150}$$

where  $R$  is the range of the force among the  $\pi$  mesons. Since the density of the Dalitz plot is shown experimentally to be homogeneous,

$$|\mathcal{F}|^2 \sim \text{const.} \tag{9.151}$$

Thus,  $l = l' = 0$  and we find that the  $3\pi$ -system is in the  $0^-$  state. If  $K^+$  is spinless, the process

$$K^+ \rightarrow \pi^+ + \gamma$$

is forbidden, as discussed in Sect. 8.7. Indeed, this decay pattern has not been observed in experiments.

Similarly, given the density of the Dalitz plot for the process  $\eta^0 \rightarrow \pi^+ + \pi^- + \pi^0$  and the existence of the process  $\eta^0 \rightarrow 2\gamma$ , the neutral meson  $\eta^0$  turns out to be  $0^-$ .

We observe that  $K^+$  can also decay into two  $\pi$  mesons, viz.,

$$K^+ \rightarrow \pi^+ + \pi^0 . \tag{9.152}$$

Since the spin of  $K^+$  is zero, the relative angular momentum between  $\pi^+$  and  $\pi^0$  must be zero. The spin and the parity of this final state with two  $\pi$  mesons can be written as  $0^+$ . This contradicts the result for the  $3\pi$  decay, which gave  $0^-$ . The decay patterns of (9.152) and (9.131) were originally referred to as  $\theta$  and  $\tau$ , respectively, and this contradiction was known as the  $\theta$ - $\tau$  puzzle. Although  $\theta$  and  $\tau$  were originally considered to be different particles, more accurate detections showed that their mass and lifetime were actually the same. Furthermore, it gradually transpired that the ratio of the frequencies of the two reactions (decay to three vs. decay to two pions) is a constant, independent of the experimental conditions for their creation, so it was finally accepted that they must be same particles. Then Lee and Yang suggested that parity might not be conserved in weak interactions. To check this idea, they proposed several kinds of experiment [45], and in 1956 it was eventually proven that their idea was correct.

# Chapter 10

## Discrete Symmetries



In Chap. 7, we discussed Noether's theorem, which shows how a conservation law can be derived from the invariance of a theory under a continuous transformation group. These conservation laws hold in both classical and quantum theories. In contrast, discrete symmetries such as space inversion and charge conjugation are specific to quantum theory, where there are several kinds of conservation law that do not arise classically. Since physically observable quantities always vary continuously in classical theory, discrete symmetries are not connected to conservation laws. However, given one classical solution, this kind of symmetry leads us to another classical solution.

### 10.1 Symmetries and Unitary Transformations

When we discussed the invariance of a theory under a group of continuous transformations in Chap. 7, we worked from the invariance under an infinitesimal transformation group. In quantum theory, for a transformation like this, the state vector transforms under a unitary transformation generated as a product of infinitesimal transformations:

$$\Phi \rightarrow U\Phi = \exp(i\lambda Q)\Phi, \quad (10.1)$$

where  $Q$  is a conserved quantity derived by Noether's theorem and represented by a Hermitian operator, and  $\lambda$  is an arbitrary real number. Therefore,  $U$  is a unitary operator.

A discrete unitary transformation can be expressed by a unitary operator that does not have a generator like this. Charge conjugation, discussed in Sects. 3.9 and 8.6, can be represented by the discrete unitary transformation  $\mathcal{C}$ . This transformation does not change the Lagrangian density, nor the form of the equation of motion,

and it commutes with the  $S$ -matrix. Thus,  $\mathcal{C}$  is a conserved quantity. In QED, for example, the field operators transform according to

$$\mathcal{C}^{-1}\psi\mathcal{C} = C\bar{\psi}, \quad \mathcal{C}^{-1}\bar{\psi}\mathcal{C} = C^{-1}\psi, \quad \mathcal{C}^{-1}A_\mu\mathcal{C} = -A_\mu, \quad (10.2)$$

under this transformation. As already discussed in Sect. 3.9, the Dirac matrix  $C$  for charge conjugation is uniquely determined up to a complex phase. Denoting  $C$  for the spinor field  $\psi_a$  by  $C_a$  and using a standard matrix  $C$ , we may write

$$C_a = \eta_a C, \quad |\eta_a| = 1. \quad (10.3)$$

When a system possesses charge conjugation invariance, the phase  $\eta_a$  is chosen in such a way that the Lagrangian density is invariant under the charge conjugation transformation. In a system which is not invariant under any choice of the phase, charge conjugation is not a symmetry. Incidentally, the relations in (10.2) imply

$$\mathcal{C}^{-2}\psi\mathcal{C}^2 = \psi, \quad \mathcal{C}^{-2}\bar{\psi}\mathcal{C}^2 = \bar{\psi}, \quad \mathcal{C}^{-2}A_\mu\mathcal{C}^2 = A_\mu, \quad (10.4)$$

that is,  $\mathcal{C}^2$  commutes with all field operators. Thus, we can deduce that  $\mathcal{C}^2$  is a  $c$ -number. By a suitable choice of phase for  $\mathcal{C}$  itself, we can thus arrange for

$$\mathcal{C}^2 = 1. \quad (10.5)$$

In general, charge conjugation is conserved under strong and weak interactions.

In a similar way, space inversion  $\mathcal{P}$  is also conserved in all the fundamental interactions except the weak interaction, i.e., the form of the Lagrangian density and the equation of motion remain invariant, and  $\mathcal{P}$  also commutes with the  $S$ -matrix. As already stated in (3.128), for the spinor field  $\psi_a$ ,

$$\mathcal{P}^{-2}\psi_a(\mathbf{x}, t)\mathcal{P} = \epsilon_a \gamma_4 \psi_a(-\mathbf{x}, t), \quad |\epsilon_a| = 1. \quad (10.6)$$

Implementing the space inversion twice, this just becomes a phase transformation:

$$\mathcal{P}^{-1}\psi(\mathbf{x}, t)\mathcal{P}^2 = \epsilon_a^2 \psi(\mathbf{x}, t). \quad (10.7)$$

If a theory is invariant under a phase transformation like this, then redefining a suitable combination of the original  $\mathcal{P}$  in (10.6) and this phase transformation to give a new  $\mathcal{P}$ , we have

$$\mathcal{P}^{-2}\psi(\mathbf{x}, t)\mathcal{P}^2 = \psi(\mathbf{x}, t). \quad (10.8)$$

In this case, the following restriction is imposed on  $\epsilon_a$ :

$$\epsilon_a^2 = 1, \quad \epsilon_a = \pm 1. \quad (10.9)$$

This  $\epsilon_a$  is called the *intrinsic parity* of the spinor field labeled by  $a$ . We chose the intrinsic parity of both the proton and the neutron to be  $+1$  in Sect. 6.4. In addition, since  $\mathcal{P}^2$  itself commutes with all field operators, we can adjust the overall phase so that

$$\mathcal{P}^2 = 1 . \quad (10.10)$$

If the square of a unitary operator is unity, as in (10.5) and (10.10), this operator is also Hermitian. The intrinsic parity of one spinor field can be fixed freely if the system is invariant under a group of continuous phase transformations, in which case the phase of parity can be absorbed by the continuous transformations. If the system is invariant under a number of phase transformations, then we can correspondingly fix the intrinsic parities of an equal number of spinor fields. A useful convention is

$$\epsilon_n = \epsilon_p = \epsilon_\Lambda = 1 . \quad (10.11)$$

This convention is a slight extension of the one mentioned above and is associated with the Sakata model. In the quark model, this convention comes down to the convention that all quarks have intrinsic parity  $+1$ .

Generalizing the above discussion, we introduce a unitary transformation  $\mathcal{U}$  corresponding to one of the discrete symmetries. This must commute with the  $S$ -matrix:

$$\mathcal{U}^{-1} S \mathcal{U} = S . \quad (10.12)$$

Now consider two eigenstates of  $\mathcal{U}$ :

$$\mathcal{U}|a\rangle = \epsilon_a|a\rangle , \quad \mathcal{U}|b\rangle = \epsilon_b|b\rangle . \quad (10.13)$$

If  $\epsilon_a \neq \epsilon_b$ , the transition  $a \rightarrow b$  is prohibited, i.e.,

$$\langle b|S|a\rangle = 0 , \quad \text{if } \epsilon_a \neq \epsilon_b . \quad (10.14)$$

This is a generalization of Furry's theorem. As a property of the unitary transformation, taking

$$|a'\rangle = \mathcal{U}|a\rangle , \quad |b'\rangle = \mathcal{U}|b\rangle , \quad (10.15)$$

we obtain

$$\langle b'|a'\rangle = \langle b|a\rangle . \quad (10.16)$$

## 10.2 Parity of Antiparticles

When discussing charge conjugation and space inversion, an important issue is the commutation relation. We assume that the vacuum is simultaneously an eigenstate of both operators with eigenvalue 1. This is a consequence of the assumption that the vacuum is not degenerate:

$$\mathcal{P}|0\rangle = \mathcal{C}|0\rangle = |0\rangle . \quad (10.17)$$

In the following, we adopt the convention

$$\mathcal{P}^2 = \mathcal{C}^2 = 1 . \quad (10.18)$$

We now carry out these two transformations in two different orders on a Dirac field, and compare the results:

$$\begin{aligned} \mathcal{P}^{-1}\mathcal{C}^{-1}\psi(\mathbf{x}, t)\mathcal{C}\mathcal{P} &= \mathcal{P}^{-1}[\eta C\bar{\psi}^T(\mathbf{x}, t)]\mathcal{P} \\ &= \eta\epsilon C[\bar{\psi}(-\mathbf{x}, t)\gamma_4]^T \\ &= \eta\epsilon C\gamma_4^T\bar{\psi}^T(-\mathbf{x}, t) \\ &= -\eta\epsilon\gamma_4 C\bar{\psi}^T(-\mathbf{x}, t) \end{aligned}$$

and

$$\begin{aligned} \mathcal{C}^{-1}\mathcal{P}^{-1}\psi(\mathbf{x}, t)\mathcal{P}\mathcal{C} &= \mathcal{C}^{-1}[\epsilon\gamma_4\psi(-\mathbf{x}, t)]\mathcal{C} \\ &= \eta\epsilon\gamma_4 C\bar{\psi}^T(-\mathbf{x}, t) \\ &= -\mathcal{P}^{-1}\mathcal{C}^{-1}\psi(\mathbf{x}, t)\mathcal{C}\mathcal{P} . \end{aligned} \quad (10.19)$$

Note that, since  $\bar{\psi}$  is a  $1 \times 4$  matrix, the transpose  $\bar{\psi}^T$  is a  $4 \times 1$  matrix. Generalizing (10.19),

$$\mathcal{C}^{-1}\mathcal{P}^{-1}\psi(x_1) \dots \psi(x_n)\mathcal{P}\mathcal{C} = (-1)^n \mathcal{P}^{-1}\mathcal{C}^{-1}\psi(x_1) \dots \psi(x_n)\mathcal{C}\mathcal{P} . \quad (10.20)$$

It is possible for  $\bar{\psi}$  to appear in the product of these  $n$  operators. Multiplying the vacuum state by (10.20) and denoting the state containing  $n$  Dirac particles by  $|n\rangle$ ,

$$\mathcal{C}\mathcal{P}|n\rangle = (-1)^n \mathcal{P}\mathcal{C}|n\rangle . \quad (10.21)$$

We assume that the eigenstate of  $\mathcal{P}$  and its eigenvalue are  $|n\rangle$  and  $\epsilon(n)$ , respectively. Moreover, taking  $\mathcal{C}|n\rangle = |\bar{n}\rangle$ , we have

$$\mathcal{P}|\bar{n}\rangle = (-1)^n \epsilon(n) |\bar{n}\rangle . \quad (10.22)$$

Thus, considering the state with one Dirac particle ( $n = 1$ ), we observe that the intrinsic parities of  $|1\rangle$  and  $|\bar{1}\rangle$  have opposite signs. Assuming that a Bose particle can transit to a state with an even number of Fermi particles, it must have the same intrinsic parity as its antiparticle. Therefore, we obtain the following theorem:

**Theorem 10.1** *A Bose particle and its antiparticle have the same intrinsic parity, while a Fermi particle and its antiparticle have opposite intrinsic parities.*

If  $N$  is the number of Fermi particles, (10.21) implies the commutation relation

$$\mathcal{C}\mathcal{P} = \mathcal{P}\mathcal{C}(-1)^N = (-1)^N \mathcal{P}\mathcal{C} . \quad (10.23)$$

Since  $(-1)^N$  commutes with both  $\mathcal{C}$  and  $\mathcal{P}$ , considering the group generated by  $\mathcal{C}$ ,  $\mathcal{P}$ , and  $(-1)^N$ , we note that all raised to the second power are equal to 1 and that  $(-1)^N$  thus belongs to the center of the group.

### $\pi^+ - \pi^-$ System

Next, we consider a two-particle system in the center-of-mass frame so that it is in an eigenstate of  $\mathcal{P}$ . We express this state in terms of the creation operator and act on it with  $\mathcal{C}$  and  $\mathcal{P}$ :

$$\Phi = \varphi(\mathbf{p})\varphi^\dagger(-\mathbf{p})|0\rangle , \quad (10.24)$$

$$\mathcal{C}\Phi = \varphi^\dagger(\mathbf{p})\varphi(-\mathbf{p})|0\rangle , \quad (10.25)$$

$$\mathcal{P}\Phi = \varphi(-\mathbf{p})\varphi^\dagger(\mathbf{p})|0\rangle = \varphi^\dagger(\mathbf{p})\varphi(-\mathbf{p})|0\rangle , \quad (10.26)$$

where  $\mathcal{C}^{-1} = \mathcal{C}$  and we have used the following:

$$\mathcal{C}^{-1}\varphi(x)\mathcal{C} = \varphi^\dagger(x) , \quad \mathcal{C}^{-1}\varphi^\dagger(x)\mathcal{C} = \varphi(x) . \quad (10.27)$$

From these, we immediately obtain

$$\mathcal{C}^{-1}\Phi = \mathcal{P}\Phi , \quad \mathcal{C}\mathcal{P}\Phi = \Phi , \quad (10.28)$$

i.e., the  $\pi^+ - \pi^-$  system is in an eigenstate of  $\mathcal{C}\mathcal{P}$  with eigenvalue 1. Assuming that a neutral K meson is an eigenstate of  $\mathcal{C}\mathcal{P}$ , it has two components, one corresponding to the eigenvalue +1 and one corresponding to the eigenvalue -1. These are denoted by  $K_S^0$  or  $K_L^0$ , respectively. The former has a short lifetime ( $\sim 10^{-10}$  s), and the latter has a long lifetime ( $\sim 5 \times 10^{-8}$  s), as indicated by the subscripts S and L, respectively. Thus, if  $\mathcal{C}\mathcal{P}$  is conserved, the component with

a long lifetime cannot decay into two  $\pi$  mesons. However, in 1964 Christenson, Cronin, Fitch, and Turlay observed this process, and thereby demonstrated that  $\mathcal{CP}$  is not conserved [108]:

$$K_L^0 \longrightarrow \pi^+ + \pi^- . \quad (10.29)$$

### $^1S$ and $^3S$ Positronium States

An electron and positron can form a bound state called *positronium*, which is unstable and decays into a number of photons. Writing this state as  $\Phi$ , the parity is determined by the quantum  $L$  of their relative orbital angular momentum, i.e.,

$$\mathcal{P}\Phi = (-1)^L \epsilon(e^-)\epsilon(e^+)\Phi = (-1)^{L+1}\Phi , \quad (10.30)$$

where we have used Theorem 10.1. Therefore, in the  $S$ -state ( $L = 0$ ), the parity is odd ( $-1$ ). In order to investigate the eigenvalues of the charge conjugation, we write  $\Phi$  in the form

$$\Phi = a^\dagger(\mathbf{p}, \sigma_z, e)a^\dagger(-\mathbf{p}, \sigma'_z, -e)|0\rangle , \quad (10.31)$$

where  $a^\dagger$  is the creation operator for a particle whose momentum, spin, and electric charge are specified. Therefore,

$$\begin{aligned} \mathcal{C}a^\dagger(\mathbf{p}, \sigma_z, e)a^\dagger(-\mathbf{p}, \sigma'_z, -e)|0\rangle &= a^\dagger(\mathbf{p}, \sigma_z, -e)a^\dagger(-\mathbf{p}, \sigma'_z, e)|0\rangle \\ &= -a^\dagger(-\mathbf{p}, \sigma'_z, e)a^\dagger(\mathbf{p}, \sigma_z, -e)|0\rangle \\ &= -P_x P_\sigma \Phi , \end{aligned} \quad (10.32)$$

where  $P_x$  and  $P_\sigma$  are operators that swap the coordinates and spins, respectively, of the two particles. The two-particle system can be in a spin triplet state  $S = 1$  or a spin singlet state  $S = 0$ . Since these correspond to a symmetric spin state and an anti-symmetric spin state, respectively, we have

$$P_x \Phi = (-1)^L \Phi , \quad P_\sigma \Phi = (-1)^{S+1} \Phi . \quad (10.33)$$

Therefore, inserting (10.33) into (10.32),

$$\mathcal{C}\Phi = (-1)^{L+S} \Phi . \quad (10.34)$$

Of course, the positronium state is only ever realized as a sum over states like (10.31) with a variety of different values of  $\mathbf{p}$ . Thus, for positronium states  $^3S$  and  $^1S$ ,

$$\mathcal{C}\Phi(^3S) = -\Phi(^3S) , \quad \mathcal{C}\Phi(^1S) = \Phi(^1S) . \quad (10.35)$$



The former is a vector and the latter is a pseudoscalar. The decay pattern predicted from Furry's theorem is

$${}^3S \rightarrow 3\gamma, \quad {}^1S \rightarrow 2\gamma. \quad (10.36)$$

### 10.3 Isospin Parity and $G$ -Conjugation

We define the isospin parity  $T$  as a rotation in isospin space:

$$T = \exp(i\pi I_2). \quad (10.37)$$

Clearly,

$$T^{-1} I_3 T = -I_3. \quad (10.38)$$

Since  $I_3$  is half of the difference between the number of protons and the number of neutrons in a nucleus, it is diagonalized. Thus,  $T$  can only be simultaneously diagonalized when the eigenvalue of  $I_3$  for a certain state is zero. In this case, the quantum number  $I$  is an integer, so all eigenvalues of  $I_2$  become integers. Thus, it turns out that we can make the replacement

$$T^2 = \exp(2\pi i I_2) \longrightarrow 1. \quad (10.39)$$

In this case, the eigenvalues of  $T$  are

$$T = \pm 1, \quad (10.40)$$

which is the isospin parity of the given nucleus. This quantity can be defined only for a nucleus in which the numbers of protons and neutrons are the same.

For a system with one nucleus,  $T$  has the representation

$$T = \exp\left(\frac{1}{2}i\pi\tau_2\right) = i\tau_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (10.41)$$

Using this matrix, we can swap the proton and the neutron.

**Theorem 10.2** *The Kroll–Foldy theorem. The isospin parity of the state with  $I_3 = 0$  is given by*

$$T = (-1)^I. \quad (10.42)$$

**Proof** Considering the isospin eigenstate  $|I', I'_3\rangle$ , the theorem states that

$$T|I', 0\rangle = (-1)^{I'}|I', 0\rangle . \quad (10.43)$$

To prove this, it is enough to prove the algebraically equivalent equation

$$e^{i\pi L_y} Y_{l,0} = (-1)^l Y_{l,0} , \quad (10.44)$$

where  $Y_{l,0}$  are spherical harmonics, i.e., common eigenstates of  $I^2$  and  $I_3$ . However, the operator corresponding to  $T$  is a rotation through  $\pi$  radians about the  $y$  axis, so

$$x \rightarrow -x , \quad y \rightarrow y , \quad z \rightarrow -z . \quad (10.45)$$

Thus,

$$\cos \theta \rightarrow -\cos \theta . \quad (10.46)$$

Replacing  $Y_{l,0}$  by the Legendre polynomials  $P_l(\cos \theta)$ , we obtain the same equation as (10.44):

$$e^{i\pi L_y} P_l(\cos \theta) = P_l(-\cos \theta) = (-1)^l P_l(\cos \theta) . \quad (10.47)$$

■

According to (10.45), the three components of the isospin transform under  $T$  as

$$T^{-1} I_1 T = -I_1 , \quad T^{-1} I_2 T = I_2 , \quad T^{-1} I_3 T = -I_3 . \quad (10.48)$$

Transforming the field theoretical expression (7.46) for the isospin of nucleons under charge conjugation, we obtain

$$\mathcal{C}^{-1} \hat{I} \mathcal{C} = \mathcal{C}^{-1} \left( \frac{1}{2} \int d^3x \psi^\dagger \hat{\tau} \psi \right) \mathcal{C} = -\frac{1}{2} \int d^3x \psi^\dagger \hat{\tau}^T \psi . \quad (10.49)$$

Now, using the Pauli representation for  $\hat{\tau}$ ,

$$\mathcal{C}^{-1} I_1 \mathcal{C} = -I_1 , \quad \mathcal{C}^{-1} I_2 \mathcal{C} = I_2 , \quad \mathcal{C}^{-1} I_3 \mathcal{C} = -I_3 . \quad (10.50)$$

This result also holds for the isospin of the meson. Comparing this with (10.48), the product  $\mathcal{C}T$  commutes with the three isospin components. Combining  $\mathcal{C}$  and  $T$  into  $G$ , we define  $G$ -conjugation by

$$G = \mathcal{C}T = \mathcal{C}e^{i\pi I_2} . \quad (10.51)$$

**Table 10.1** Transformation properties of quantities specifying particle states

	Angular momentum	Momentum	Coordinate	Charge
Charge conjugation	+	−	−	+
Space inversion	+	−	−	+
Time reversal	−	−	+	+

Therefore,

$$G^{-1} \hat{I} G = \hat{I} . \tag{10.52}$$

For the meson field,

$$G^{-1} \hat{\phi} G = -\hat{\phi} . \tag{10.53}$$

As just described, all isospin states of the  $\pi$  meson are eigenstates of  $G$  with eigenvalue  $-1$ . This is useful, since in contrast with the charge conjugation case it can be used not only for the neutral case but also for the charged states. In this regard, however, the law of conservation of isospin is true only for the nuclear force. It is broken by electromagnetic interactions, and the same goes for  $G$ -invariance.

On the other hand, in the case of strong interactions, Furry’s theorem holds for  $G$ -invariance:

$$\begin{array}{ccccccc} \rho & \rightarrow & \pi & + & \pi & , & \rho \not\rightarrow \eta + \pi , \\ G & + & - & - & + & + & - . \end{array}$$

The former reaction is realized and the latter is forbidden. This is thus another example of Furry’s theorem.

Although the transformations discussed above, namely, space inversion, charge conjugation, and  $G$ -conjugation, are unitary transformations, time reversal is an *antiunitary transformation*. This is the subject of the next section. For later convenience, Table 10.1 shows how several quantities used to specify elementary particle states are affected by the various transformations. This table will be useful in the next section.

## 10.4 Antiunitary Transformations

Charge conjugation and space inversion discussed in the previous section are unitary transformations, but time reversal is an antiunitary transformation. Denoting charge conjugation, space inversion, and time reversal by  $C$ ,  $P$ , and  $T$ , respectively, several

kinds of product transformation can be defined and categorized as follows:

$$\begin{aligned} &\text{unitary transformations} \quad C, P, CP, \\ &\text{antiunitary transformations} \quad T, CT, PT, CPT. \end{aligned}$$

We first discuss the general characteristics of unitary and antiunitary transformations:

1. Unitary transformation. This is defined by

$$\Phi \rightarrow \Phi' = \mathcal{U} \Phi, \quad (10.54)$$

where  $\mathcal{U}$  is a unitary operator. The inner product of two state vectors is invariant, i.e., expressing the inner product by a bracket, we have

$$(\Phi', \Psi') = (\Phi, \Psi). \quad (10.55)$$

2. Antiunitary transformation. This is defined by

$$\Phi \rightarrow \Phi' = \mathcal{U}^T \Phi^* \equiv \Phi^*, \quad (10.56)$$

which is clearly a generalization of complex conjugation and has thus been indicated by the notation  $\star$ . In this case,

$$(\Phi', \Psi') = (\Phi^*, \Psi^*) = (\Psi, \Phi). \quad (10.57)$$

This equation clearly holds true because  $\star$  is a generalization of complex conjugation. This transformation clearly converts a ket vector into a bra vector.

We can also define the transformation of an operator by

$$(\Phi^*, Q\Psi^*) = (\Psi, Q^*\Psi). \quad (10.58)$$

We can thus choose to apply the antiunitary transformation to the state vector or to the operator, which corresponds to the difference between the Schrödinger picture and the Heisenberg picture, where we choose to put the time dependence on the state vector or on the operator. Since in the Heisenberg picture the form of the equation of motion is the same for both the quantum and classical theories, the latter choice is useful in the sense that knowledge of classical theory can be used. Moreover, it is straightforward to show that

$$Q^* = \mathcal{U} Q^T \mathcal{U}^{-1}. \quad (10.59)$$

Hence, for operator transformations, we have

$$(aA + bB)^* = aA^* + bB^* \quad (\text{linearity}) \quad (10.60)$$

and

$$(AB)^* = B^*A^* , \quad (10.61)$$

where  $a$  and  $b$  are  $c$ -numbers. This linearity does not hold for transformations of state vectors, where coefficients are replaced by their complex conjugates. This is another advantage of considering operator transformations.

**Theorem 10.3** *If an operator  $Q$  corresponding to a physical quantity satisfies  $Q^* = \pm Q$  and if  $\Phi$  is an eigenstate of  $Q$  with eigenvalue  $q$ , then  $\Phi^*$  is also an eigenstate of  $Q$  and its eigenvalues are  $\pm q$ .*

**Proof** We have

$$(Q\Phi^*, \Psi^*) = (\Phi^*, Q\Psi^*) = (\Psi, Q^*\Phi) = \pm(\Psi, Q\Phi) = \pm q(\Psi, \Phi) = \pm q(\Phi^*, \Psi^*) .$$

Therefore, since  $\Psi^*$  is an arbitrary state,

$$Q\Phi^* = \pm q\Phi^* .$$

■

If a system is invariant under an antiunitary transformation  $\star$ , then we have

$$S^* = S . \quad (10.62)$$

In the Schrödinger picture, we decompose the Hamiltonian into the free part and the interacting part:

$$H = H_f + H_{\text{int}} . \quad (10.63)$$

The condition for invariance (10.62) is satisfied if

$$H_f^* = H_f , \quad H_{\text{int}}^* = H_{\text{int}} . \quad (10.64)$$

In order to prove this, we change to the interaction picture. In this picture, the interaction Hamiltonian is

$$H_{\text{int}}(t) = e^{itH_f} H_{\text{int}} e^{-itH_f} . \quad (10.65)$$

Thus,

$$\begin{aligned}
 H_{\text{int}}(t)^\star &= e^{-itH_f^\star} H_{\text{int}}^\star e^{itH_f^\star} \\
 &= e^{-itH_f} H_{\text{int}} e^{itH_f} \\
 &= H_{\text{int}}(-t) .
 \end{aligned} \tag{10.66}$$

This will be our starting point. We write Dyson's formula in the form

$$S = T_t \exp \left[ -i \int dt H_{\text{int}}(t) \right] , \tag{10.67}$$

where  $T_t$  indicates that we take into account the ordering with respect to the parameter  $t$ . Since the ordering is reversed when we apply  $\star$ , we introduce the operator  $\tilde{T}_t$  :

$$\begin{aligned}
 S^\star &= \tilde{T}_t \exp \left[ -i \int dt H_{\text{int}}(t)^\star \right] \\
 &= \tilde{T}_t \exp \left[ -i \int dt H_{\text{int}}(-t) \right] \\
 &= T_{t'} \exp \left[ -i \int dt' H_{\text{int}}(t') \right] = S ,
 \end{aligned} \tag{10.68}$$

where we have used  $t' = -t$  and  $\tilde{T}_t = T_{-t}$ .

Although invariance under the antiunitary transformation does not imply the selection rule, the principle of detailed balance can be derived, as discussed in Sect. 6.4. Equation (10.68) implies that

$$(\Phi^\star, S\Psi^\star) = (\Psi, S\Phi) , \tag{10.69}$$

so summing over spin states,

$$\sum_{S_i} \sum_{S_f} |\mathcal{T}(i \rightarrow f)|^2 = \sum_{S_i} \sum_{S_f} |\mathcal{T}(f^\star \rightarrow i^\star)|^2 . \tag{10.70}$$

Combining this with (9.27), we obtain a relation between differential cross-sections for inverse processes:

1. Invariance under  $T$  and  $TP$  :

$$\frac{\frac{d}{d\Omega} \sigma(a + b \rightarrow c + d)}{\frac{d}{d\Omega} \sigma(c + d \rightarrow a + b)} = \frac{(2S_c + 1)(2S_d + 1)}{(2S_a + 1)(2S_b + 1)} \left( \frac{p_{cd}}{p_{ab}} \right)^2 , \tag{10.71}$$

where  $S_a$  stands for the spin of particle  $a$ , etc., and  $p_{ab}$  ( $p_{cd}$ ) are the moduli of the relative momenta between particles  $a$  and  $b$  ( $c$  and  $d$ ) in the center-of-mass systems. Equation (6.88) used to determine the spin of the  $\pi$  meson or pion is a special case of (10.71), derived on the assumption of invariance under the time reversal of strong interactions.

2. Invariance under  $CT$  and  $CPT$  :

$$\frac{\frac{d}{d\Omega}\sigma(a+b \rightarrow c+d)}{\frac{d}{d\Omega}\sigma(\bar{c}+\bar{d} \rightarrow \bar{a}+\bar{b})} = \frac{(2S_c+1)(2S_d+1)}{(2S_a+1)(2S_b+1)} \left(\frac{p_{cd}}{p_{ab}}\right)^2. \quad (10.72)$$

## 10.5 CPT Theorem

We now discuss the CPT theorem or Pauli–Lüders theorem, a theorem about antiunitary transformations which always holds true in a local field theory. This theorem was given by Lüders in [109] and by Pauli in [110]. It attracted a lot of attention in 1956, when it was discovered that parity is not always conserved.

In this section,  $\star$  stands for the  $CPT$  transformation. It is defined by the following equations:

$$\psi(x)^\star = i\gamma_5\psi(-x), \quad \bar{\psi}(x)^\star = i\bar{\psi}(-x)\gamma_5 \quad (\text{spinor}), \quad (10.73)$$

$$\varphi_{\lambda_1\dots\lambda_n}(x)^\star = (-1)^n\varphi_{\lambda_1\dots\lambda_n}(-x) \quad (\text{tensor}). \quad (10.74)$$

Here the field operators are in time-dependent representations, such as the Heisenberg picture or the interaction picture.

**Theorem 10.4** *CPT theorem. Any local field theory is invariant under the CPT transformation.*

The content of this theorem can be expressed by

$$\mathcal{L}(x)^\star = \mathcal{L}(-x), \quad (10.75)$$

or more specifically by

$$\mathcal{L}_f(x)^\star = \mathcal{L}_{\text{int}}(-x), \quad \mathcal{L}_{\text{int}}(x)^\star = \mathcal{L}_{\text{int}}(-x). \quad (10.76)$$

To prove this, we consider a simple Lagrangian density:

$$\mathcal{L}_{\text{int}}(x) = g\bar{\psi}_{\alpha_1}(x) \dots \bar{\psi}_{\alpha_n}(x) O_{\alpha_1\dots\alpha_n\beta_1\dots\beta_n}^{\lambda_1\dots\lambda_n} \psi_{\beta_1}(x) \dots \psi_{\beta_n}(x) \varphi_{\lambda_1\dots\lambda_m}(x), \quad (10.77)$$

where we have assumed that there are as many occurrences of  $\bar{\psi}$  as of  $\psi$ , and  $O$  is a product of the Dirac matrices, e.g., if  $m = n$ ,

$$(\gamma_{\lambda_1})_{\alpha_1\beta_1} \cdots (\gamma_{\lambda_n})_{\alpha_n\beta_n} (1 + \gamma_5) . \quad (10.78)$$

Furthermore,  $\varphi_{\lambda_1 \dots \lambda_m}$  is a bosonic part and if  $\lambda_1, \dots, \lambda_m$  are vector indices, it can be a product of multi-operators, e.g.,

$$\varphi_{\alpha\beta\gamma} = \varphi_{\alpha}(x)\varphi_{\beta}(x) \frac{\partial\varphi(x)}{\partial x_{\gamma}} . \quad (10.79)$$

Clearly, however, we always have

$$\varphi_{\lambda_1 \dots \lambda_m}(x)^* = (-1)^m \varphi_{\lambda_1 \dots \lambda_m}(-x) . \quad (10.80)$$

The order of spinor operators is reversed under  $CPT$ . Assuming that the bosonic part is defined by the normal product, the ordering makes no difference under  $CPT$ . Restoring the order of the spinor part introduces a sign factor  $(-1)^n$ . Thus, in order to prove (10.75), we must prove that

$$(-1)^{m+n} (i\gamma_5)_{\alpha_1\alpha'_1} \cdots (i\gamma_5)_{\alpha_n\alpha'_n} O_{\alpha'_1 \dots \alpha'_n \beta'_1 \dots \beta'_n}^{\lambda_1 \dots \lambda_m} (i\gamma_5)_{\beta'_1\beta_1} \cdots (i\gamma_5)_{\beta'_n\beta_n} = O_{\alpha'_1 \dots \alpha'_n \beta'_1 \dots \beta'_n}^{\lambda_1 \dots \lambda_m} . \quad (10.81)$$

To do this, we use the following:

$$\begin{aligned} i\gamma_5(1 \text{ or } \gamma_5)i\gamma_5 &= -(1 \text{ or } \gamma_5) && (n - m) \text{ combinations,} \\ i\gamma_5(\gamma_{\mu} \text{ or } \gamma_{\mu}\gamma_5)i\gamma_5 &= \gamma_{\mu} \text{ or } \gamma_{\mu}\gamma_5 && m \text{ combinations.} \end{aligned}$$

Therefore, the change in the overall sign is  $(-1)^{n-m} = (-1)^{n+m}$ , which proves (10.81).

Next, we consider the case where the number of occurrences of  $\psi$  and  $\bar{\psi}$  in the interaction Lagrangian are different. For example,

$$\mathcal{L}_{\text{int}}(x) = g \bar{\psi}(x) i\gamma_5 C \bar{\psi}^T(x) \varphi(x) + \text{h.c.} \quad (10.82)$$

It is clear from the discussion about charge conjugation that  $C\bar{\psi}^T$  transforms in the same way as  $\psi$  under a proper Lorentz transformation. Hence, (10.82) is invariant under a proper Lorentz transformation. Then, taking

$$\psi'(x) = C \bar{\psi}^T(x) , \quad (10.83)$$



the fields  $\psi'$  and  $\psi$  transform in the same way under the CPT transformation. This can be proven as follows:

$$\begin{aligned}\psi'(x)^\star &= [C\bar{\psi}^\text{T}(x)]^\star = iC[\bar{\psi}(-x)\gamma_5]^\text{T} = iC\gamma_5^\text{T}\bar{\psi}^\text{T}(-x) \\ &= i\gamma_5 C\bar{\psi}^\text{T}(-x) = i\gamma_5\psi'(-x).\end{aligned}\tag{10.84}$$

Thus, the assumption that the numbers of  $\psi$  and  $\bar{\psi}$  are the same is not necessary if one rewrites  $\bar{\psi}$  as  $\psi'$ . Let us examine what we have said above regarding the transformation properties of  $C\bar{\psi}^\text{T}$ .

### Transformation Properties of $C\bar{\psi}^\text{T}$

In Sect. 3.6, we looked carefully at the way  $\psi$  transforms under proper Lorentz transformations. Now we investigate the transformation of  $C\bar{\psi}^\text{T}$ . From (3.116),

$$C\bar{\psi}^\text{T} \rightarrow C(\bar{\psi}S^{-1})^\text{T} = C(S^\text{T})^{-1}\bar{\psi}^\text{T}.\tag{10.85}$$

We consider first an infinitesimal transformation (3.119), whence

$$S^\text{T} = 1 + \frac{1}{2}\epsilon_{\mu\nu}T_{\mu\nu}^\text{T}.\tag{10.86}$$

Using (3.121),

$$CS^\text{T}C^{-1} = 1 + \frac{1}{2}\epsilon_{\mu\nu}CT_{\mu\nu}^\text{T}C^{-1} = 1 - \frac{1}{2}\epsilon_{\mu\nu}T_{\mu\nu} = S^{-1}.\tag{10.87}$$

Thus, since this relation should also hold true for a finite proper Lorentz transformation, we have

$$C(S^\text{T})^{-1}\bar{\psi}^\text{T} = C(S^\text{T})^{-1}C^{-1}C\bar{\psi}^\text{T} = S(C\bar{\psi}^\text{T}),\tag{10.88}$$

i.e., for a proper Lorentz transformation,

$$C\bar{\psi}^\text{T} \rightarrow S(C\bar{\psi}^\text{T}).\tag{10.89}$$

This shows that it obeys the same transformation law as  $\psi$ .

In contrast, for space inversion, (10.6) implies

$$C\bar{\psi}^\text{T} \rightarrow \epsilon C(\bar{\psi}\gamma_4)^\text{T} = \epsilon C\gamma_4^\text{T}\bar{\psi}^\text{T} = -\epsilon\gamma_4(C\bar{\psi}^\text{T}).\tag{10.90}$$

The sign is different compared with the transformation of  $\psi$  under space inversion. This confirms Theorem 10.1 in Sect. 10.2.

### Origin of the Name ‘CPT Transformation’

We now prove that the combination of the three transformations  $C$ ,  $P$ , and  $T$  yields  $\star$ . Since for each transformation the phase is not necessarily determined, we apply the simplest one. Since time reversal swaps the creation and annihilation of a particle, it turns out that  $\psi \rightleftharpoons \bar{\psi}$  and  $t \rightarrow -t$ . Furthermore, the sign of the spatial components of the four-current  $i\bar{\psi}\gamma_\mu\psi$  is reversed. Thus, the time reversal of the spinor becomes

$$T: \quad \psi(\mathbf{x}, t) \rightarrow \bar{\psi}(\mathbf{x}, -t)\gamma_4\gamma_5C, \quad \bar{\psi}(\mathbf{x}, t) \rightarrow C^{-1}\gamma_5\gamma_4\psi(\mathbf{x}, -t), \quad (10.91)$$

up to a phase. Next, consecutively carrying out space inversion and then charge conjugation,

$$PT: \quad \psi(\mathbf{x}, t) \rightarrow \bar{\psi}(-\mathbf{x}, -t)\gamma_5C, \quad \bar{\psi}(\mathbf{x}, t) \rightarrow C^{-1}\gamma_5\psi(-\mathbf{x}, -t), \quad (10.92)$$

and

$$CPT: \quad \psi(\mathbf{x}, t) \rightarrow (\gamma_5C)^T C^{-1}\psi(-\mathbf{x}, -t) = -\gamma_5\psi(-\mathbf{x}, -t). \quad (10.93)$$

Similarly,

$$CPT: \quad \bar{\psi}(\mathbf{x}, t) \rightarrow \bar{\psi}(-\mathbf{x}, -t)\gamma_5. \quad (10.94)$$

Apart from the undetermined factor, this coincides with  $\star$  in (10.73). This is the reason why  $\star$  is called the CPT transformation. Although the CPT transformation in (10.73) is uniquely determined, the phase of each transformation is not necessarily uniquely determined.

From the fact that the product of these three transformations is always invariant, it follows that, if one of them does not make the theory invariant, at least one more must also break the invariance. The non-conservation of parity was discovered in 1956, implying the violation of invariance under space inversion. It then turned out that charge conjugation invariance was also broken. At the time, it was expected that invariance under  $CP$  would hold. But as already mentioned when discussing the decay pattern in (10.29), this too turned out to be incorrect.

### Transformation Properties of the Hamiltonian

The energy–momentum four-vector is

$$P_\nu = \int d^3x \mathcal{T}_{0\nu}(x). \quad (10.95)$$

Under the CPT transformation,

$$\mathcal{T}_{\mu\nu}(x)^* = \mathcal{T}_{\mu\nu}(-x) . \quad (10.96)$$

Thus,

$$P_v^* = P_v . \quad (10.97)$$

At  $t = 0$  or in the Schrödinger picture, this implies

$$H_f^* = H_f , \quad H_{\text{int}}^* = H_{\text{int}} . \quad (10.98)$$

Since (10.58) can be derived from (10.64), this equation ensures that the  $S$ -matrix is invariant under the CPT transformation.

### Equality of Particle and Antiparticle Masses

We use Theorem 10.2 in Sect. 10.3. Assuming that  $\Phi$  is a state containing a particle of mass  $m$ ,  $\Phi^*$  is a state containing its antiparticle, i.e.,

$$(P_v^2 + m^2)\Phi = 0 . \quad (10.99)$$

Therefore, from (10.97),

$$(P_v^2 + m^2)\Phi^* = 0 . \quad (10.100)$$

These equations imply that the masses of a particle and its antiparticle are the same.

### Equality of Particle and Antiparticle Lifetimes

We have

$$\mathcal{T}(a \rightarrow a) = \mathcal{T}(a^* \rightarrow a^*) , \quad (10.101)$$

due to the invariance of  $S$ . Taking the imaginary part as an application of the optical theorem to the decay process, it turns out that the total decay widths of a particle and its antiparticle are the same, whence their lifetimes are also the same.

# Chapter 11

## Green's Functions



So far we have talked a lot about the properties of the  $S$ -matrix, but there is an important concept closely related to the  $S$ -matrix, namely, the concept of Green's functions. We first define Green's functions in the interaction picture, and then rewrite them in the Heisenberg picture. This leads to the so-called *Gell-Mann–Low relation* [111]. We then discuss Matthews' theorem [112] for the relation between the Hamiltonian formalism and the Lagrangian formalism. Finally, we turn our attention to the reduction formula connecting the  $S$ -matrix and Green's functions. This formula relates to fundamental conditions called *asymptotic conditions* in the Heisenberg picture.

### 11.1 Gell-Mann–Low Relation

As in Chap. 3, writing the free-field energy–momentum four-vector in the interaction picture as  $P_\mu^{(0)}$ , we find that (3.29) holds true for a field operator  $\mathcal{O}(x)$  without external fields, i.e.,

$$[P_\mu^{(0)}, \mathcal{O}(x)] = i \frac{\partial}{\partial x_\mu} \mathcal{O}(x). \quad (11.1)$$

In particular, for  $\mu = 0$ ,

$$[P_0^{(0)}, \mathcal{O}(x)] = -i \frac{\partial}{\partial t} \mathcal{O}(x). \quad (11.2)$$

We now introduce the following transformation function as  $\mathcal{O}$  :

$$U(t, t_0) = T \exp \left[ -i \int_{t_0}^t dt' H_{\text{int}}(t') \right]. \quad (11.3)$$

We insert this into (11.2), but bearing in mind that

$$\begin{aligned} \left[ P_0^{(0)}, -i \int_{t_0}^t dt' H_{\text{int}}(t') \right] &= - \int_{t_0}^t dt' \frac{\partial}{\partial t'} H_{\text{int}}(t') \\ &= -[H_{\text{int}}(t) - H_{\text{int}}(t_0)]. \end{aligned} \quad (11.4)$$

This immediately yields

$$\begin{aligned} [P_0^{(0)}, U(t, t_0)] &= -T[[H_{\text{int}}(t) - H_{\text{int}}(t_0)]U(t, t_0)] \\ &= -H_{\text{int}}(t)U(t, t_0) + U(t, t_0)H_{\text{int}}(t_0). \end{aligned} \quad (11.5)$$

We now set  $t = 0$  and take the limit  $t_0 \rightarrow -\infty$ . In this case, putting

$$P_0^{(0)} + H_{\text{int}}(0) = H_{\text{total}}, \quad (11.6)$$

we obtain

$$H_{\text{total}}U(0, -\infty) = U(0, -\infty)[P_0^{(0)} + H_{\text{int}}(-\infty)].$$

Assuming that interactions exist only in a finite space-time region, as in Sects. 4.4 and 7.3, we consider the limit as this space-time region extends to infinity. This corresponds to introducing interactions adiabatically, and a variety of consequences depend on how the limit is actually taken. With some loss of rigour, we assume that

$$H_{\text{int}}(-\infty) = 0, \quad (11.7)$$

whence

$$H_{\text{total}}U(0, -\infty) = U(0, -\infty)P_0^{(0)}. \quad (11.8)$$

Then writing the vacuum in the interaction picture as  $\Phi_0$ ,

$$P_0^{(0)}\Phi_0 = 0. \quad (11.9)$$

Hence, writing the vacuum in the Heisenberg picture as  $\Psi_0$ , we find

$$H_{\text{total}}\Psi_0 = 0, \quad \Psi_0 = U(0, -\infty)\Phi_0. \quad (11.10)$$

As shown in Chap. 6, the Heisenberg picture and the interaction picture are related by (6.30), viz.,

$$U(t, t_0)^{-1}\varphi_\alpha(\mathbf{x}, t)U(t, t_0) = \varphi_\alpha^{(\text{H})}(\mathbf{x}, t). \quad (11.11)$$

In this section, we set  $t_0 = 0$ . Therefore, for  $t_1 > t_2$ ,

$$\begin{aligned}
 & (\Phi_0, T[U(\infty, -\infty)A(x_1, t_1)B(x_2, t_2)]\Phi_0) \\
 &= (\Phi_0, U(\infty, t_1)A(x_1, t_1)U(t_1, t_2)B(x_2, t_2)U(t_2, -\infty)\Phi_0) \\
 &= (\Phi_0, U(\infty, 0)A^{(H)}(x_1, t_1)B^{(H)}(x_2, t_2)U(0, -\infty)\Phi_0) \\
 &= (\Phi_0, U(\infty, -\infty)U(0, -\infty)^{-1}A^{(H)}(x_1, t_1)B^{(H)}(x_2, t_2)U(0, -\infty)\Phi_0) \\
 &= (\Phi_0, U(\infty, -\infty)\Phi_0)(\Phi_0, U(0, -\infty)^{-1}A^{(H)}(x_1, t_1)B^{(H)}(x_2, t_2)U(0, -\infty)\Phi_0) \\
 &= (\Phi_0, U(\infty, -\infty)\Phi_0)(\Psi_0, A^{(H)}(x_1, t_1)B^{(H)}(x_2, t_2)\Psi_0), \tag{11.12}
 \end{aligned}$$

where we have used the composition rule (6.27) for the transformation function and  $U(\infty, -\infty)$ , and the fact that the vacuum  $\Phi_0$  transforms to the vacuum up to a phase. This can be written in the form

$$(\Psi_0, T[A^{(H)}(x_1)B^{(H)}(x_2)]\Psi_0) = \frac{(\Phi_0, T[U(\infty, -\infty)A(x_1)B(x_2)]\Phi_0\Phi_0)}{(\Phi_0, U(\infty, -\infty)\Phi_0)}. \tag{11.13}$$

The presence of the denominator corresponds to neglecting bubble graphs, as discussed in connection with (8.70). This equation is easily generalized:

$$\begin{aligned}
 & (\Psi_0, T[A^{(H)}(x_1)B^{(H)}(x_2)\dots Z^{(H)}(x_n)]\Psi_0) \tag{11.14} \\
 &= \frac{(\Phi_0, T[U(\infty, -\infty)A(x_1)B(x_2)\dots Z(x_n)]\Phi_0)}{(\Phi_0, U(\infty, -\infty)\Phi_0)}.
 \end{aligned}$$

This is called the *Gell-Mann–Low relation* [111]. It gives the relationship between the Heisenberg picture and the interaction picture. Additionally, any quantity like the left-hand side, i.e., a vacuum expectation value of a product of time-ordered field operators, is called a *Green's function*.

## 11.2 Green's Functions and Their Generating Functionals

When we discuss properties of a Green's function, instead of considering them one by one, it is sometimes useful to discuss their generating function, or more precisely their generating functional. In this section, we describe the simplest generating functional, which produces the Green's functions of a neutral scalar field.

We consider the Lagrangian density

$$\mathcal{L} = -\frac{1}{2}[(\partial_\mu\varphi)^2 + m^2\varphi^2] - \frac{g}{4!}\varphi^4. \tag{11.15}$$

We split this into the free part and the interaction part. In fact, such a separation is not trivial, and becomes particularly complicated when we take into account renormalization in the next chapter. However, in this section, we simply choose

$$\mathcal{H}_{\text{int}}(x) = -\mathcal{L}_{\text{int}}(x) = \frac{g}{4!}\varphi^4. \quad (11.16)$$

In the following, we shall express the vacuum  $\Phi_0$  simply as  $|0\rangle$ . We introduce a  $c$ -number external field  $J(x)$  and define the functional

$$\mathcal{F}^{(0)}[J] = \frac{\langle 0|T \exp \left\{ -i \int d^4x [\mathcal{H}_{\text{int}}(x) + J(x)\varphi(x)] \right\} |0\rangle}{\langle 0|T \exp \left[ -i \int d^4x \mathcal{H}_{\text{int}}(x) \right] |0\rangle}. \quad (11.17)$$

We now derive the equation satisfied by the generating functional of the Green's functions  $\mathcal{F}^{(0)}[J]$ . To do this, we just have to use the reduction formula (8.57), i.e.,

$$(\square_x - m^2)T[\varphi(x)AB \dots] = i \frac{\delta}{\delta\varphi(x)} T[AB \dots]. \quad (11.18)$$

Writing the denominator of (11.17) as  $\langle 0|U(\infty, -\infty)|0\rangle$ ,

$$\begin{aligned} (\square_x - m^2) i \frac{\delta}{\delta J(x)} \mathcal{F}^{(0)}[J] &= \frac{(\square_x - m^2) \langle 0|T(\varphi(x) \exp \{ -i \int d^4y [\mathcal{H}_{\text{int}}(y) + J(y)\varphi(y)] \}) |0\rangle}{\langle 0|U(\infty, -\infty)|0\rangle} \\ &= \frac{\langle 0| i \frac{\delta}{\delta\varphi(x)} T(\exp \{ -i \int d^4y [\mathcal{H}_{\text{int}}(y) + J(y)\varphi(y)] \}) |0\rangle}{\langle 0|U(\infty, -\infty)|0\rangle} \\ &= \frac{\langle 0|T \left( \left[ \frac{g}{3!}\varphi^3(x) + J(x) \right] \exp \{ -i \int d^4y [\mathcal{H}_{\text{int}}(y) + J(y)\varphi(y)] \} \right) |0\rangle}{\langle 0|U(\infty, -\infty)|0\rangle} \\ &= \left\{ \frac{g}{3!} \left[ i \frac{\delta}{\delta J(x)} \right]^3 + J(x) \right\} \mathcal{F}^{(0)}[J]. \end{aligned} \quad (11.19)$$

Integrating this equation, we obtain the functional equation for  $\mathcal{F}^{(0)}[J]$ :

$$i \frac{\delta}{\delta J(x)} \mathcal{F}^{(0)}[J] = -i \int d^4y \Delta_F(x-y) \left\{ \frac{g}{3!} \left[ i \frac{\delta}{\delta J(y)} \right]^3 + J(y) \right\} \mathcal{F}^{(0)}[J]. \quad (11.20)$$

Differentiating  $\mathcal{T}^{(0)}[J]$   $n$  times with respect to  $J$  and setting  $J = 0$ , we obtain the  $n$ -point Green's function:

$$i^n \frac{\delta^n}{\delta J(x_1) \dots \delta J(x_n)} \mathcal{T}^{(0)}[J] \Big|_{J=0} = \frac{\langle 0|T[\varphi(x_1) \dots \varphi(x_n)U(\infty, -\infty)]|0\rangle}{\langle 0|U(\infty, -\infty)|0\rangle}. \quad (11.21)$$

We now turn to the Heisenberg picture, defining the generating functional by

$$\mathcal{T}[J] = \langle \mathbf{0}|T \exp \left[ -i \int d^4x J(x)\varphi(x) \right]|\mathbf{0}\rangle, \quad (11.22)$$

where we have written  $\Psi_0$  and  $\varphi^{(H)}$  in bold face as  $|\mathbf{0}\rangle$  and  $\varphi$ , respectively. Therefore,

$$(\square_x - m^2)i \frac{\delta}{\delta J(x)} \mathcal{T}[J] = (\square_x - m^2)\langle \mathbf{0}|T \left\{ \varphi(x) \exp \left[ -i \int d^4y J(y)\varphi(y) \right] \right\}|\mathbf{0}\rangle. \quad (11.23)$$

In this picture,

$$(\square_x - m^2)T[\varphi(x)\mathbf{A}\mathbf{B} \dots] = T[(\square_x - m^2)\varphi(x), \mathbf{A}\mathbf{B} \dots] + i \frac{\delta}{\delta \varphi(x)} T[\mathbf{A}\mathbf{B} \dots]. \quad (11.24)$$

Inserting this into (11.23) and using the equation of motion for  $\varphi$ , we recover (11.19). Hence, (11.20) can also be recovered, i.e.,

$$i \frac{\delta}{\delta J(x)} \mathcal{T}[J] = -i \int d^4y \Delta_F(x-y) \left\{ \frac{g}{3!} \left[ i \frac{\delta}{\delta J(y)} \right]^3 + J(y) \right\} \mathcal{T}[J]. \quad (11.25)$$

This implies that, even if the boundary conditions are the same, we can expect

$$\mathcal{T}[J] = \mathcal{T}^{(0)}[J]. \quad (11.26)$$

In fact, differentiating this  $n$  times with respect to  $J$  and setting  $J = 0$ , it becomes

$$\langle \mathbf{0}|T[\varphi(x_1) \dots \varphi(x_n)]|\mathbf{0}\rangle = \frac{\langle 0|T[\varphi(x_1) \dots \varphi(x_n)U(\infty, -\infty)]|0\rangle}{\langle 0|U(\infty, -\infty)|0\rangle}, \quad (11.27)$$

which is nothing other than the Gell-Mann–Low relation. Thus, it turns out that (11.26) always holds true.

When a Green's function is expressed by a Feynman diagram, it will generally give a set of non-connected parts. We focus only on the graphs in which  $n$  points are connected to each other by lines. We write the contribution corresponding to such a



graph as

$$\langle \mathbf{0} | T[\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle_{\text{conn}}, \quad (11.28)$$

where conn stands for 'connected.' So what is the relationship between the connected Green's function and the original Green's function? To find out, we start with some point  $x$  and separate into the part connected with  $x$  and the other parts. This gives a recursion formula:

$$\begin{aligned} \langle \mathbf{0} | T[\varphi(x)\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle \\ = \sum_{\text{comb}} \langle \mathbf{0} | T[\varphi(x)\varphi(x'_1) \dots \varphi(x'_k)] | \mathbf{0} \rangle_{\text{conn}} \langle \mathbf{0} | T[\varphi(x'_{k+1}) \dots \varphi(x'_n)] | \mathbf{0} \rangle, \end{aligned} \quad (11.29)$$

where the summation has been taken over all combinations separating  $x_1, \dots, x_n$  into  $x'_1, \dots, x'_k$  and  $x'_{k+1}, \dots, x'_n$ . In order to express this relation in a closed form, we introduce a generating functional  $\mathcal{R}[J]$  for the connected Green's functions:

$$\mathcal{R}[J] = \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \dots \int d^4x_n \langle \mathbf{0} | T[\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle_{\text{conn}} J(x_1) \dots J(x_n), \quad (11.30)$$

$$i^n \frac{\delta^n}{\delta J(x_1) \dots \delta J(x_n)} \mathcal{R}[J] \Big|_{J=0} = \langle \mathbf{0} | T[\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle_{\text{conn}}. \quad (11.31)$$

The recursion formula can now be expressed in closed form:

$$i \frac{\delta}{\delta J(x)} \mathcal{T}[J] = \left( i \frac{\delta}{\delta J(x)} \mathcal{R}[J] \right) \mathcal{T}[J]. \quad (11.32)$$

We solve this functional equation with the boundary conditions

$$\mathcal{R}[0] = 0, \quad \mathcal{T}[0] = 1. \quad (11.33)$$

The solution is

$$\mathcal{T}[J] = \exp \mathcal{R}[J]. \quad (11.34)$$

Then writing the expectation value of  $\varphi(x)$  when there exists an external field  $J$  as  $\langle \varphi(x) \rangle$ , we have

$$\langle \varphi(x) \rangle = \frac{i \frac{\delta}{\delta J(x)} \mathcal{T}[J]}{\mathcal{T}[J]} = i \frac{\delta}{\delta J(x)} \mathcal{R}[J]. \quad (11.35)$$

In some situations, it is more useful to consider  $\langle\varphi(x)\rangle$  as an independent external field instead of  $J(x)$ , so we introduce the following Legendre transformation:

$$\mathcal{F} = \mathcal{R} + i \int d^4x J(x) \langle\varphi(x)\rangle . \quad (11.36)$$

This yields

$$\begin{aligned} \delta\mathcal{F} &= \delta\mathcal{R} + i \int d^4x [J(x)\delta\langle\varphi(x)\rangle + \delta J(x)\langle\varphi(x)\rangle] \\ &= -i \int d^4x \delta J(x)\delta\langle\varphi(x)\rangle + i \int d^4x [J(x)\delta\langle\varphi(x)\rangle + \delta J(x)\langle\varphi(x)\rangle] \\ &= i \int d^4x J(x)\delta\langle\varphi(x)\rangle , \end{aligned} \quad (11.37)$$

i.e., taking  $\langle\varphi(x)\rangle$  as an independent variable instead of  $J(x)$ ,

$$\frac{\delta\mathcal{F}}{\delta\langle\varphi(x)\rangle} = iJ(x) . \quad (11.38)$$

Variable transformations like this will play an important role in the discussion of spontaneous symmetry breaking later on. Moreover, differentiating (11.35) again,

$$i \frac{\delta\langle\varphi(x)\rangle}{\delta J(y)} = - \frac{\delta^2\mathcal{R}}{\delta J(x)\delta J(y)} = \Delta_F(x, y) . \quad (11.39)$$

This is the two-point Green's function in the case where an external field exists. Differentiating (11.38),

$$\frac{\delta^2\mathcal{F}}{\delta\langle\varphi(x)\rangle\delta\langle\varphi(y)\rangle} = i \frac{\delta J(y)}{\delta\langle\varphi(x)\rangle} = -\Delta_F^{-1}(x, y) , \quad (11.40)$$

which is the inverse of the two-point Green's function.

### 11.3 Different Time-Orderings in the Lagrangian Formalism

In the previous section, we investigated the relationship between representations of Green's functions in the Heisenberg picture and in the interaction picture. In this section, we shall discuss the difference between time-orderings in the Hamiltonian formalism and the Lagrangian formalism.

For simplicity, we start with particle dynamics. Considering a one-dimensional system, we assume that

$$p(t) = m\dot{q}(t) . \quad (11.41)$$

Therefore, under the time-ordering operator  $T$ ,  $p(t)$  is ordered as an operator corresponding to time  $t$ . This is because in the Hamiltonian formalism, the operators  $q(t)$  and  $p(t)$  are treated as independent variables, so these are considered to be quantities that need no further temporal decomposition. However, in the Lagrangian formalism, only  $q(t)$  is an independent quantity, so  $\dot{q}(t)$  becomes

$$\dot{q}(t) = \lim_{\epsilon \rightarrow 0} \frac{q(t + \epsilon) - q(t)}{\epsilon} . \quad (11.42)$$

Assuming that  $\epsilon$  is small but finite, it turns out that, in the Lagrangian formalism,  $\dot{q}(t)$  is associated with two clock times in time-orderings. We denote the time-ordering operator in this treatment by  $T^*$ . It should be noted here that, in the path-integral method, the same decomposition (11.42) is used. Therefore, the method using  $T^*$  is closely related to the path-integral method. So what is the difference between using  $T$  or  $T^*$ ?

In the Hamiltonian formalism and hence under  $T$ , we assume that  $\dot{q}(t)$  can be treated as a one-clock-time quantity, as in (11.41). Now,

$$\begin{aligned} T^*[\dot{q}(t), q(t')] &= T^* \left[ \lim_{\epsilon \rightarrow 0} \frac{q(t + \epsilon) - q(t)}{\epsilon}, q(t') \right] \\ &= \lim_{\epsilon \rightarrow 0} T \left[ \frac{q(t + \epsilon) - q(t)}{\epsilon}, q(t') \right] \\ &= \frac{\partial}{\partial t} T[q(t), q(t')] . \end{aligned} \quad (11.43)$$

However, if  $q$  obeys a second order differential equation, then only  $\dot{q}$  can enter in  $T^*$ , because  $\ddot{q}$  is not independent of  $q$ . Another example is

$$T^*[\dot{q}(t), \dot{q}(t')] = \frac{\partial^2}{\partial t \partial t'} T[q(t), q(t')] . \quad (11.44)$$

The derivation of this equation is perfectly analogous to the one in the previous example. However, using  $T$ , we find

$$\begin{aligned} \frac{\partial}{\partial t} T[q(t), q(t')] &= T[\dot{q}(t), q(t')] + \frac{\partial}{\partial t} \left[ \frac{1}{2} \epsilon(t - t') \right] [q(t), q(t')] \\ &= T[\dot{q}(t), q(t')] + \delta(t - t') [q(t), q(t')] \\ &= T[\dot{q}(t), q(t')] . \end{aligned} \quad (11.45)$$

Therefore, this is the same as  $T^*$  in (11.43). However, assuming (11.41),

$$\begin{aligned} \frac{\partial}{\partial t} T[q(t), \dot{q}(t')] &= T[\dot{q}(t), \dot{q}(t')] + \delta(t - t')[q(t), \dot{q}(t')] \\ &= T[\dot{q}(t), \dot{q}(t')] + \frac{i}{m} \delta(t - t'). \end{aligned} \quad (11.46)$$

The left-hand side is equal to the right-hand side of (11.44). So setting the left-hand side of (11.46) equal to the left-hand side of (11.44),

$$T[\dot{q}(t), \dot{q}(t')] = T^*[\dot{q}(t), \dot{q}(t')] - \frac{i}{m} \delta(t - t'). \quad (11.47)$$

The difference between  $T$  and  $T^*$  becomes clear in this way.

This argument can be extended to field theory. For simplicity, we consider a neutral scalar field in the interaction picture:

$$\begin{aligned} \frac{\partial}{\partial x_\mu} T[\partial_\mu \varphi(x), \varphi(y)] &= T[\partial_\mu \varphi(x), \varphi(y)] + \frac{1}{i} \delta_{\mu 4} \delta(x_0 - y_0) [\varphi(x), \varphi(y)] \\ &= T[\partial_\mu \varphi(x), \varphi(y)]. \end{aligned} \quad (11.48)$$

In order to differentiate this one more time, we introduce a unit time-like vector  $n_\mu$ . Here we assume that  $n_1 = n_2 = n_3 = 0$  and  $n_4 = i$ :

$$\frac{\partial}{\partial y_\nu} T[\partial_\mu \varphi(x), \varphi(y)] = T[\partial_\mu \varphi(x), \varphi(y)] + i \delta_{\nu 4} \delta(x_0 - y_0) [\partial_\mu \varphi(x), \varphi(y)].$$

The second term becomes

$$\begin{aligned} \delta_{\mu 4} \delta_{\nu 4} \delta(x_0 - y_0) [\dot{\varphi}(x), \varphi(y)] &= -i \delta_{\mu 4} \delta_{\nu 4} \delta^4(x - y) \\ &= i n_\mu n_\nu \delta^4(x - y). \end{aligned}$$

Taking the vacuum expectation value,

$$\langle 0 | T[\partial_\mu \varphi(x), \partial_\nu \varphi(y)] | 0 \rangle = \frac{\partial^2}{\partial x_\mu \partial y_\nu} \Delta_F(x - y) - i n_\mu n_\nu \delta^4(x - y). \quad (11.49)$$

On the other hand,

$$\langle 0 | T^*[\partial_\mu \varphi(x), \partial_\nu \varphi(y)] | 0 \rangle = \frac{\partial^2}{\partial x_\mu \partial y_\nu} \Delta_F(x - y). \quad (11.50)$$

This difference corresponds to (11.47). Comparing the two equations above, we see that  $T^*$  is covariant and simpler. When there are derivatives of field operators in

the interaction term, the Hamiltonian density is not a scalar, but a tensor involving the time-like vector  $n$ . Although in this case both the Hamiltonian density and the contraction function depend on  $n$ , when we compute the  $S$ -matrix, the two types of  $n$ -dependence cancel out, so the final result does not depend on  $n$ . This is called *Matthew's theorem* [112].

## 11.4 Matthews' Theorem

If interactions do not include derivatives of the field operators, we have

$$\mathcal{H}_{\text{int}}(x) = -\mathcal{L}_{\text{int}}(x) . \quad (11.51)$$

Clearly, in this case,

$$S = T \exp \left[ -i \int d^4x \mathcal{H}_{\text{int}}(x) \right] = T^* \exp \left[ i \int d^4x \mathcal{L}_{\text{int}}(x) \right] . \quad (11.52)$$

This is because the difference between  $T$  and  $T^*$  does not appear anywhere here.

When derivatives are included in the interactions, the situation becomes more complicated. For example, when a charged scalar field and the electromagnetic field interact with each other,

$$\mathcal{L} = -[(\partial_\mu + ieA_\mu)\varphi^\dagger \cdot (\partial_\mu - ieA_\mu)\varphi + m^2\varphi^\dagger\varphi] + \mathcal{L}_{\text{em}} , \quad (11.53)$$

$$\mathcal{L}_{\text{int}} = -ieA_\mu(\varphi^\dagger \cdot \partial_\mu\varphi - \partial_\mu\varphi^\dagger \cdot \varphi) - e^2A_\mu^2\varphi^\dagger\varphi . \quad (11.54)$$

Thus, when derivatives of field operators are included,  $\mathcal{H}_{\text{int}}$  differs from  $-\mathcal{L}_{\text{int}}$ .

First, we decompose the Lagrangian density:

$$\mathcal{L}(x) = \mathcal{L}_{\text{f}}(x) + \mathcal{L}_{\text{int}}(x) . \quad (11.55)$$

If  $\varphi_\alpha$  is a real scalar field, the quantity canonically conjugate to  $\varphi_\alpha$  in the free field case is

$$\pi_\alpha(x) = \frac{\partial \mathcal{L}_{\text{f}}(x)}{\partial \dot{\varphi}_\alpha(x)} = \dot{\varphi}_\alpha(x) . \quad (11.56)$$

If we now consider interactions, the canonically conjugate field is

$$\pi'_\alpha(x) = \frac{\partial \mathcal{L}(x)}{\partial \dot{\varphi}_\alpha(x)} = \dot{\varphi}_\alpha(x) + \frac{\partial \mathcal{L}_{\text{int}}(x)}{\partial \dot{\varphi}_\alpha(x)} . \quad (11.57)$$

By the Yang–Feldman equation (6.34), the relation between the Heisenberg picture and the interaction picture is

$$\varphi_\alpha(x) = U(x_0, -\infty)^{-1} \varphi_\alpha^{\text{in}}(x) U(x_0, -\infty), \quad (11.58)$$

$$\pi'_\alpha(x) = U(x_0, -\infty)^{-1} \pi_\alpha^{\text{in}}(x) U(x_0, -\infty). \quad (11.59)$$

The reason why  $\pi$  transforms into  $\pi'$  is that the transformation by  $U$  should not change the canonical commutation relation. In the following, we shall drop the superscript ‘in’ on operators in the interaction picture, and express Heisenberg operators in bold face. Therefore, the Hamiltonian density is

$$\mathcal{H}(x) = \sum_\alpha \pi'_\alpha(x) \dot{\boldsymbol{\varphi}}_\alpha(x) - \mathcal{L}(\boldsymbol{\varphi}_\alpha(x), \dot{\boldsymbol{\varphi}}_\alpha(x)), \quad (11.60)$$

where we have assumed that a spatial derivative is a linear combination of the  $\boldsymbol{\varphi}_\alpha(x)$ . Then,

$$\mathcal{H}(x) = U(x_0, -\infty) \mathcal{H}(x) U(x_0, -\infty)^{-1}. \quad (11.61)$$

In this computation we use the inverse transformations of (11.58) and (11.59). Therefore,

$$\begin{aligned} U(x_0, -\infty) \dot{\boldsymbol{\varphi}}_\alpha(x) U(x_0, -\infty)^{-1} &= U(x_0, -\infty) [\boldsymbol{\pi}'_\alpha(x) - \boldsymbol{\sigma}_\alpha(x)] U(x_0, -\infty)^{-1} \\ &= \boldsymbol{\pi}_\alpha(x) - \boldsymbol{\sigma}_\alpha(x) \\ &= \dot{\boldsymbol{\varphi}}_\alpha(x) - \boldsymbol{\sigma}_\alpha(x), \end{aligned} \quad (11.62)$$

where  $\boldsymbol{\sigma}_\alpha$  is defined by

$$\boldsymbol{\sigma}_\alpha(x) = \frac{\partial \mathcal{L}_{\text{int}}(x)}{\partial \dot{\boldsymbol{\varphi}}_\alpha(x)} = -n_\mu \frac{\partial \mathcal{L}_{\text{int}}(x)}{\partial \varphi_{\alpha,\mu}(x)}. \quad (11.63)$$

Thus,

$$\mathcal{H}(x) = \sum_\alpha \pi_\alpha(x) [\pi_\alpha(x) - \sigma_\alpha(x)] - \mathcal{L}(\varphi_\alpha(x), \dot{\boldsymbol{\varphi}}_\alpha(x) - \boldsymbol{\sigma}_\alpha(x)). \quad (11.64)$$

We use a Taylor expansion for the second term. In this case, since  $\mathcal{L}$  includes at most second order terms in  $\dot{\boldsymbol{\varphi}}_\alpha(x)$  and using

$$\frac{\partial^2}{\partial \dot{\boldsymbol{\varphi}}_\alpha \partial \dot{\boldsymbol{\varphi}}_\beta} \mathcal{L} = \delta_{\alpha\beta}, \quad \frac{\partial}{\partial \dot{\boldsymbol{\varphi}}_\alpha} \mathcal{L} = \pi_\alpha + \boldsymbol{\sigma}_\alpha, \quad (11.65)$$

we obtain the following power series expansion in  $\sigma$  :

$$\mathcal{L}(\varphi_\alpha, \dot{\varphi}_\alpha - \sigma_\alpha) = \mathcal{L}(\varphi_\alpha, \dot{\varphi}_\alpha) - \sum_\alpha \sigma_\alpha \pi_\alpha - \frac{1}{2} \sum_\alpha \sigma_\alpha^2(x) . \quad (11.66)$$

Inserting this into (11.64),

$$\mathcal{H}(x) = \sum_\alpha \pi_\alpha^2(x) - \mathcal{L}(\varphi_\alpha(x), \dot{\varphi}_\alpha(x)) + \frac{1}{2} \sum_\alpha \sigma_\alpha^2(x) . \quad (11.67)$$

Therefore,

$$\mathcal{H}_{\text{int}}(x) = -\mathcal{L}_{\text{int}}(x) + \frac{1}{2} \sum_\alpha n_\mu n_\nu \frac{\partial \mathcal{L}_{\text{int}}(x)}{\partial \varphi_{\alpha,\mu}(x)} \frac{\partial \mathcal{L}_{\text{int}}(x)}{\partial \varphi_{\alpha,\nu}(x)} . \quad (11.68)$$

It turns out that the second term on the right-hand side expresses a shift from (11.51). We now prove the equality (11.52) in this case, i.e., Matthews' theorem. To do so, we first specify the relation between the T-product and the normal product for the simple neutral scalar field theory.

We expand the T-product  $T[AB \dots Z]$  into normal products in the neutral scalar theory. In this case, we write the contraction function  $\Delta_F$ , replaced by  $\lambda \Delta_F$ , as  $T_\lambda[AB \dots Z]$ . Since differentiating this with respect to  $\lambda$  is equivalent to contracting  $\varphi(x)$  and  $\varphi(y)$  and multiplying by  $\Delta_F$ , we have

$$\frac{\partial}{\partial \lambda} T_\lambda[AB \dots Z] = \frac{1}{2} \int d^4x \int d^4y \frac{\delta}{\delta \varphi(x)} \Delta_F(x-y) \frac{\delta}{\delta \varphi(y)} T_\lambda[AB \dots Z] , \quad (11.69)$$

where we have treated  $\varphi$ 's in the normal product as c-numbers and differentiated with respect to them. Moreover, if we take  $\lambda = 0$ , then since this is exactly the same as that with no contraction, we have

$$T_0[AB \dots Z] = :AB \dots Z: . \quad (11.70)$$

Solving the differential equation (11.69) under the initial condition (11.70) and taking  $\lambda = 1$ , we find

$$T[AB \dots Z] = \exp \left[ \frac{1}{2} \int d^4x \int d^4y \frac{\delta}{\delta \varphi(x)} \Delta_F(x-y) \frac{\delta}{\delta \varphi(y)} \right] :AB \dots Z: . \quad (11.71)$$

Next we take the vacuum expectation value of this equation. Since the expectation value of the normal product vanishes, the right-hand side survives only if all field operators are contracted. Thus, setting to zero the operators  $\varphi$  which are not

contracted on the right-hand side, this gives only those contributions which appear when all operators are contracted:

$$\langle 0|T[AB\dots Z]|0\rangle = \exp\left[\frac{1}{2}\int d^4x\int d^4y\frac{\delta}{\delta\varphi(x)}\Delta_F(x-y)\frac{\delta}{\delta\varphi(y)}\right]:AB\dots Z:\Big|_{\varphi=0}. \quad (11.72)$$

In fact, (11.52) is only true if no derivative of  $\varphi$  is included in  $AB\dots Z$ . This operator appearing in the argument of the exponential function is related to the field quantization. It tells us how to contract operators. Denoting this by  $D$ , the operation which converts the normal product including  $\varphi$  and its derivative to the T-product is

$$T[AB\dots Z] = e^D :AB\dots Z: . \quad (11.73)$$

When the normal product includes  $\varphi$ , its derivative, and fermionic fields  $\psi$  and  $\bar{\psi}$ , we have

$$\begin{aligned} D = & \frac{1}{2}\int d^4x\int d^4y\left\{\frac{\delta}{\delta\varphi(x)}\Delta_F(x-y)\frac{\delta}{\delta\varphi(y)}\right. \\ & + 2\frac{\delta}{\delta\varphi,\mu(x)}\left[\frac{\partial}{\partial x_\mu}\Delta_F(x-y)\right]\frac{\delta}{\delta\varphi(y)} \\ & \left. + \frac{\delta}{\delta\varphi,\mu(x)}\left[\frac{\partial^2}{\partial x_\mu\partial y_\nu}\Delta_F(x-y) - in_\mu n_\nu\delta^4(x-y)\right]\frac{\delta}{\delta\varphi,\nu(y)}\right\} \\ & + \int d^4x\int d^4y\frac{\delta}{\delta\bar{\psi}_\beta(y)}S_F(x-y)_{\alpha\beta}\frac{\delta}{\delta\psi_\alpha(x)}, \end{aligned} \quad (11.74)$$

where we have used (11.49) as a contraction function. Regarding the functional derivatives with respect to the Dirac field, the reader is referred to the caution after (8.59). Likewise for the T\*-product,

$$T^*[AB\dots Z] = e^{D^*} :AB\dots Z: , \quad (11.75)$$

where  $D^*$  is defined so that (11.50) applies to contractions of the  $\varphi$  derivatives in (11.74), whence

$$D = D^* + \frac{1}{2}\int d^4x\int d^4y\frac{\delta}{\delta\varphi,\mu(x)}[-in_\mu n_\nu\delta^4(x-y)]\frac{\delta}{\delta\varphi,\nu(y)}. \quad (11.76)$$

In both cases, we have treated  $\varphi$  and its derivative as independent when taking functional derivatives.

We now choose the interaction Lagrangian density to have the form

$$\mathcal{L}_{\text{int}}(x) = -j_\mu(x)\partial_\mu\varphi(x) + \mathcal{L}_{\text{int}}^{(0)}(x), \quad (11.77)$$



where  $\mathcal{L}_{\text{int}}^{(0)}$  is a term which includes no derivatives. Therefore, from (11.68),

$$\mathcal{H}_{\text{int}}(x) = -\mathcal{L}_{\text{int}}(x) + \frac{1}{2}[n_{\mu}j_{\mu}(x)]^2. \quad (11.78)$$

In order to compute the  $S$ -matrix, we consider

$$\begin{aligned} T \exp \left[ -i \int d^4x \mathcal{H}_{\text{int}}(x) \right] &= e^D : \exp \left[ -i \int d^4x \mathcal{H}_{\text{int}}(x) \right] : \\ &= e^{D^*} e^{D-D^*} : \exp \left[ -i \int d^4x \mathcal{H}_{\text{int}}(x) \right] : \end{aligned} \quad (11.79)$$

where  $D - D^*$  is the second term on the right-hand side of (11.76). This is a functional derivative with respect to the derivative of  $\varphi$ . Hence, this acts only on the term  $j_{\mu} \partial_{\mu} \varphi$  in  $\mathcal{H}_{\text{int}}(x)$  which includes the derivative of  $\varphi$ . We must therefore compute

$$e^{D-D^*} : \exp \left[ -i \int d^4x j_{\mu}(x) \partial_{\mu} \varphi(x) \right] : . \quad (11.80)$$

To carry out this computation, we use

$$\exp \left( \lambda \frac{d^2}{dx^2} \right) e^{ax} = e^{ax+a^2\lambda} . \quad (11.81)$$

Generalizing this formula and using it to calculate (11.80), we find that (11.80) is equivalent to something of the form

$$: \exp \left[ -i \int d^4x \left( j_{\mu}(x) \partial_{\mu} \varphi(x) - \frac{1}{2} [n_{\mu}j_{\nu}(x)]^2 \right) \right] : . \quad (11.82)$$

Therefore, (11.79) can be written in the form

$$\begin{aligned} T \exp \left[ -i \int d^4x \mathcal{H}_{\text{int}}(x) \right] &= e^{D^*} : \exp \left[ -i \int d^4x \left( \mathcal{H}_{\text{int}}(x) - \frac{1}{2} [n_{\mu}j_{\nu}(x)]^2 \right) \right] : \\ &= e^{D^*} : \exp \left[ i \int d^4x \mathcal{L}_{\text{int}}(x) \right] : \\ &= T^* \exp \left[ i \int d^4x \mathcal{L}_{\text{int}}(x) \right] . \end{aligned} \quad (11.83)$$

Dividing both sides of this by the vacuum expectation value yields the  $S$ -matrix:

$$S = \frac{T^* \exp \left[ i \int d^4x \mathcal{L}_{\text{int}}(x) \right]}{\langle 0 | T^* \exp \left[ i \int d^4x \mathcal{L}_{\text{int}}(x) \right] | 0 \rangle} . \quad (11.84)$$

This is known as *Matthews' theorem*. Note that this holds only if derivatives of the field operator are included linearly in (11.77).

## 11.5 Example of Matthews' Theorem with Modification

In the last section, we considered a situation where Matthews' theorem holds true and  $T$  and  $\mathcal{H}_{\text{int}}$  can be replaced by  $T^*$  and  $-\mathcal{L}_{\text{int}}$ , respectively. In general, we have to replace  $\mathcal{H}_{\text{int}}$  by something slightly different from  $-\mathcal{L}_{\text{int}}$ . We thus generalize Matthews' theorem to

$$T \exp \left[ -i \int d^4x \mathcal{H}_{\text{int}}(x) \right] = T^* \exp \left[ i \int d^4x \mathcal{L}_{\text{eff}}(x) \right] , \quad (11.85)$$

where  $\mathcal{L}_{\text{eff}}$  is an effective interaction Lagrangian density, which is equivalent to  $\mathcal{L}_{\text{int}}$  only when Matthews' theorem holds true. As an example of a situation where the theorem does not hold true, we consider the Lagrangian density

$$\mathcal{L} = -\frac{1}{2} D_{ab}(\varphi) \partial_\lambda \varphi_a \partial_\lambda \varphi_b - V(\varphi) , \quad (11.86)$$

where  $\varphi_a$  ( $a = 1, 2, \dots, N$ ) are real scalar fields and  $D_{ab}$  is a real positive-definite matrix which is a function of  $\varphi$ . In addition, we assume that  $D$  and  $V$  do not include derivatives of  $\varphi$ . We shall now derive the effective interaction Lagrangian density for this theory.

When  $x_0 = y_0$ , the canonical commutation relations read

$$[\varphi_a(x), \varphi_b(y)] = 0 , \quad (11.87)$$

$$[\varphi_a(x), \dot{\varphi}_b(y)] = i C_{ab}(\varphi) \delta^3(x - y) , \quad (11.88)$$

where  $C$  is the inverse matrix of  $D$ , i.e.,

$$\sum_b C_{ab} D_{bc} = \sum_b D_{ab} C_{bc} = \delta_{ac} . \quad (11.89)$$

With summing over repeated indices, the Euler–Lagrange equation gives

$$[\mathcal{L}]_{\varphi_a} = D_{ab}(\square\varphi_b - \mathbf{j}_b) = 0, \quad (11.90)$$

$$\mathbf{j}_a = C_{ab} \left( \frac{1}{2} \frac{\partial}{\partial\varphi_b} D_{cd} - \frac{\partial}{\partial\varphi_d} D_{bc} \right) \partial_\lambda\varphi_c \partial_\lambda\varphi_d + C_{ab} \frac{\partial}{\partial\varphi_b} V. \quad (11.91)$$

We now introduce a generating functional for the Green's functions:

$$\mathcal{T}[J] = \left\langle \mathbf{0} \left| T \exp \left[ -i \int d^4x J_a(x) \varphi_a(x) \right] \right| \mathbf{0} \right\rangle. \quad (11.92)$$

Therefore, combining (11.90) and the equal-time commutation relation (11.88),

$$\begin{aligned} \square_x \langle \mathbf{0} | T[\varphi_a(x), \mathbf{A}, \mathbf{B}, \dots] | \mathbf{0} \rangle &= \langle \mathbf{0} | T[\mathbf{j}_a(x), \mathbf{A}, \mathbf{B}, \dots] | \mathbf{0} \rangle \\ &+ i \langle \mathbf{0} | T \left[ C_{ac}(\varphi(x)) \frac{\delta}{\delta\varphi_b(x)}, \mathbf{A}, \mathbf{B}, \dots \right] | \mathbf{0} \rangle + \dots \end{aligned} \quad (11.93)$$

Expressing this in terms of the generating functional,

$$\begin{aligned} \square_x i \frac{\delta}{\delta J_a(x)} \mathcal{T}[J] &= \square_x \left\langle \mathbf{0} \left| T \left[ \varphi_a(x), \exp \left( -i \int d^4y J_c(y) \varphi_c(y) \right) \right] \right| \mathbf{0} \right\rangle \\ &= \left\langle \mathbf{0} \left| T \left[ \mathbf{j}_a(x), \exp \left( -i \int d^4y J_c(y) \varphi_c(y) \right) \right] \right| \mathbf{0} \right\rangle \\ &+ J_b(x) \left\langle \mathbf{0} \left| T \left[ C_{ab}(\varphi(x)), \exp \left( -i \int d^4y J_c(y) \varphi_c(y) \right) \right] \right| \mathbf{0} \right\rangle. \end{aligned} \quad (11.94)$$

Using  $\mathcal{T}[J]$ , the second term on the right-hand side can be expressed as

$$J_b(x) C_{ab} \left( i \frac{\delta}{\delta J(x)} \right) \mathcal{T}[J]. \quad (11.95)$$

In the interaction picture, considering the T-product which includes only two derivatives of  $\varphi$ ,

$$\begin{aligned} T[\partial_\mu\varphi(x), \partial_\nu\varphi(y), \varphi(z), \dots] &= T^*[\partial_\mu\varphi(x), \partial_\nu\varphi(y), \varphi(z), \dots] \\ &- i n_\mu n_\nu \delta^4(x-y) T[\varphi(z), \dots]. \end{aligned} \quad (11.96)$$

Noting that, in this derivation, we have only used the equal-time commutation relations and we have not used the field equation, we can easily make the extension

to the Heisenberg picture:

$$T[\partial_\mu \boldsymbol{\varphi}_a(x), \partial_\nu \boldsymbol{\varphi}_b(y), \boldsymbol{\varphi}_c(z), \dots] = T^*[\partial_\mu \boldsymbol{\varphi}_a(x), \partial_\nu \boldsymbol{\varphi}_b(y), \boldsymbol{\varphi}_c(z), \dots] \quad (11.97)$$

$$-in_\mu n_\nu \delta^4(x-y) T[C_{ab}(\boldsymbol{\varphi}(x)), \boldsymbol{\varphi}_c(z), \dots].$$

In particular, if we take  $y \rightarrow x$  and  $\mu = \nu = \lambda$ , then since  $n_\lambda n_\lambda = -1$ ,

$$T[\partial_\lambda \boldsymbol{\varphi}_a(x), \partial_\lambda \boldsymbol{\varphi}_b(y), \boldsymbol{\varphi}_c(z), \dots] = T^*[\partial_\lambda \boldsymbol{\varphi}_a(x), \partial_\lambda \boldsymbol{\varphi}_b(y), \boldsymbol{\varphi}_c(z), \dots] \quad (11.98)$$

$$+i\delta^4(0) T[C_{ab}(\boldsymbol{\varphi}(x)), \boldsymbol{\varphi}_c(z), \dots].$$

Using this result, the first term on the right-hand side of (11.94) can be written

$$\langle \mathbf{0} | T^* \left[ j_a(x), \exp \left( -i \int d^4 y J_c(y) \boldsymbol{\varphi}_c(y) \right) \right] | \mathbf{0} \rangle \quad (11.99)$$

$$+i\delta^4(0) \langle \mathbf{0} | T \left[ C_{ab}(x) \left( \frac{1}{2} \frac{\partial D_{cd}(x)}{\partial \boldsymbol{\varphi}_b(x)} - \frac{\partial D_{bc}(x)}{\partial \boldsymbol{\varphi}_d(x)} \right) C_{cd}(x) \right. \right.$$

$$\left. \left. \times \exp \left( -i \int d^4 y J_c(y) \boldsymbol{\varphi}_c(y) \right) \right] | \mathbf{0} \rangle,$$

where  $C_{ab}(x)$  is an abbreviation for  $C_{ab}(\boldsymbol{\varphi}(x))$ , and the same goes for  $D_{cd}(x)$ .

Thus,  $\mathcal{F}[J]$  satisfies

$$\square_x i \frac{\delta}{\delta J_a(x)} \mathcal{F}[J] = \left[ j_a \left( i \frac{\delta}{\delta J(x)} \right) + i\delta^4(0) F_a \left( i \frac{\delta}{\delta J(x)} \right) + J_b(x) C_{ab} \left( i \frac{\delta}{\delta J(x)} \right) \right] \mathcal{F}[J], \quad (11.100)$$

where

$$F_a(\boldsymbol{\varphi}(x)) = C_{ab} \left( \frac{1}{2} \frac{\partial D_{cd}(x)}{\partial \boldsymbol{\varphi}_b(x)} - \frac{\partial D_{bc}(x)}{\partial \boldsymbol{\varphi}_d(x)} \right) C_{cd}(x). \quad (11.101)$$

Note that we can factorize the functional derivative  $j_a$  using the  $T^*$ -product. Rewriting the third term on the right-hand side of (11.100), we have

$$J_b(x) C_{ab} \left( i \frac{\delta}{\delta J(x)} \right) = C_{ab} \left( i \frac{\delta}{\delta J(x)} \right) J_b(x) - i\delta^4(0) \left. \frac{\partial C_{ab}(x)}{\partial \boldsymbol{\varphi}_b(x)} \right|_{\boldsymbol{\varphi} \rightarrow i\delta/\delta J}. \quad (11.102)$$

Then the coefficients of  $i\delta^4(0)$  in (11.100) are

$$\begin{aligned}
 F_a - \frac{\partial C_{ab}}{\partial \varphi_b} &= \frac{1}{2} C_{ab} \frac{\partial D_{cd}}{\partial \varphi_b} C_{cd} - C_{ab} \left( \frac{\partial D_{bc}}{\partial \varphi_d} C_{cd} + D_{bc} \frac{\partial C_{cd}}{\partial \varphi_d} \right) \\
 &= \frac{1}{2} C_{ab} \frac{\partial}{\partial \varphi_b} \ln(\det D) - C_{ab} \frac{\partial}{\partial \varphi_d} \delta_{bd} \\
 &= \frac{1}{2} C_{ab} \frac{\partial}{\partial \varphi_b} \ln(\det D) \\
 &= \frac{1}{2} G_a .
 \end{aligned} \tag{11.103}$$

Consequently, (11.100) can be written as

$$\square_x i \frac{\delta}{\delta J_a(x)} \mathcal{F}[J] = \left[ j_a \left( i \frac{\delta}{\delta J(x)} \right) + \frac{1}{2} i \delta^4(0) G_a \left( i \frac{\delta}{\delta J(x)} \right) + C_{ab} \left( i \frac{\delta}{\delta J(x)} \right) J_b(x) \right] \mathcal{F}[J] . \tag{11.104}$$

Multiplying this on the left by  $D$ , we obtain

$$\begin{aligned}
 D_{ab} \left( i \frac{\delta}{\delta J(x)} \right) \square_x i \frac{\delta}{\delta J_b(x)} \mathcal{F}[J] & \tag{11.105} \\
 = \left[ D_{ab} j_b \left( i \frac{\delta}{\delta J(x)} \right) + J_a(x) + \frac{1}{2} i \delta^4(0) \frac{\partial}{\partial \varphi_a(x)} \ln(\det D(x)) \Big|_{\varphi \rightarrow i\delta/\delta J} \right] \mathcal{F}[J] .
 \end{aligned}$$

This is the equation satisfied by the generating functional  $\mathcal{F}[J]$ .

In order to derive a Feynman–Dyson-like formula in a theory like this, we must express the solution of the equation for  $\mathcal{F}[J]$  in terms of  $\varphi$  in the interaction picture. We thus test the following quantity:

$$\text{“}\mathcal{F}[J]\text{”} = \left\langle 0 \left| T^* \exp \left[ i \int d^4x [\mathcal{L}_{\text{int}}(x) - J_a(x) \varphi_a(x)] \right] \right| 0 \right\rangle , \tag{11.106}$$

where

$$\mathcal{L}_{\text{int}} = \mathcal{L} + \frac{1}{2} (\partial_\lambda \varphi_a)^2 . \tag{11.107}$$

To obtain the functional equation satisfied by this “ $\mathcal{F}[J]$ ”, we begin with

$$\begin{aligned}
 \text{“}\mathcal{F}[J]\text{”} &= \exp \left[ i \int d^4x \mathcal{L}_{\text{int}} \left( i \frac{\delta}{\delta J(x)} \right) \right] \left\langle 0 \left| \exp \left[ -i \int d^4x J_b(x) \varphi_b(x) \right] \right| 0 \right\rangle \\
 &= \exp \left[ i \int d^4x \mathcal{L}_{\text{int}} \left( i \frac{\delta}{\delta J(x)} \right) \right] \exp \left[ -\frac{1}{2} \int d^4x \int d^4y J_b(x) \Delta_{\text{F}}(x-y) J_b(y) \right] ,
 \end{aligned} \tag{11.108}$$

where we replace  $\varphi$  by  $i\delta/\delta J$  in  $\mathcal{L}_{\text{int}}$ . For the transformation on the right-hand side, we have used (11.72). Hence,

$$\begin{aligned}
 \square_x i \frac{\delta}{\delta J_a(x)} \mathcal{T}[J] &= \exp \left[ i \int d^4x \mathcal{L}_{\text{int}} \left( i \frac{\delta}{\delta J(x)} \right) \right] J_a(x) \\
 &\quad \times \exp \left[ -\frac{1}{2} \int d^4x \int d^4y J_b(x) \Delta_F(x-y) J_b(y) \right] \\
 &= J_a(x) \mathcal{T}[J] + \left[ i \int d^4x' \mathcal{L}_{\text{int}} \left( i \frac{\delta}{\delta J(x')} \right), J_a(x) \right] \mathcal{T}[J] \\
 &= J_a(x) \mathcal{T}[J] - [\mathcal{L}_{\text{int}}]_{\varphi_a} \left( i \frac{\delta}{\delta J(x)} \right) \mathcal{T}[J] \\
 &= \left\{ J_a(x) - D_{ab} \left( i \frac{\delta}{\delta J(x)} \right) \left[ \square_x i \frac{\delta}{\delta J_b(x)} - j_b \left( i \frac{\delta}{\delta J(x)} \right) \right] \right. \\
 &\quad \left. + \square_x i \frac{\delta}{\delta J_a(x)} \right\} \mathcal{T}[J]. \tag{11.109}
 \end{aligned}$$

Since the last term on the right-hand side is the same as the left-hand side, these terms cancel out:

$$D_{ab} \left( i \frac{\delta}{\delta J(x)} \right) \square_x i \frac{\delta}{\delta J_b(x)} \mathcal{T}[J] = D_{ab} j_b \left( i \frac{\delta}{\delta J(x)} \right) \mathcal{T}[J] + J_a(x) \mathcal{T}[J]. \tag{11.110}$$

If we replace  $V$  in this equation by

$$V' = V + \frac{i}{2} \delta^4(0) \ln(\det D), \tag{11.111}$$

it coincides with (11.105). Hence,  $\mathcal{T}[J]$  can be obtained by replacing  $V$  by  $V'$  in  $\mathcal{T}[J]$  and normalizing in such a way that it is equal to unity when  $J = 0$ . Thus,

$$\mathcal{T}[J] = \frac{\langle 0 | T^* \exp \left\{ i \int d^4x [\mathcal{L}_{\text{eff}}(\varphi(x)) - J_a(x) \varphi_a(x)] \right\} | 0 \rangle}{\langle 0 | T^* \exp \left[ i \int d^4x \mathcal{L}_{\text{eff}}(\varphi(x)) \right] | 0 \rangle}, \tag{11.112}$$

where

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{int}} - \frac{i}{2} \delta^4(0) \ln(\det D). \tag{11.113}$$

It thus turns out that, in this example, Matthews' theorem has been modified.

The next question concerns properties of the additional term. First of all, this term is an imaginary number and includes the divergence  $\delta^4(0)$ . We have the following

representation of the  $\delta$ -function in the space-time coordinates:

$$\delta^4(x) = \frac{1}{(2\pi)^4} \int d^4k e^{ik \cdot x} \rightarrow \delta^4(0) = \frac{1}{(2\pi)^4} \int d^4k . \quad (11.114)$$

This therefore gives a fourth-order divergence in momentum space. The imaginary coefficient is connected with the fourth-order divergence. If we compute the closed loop without the additional term, a fourth-order divergence comes about in this theory. Introducing the cutoff  $\Lambda$  in momentum space, we obtain

$$\int d^4k \frac{k^2}{k^2 + m^2 - i\epsilon} \sim \frac{\pi^2}{2} \Lambda^4 - i\pi^2 m^2 \Lambda^2 . \quad (11.115)$$

This implies that the divergences up to second order will give divergent contributions to the mass and coupling constants, as will be discussed later, but those are real numbers. On the other hand, the fourth-order divergence implied by the equation above, compared to the divergences up to second order, is an imaginary number. Such a contribution cannot be removed by renormalization and breaks the unitarity. Fortunately, the additional term mentioned above automatically cancels this fourth-order divergence, and in this sense a safety mechanism is automatically introduced into the theory.

Although the result mentioned above can also be derived by the path-integral method, this gives the additional term a different interpretation. The path-integral method produces the above result more easily than the method used here. This implies that, since the path-integral method is based on the Lagrangian, we can say that it is more suitable to derive the result including the Lagrangian and the  $T^*$ -product.

## 11.6 Reduction Formula in the Interaction Picture

So far we have discussed the  $S$ -matrix computational method in the interaction picture. Combining the reduction formula given in Sect. 8.3 and the Gell-Mann–Low relation, we can also express the  $S$ -matrix elements in terms of the Green's function in the Heisenberg picture.

In Dyson's formula, the interaction Hamiltonian density or Lagrangian density appear when we express the  $S$ -matrix. We can ask ourselves whether it is possible to derive an equation which does not depend explicitly on the form of the interaction.

For simplicity, we consider the charged scalar field and analyze the  $S$ -matrix elements for the scattering process

$$a + b \rightarrow b + a . \quad (11.116)$$

To do so, we expand  $U(\infty, -\infty)$  in normal products using Wick's theorem and determine the coefficient of the term

$$:\varphi_a^\dagger \varphi_b^\dagger \varphi_a \varphi_b: . \quad (11.117)$$

Since this requires us to read off the normal product from  $U(\infty, -\infty)$  and contract the rest, we need to calculate

$$\left\langle 0 \left| \frac{\delta}{\delta \varphi_a^\dagger(x'_1)} \frac{\delta}{\delta \varphi_b^\dagger(x'_2)} \frac{\delta}{\delta \varphi_a(x_1)} \frac{\delta}{\delta \varphi_b(x_2)} U(\infty, -\infty) \right| 0 \right\rangle . \quad (11.118)$$

Writing the final state as  $|a', b'\rangle$ , the  $S$ -matrix element is given by

$$\begin{aligned} \langle a', b' | S - 1 | a, b \rangle &= \int d^4 x'_1 d^4 x'_2 d^4 x_1 d^4 x_2 \langle a' | \varphi_a^\dagger(x'_1) | 0 \rangle \langle b' | \varphi_b^\dagger(x'_2) | 0 \rangle \\ &\quad \times \left\langle 0 \left| \frac{\delta}{\delta \varphi_a^\dagger(x'_1)} \frac{\delta}{\delta \varphi_b^\dagger(x'_2)} \frac{\delta}{\delta \varphi_a(x_1)} \frac{\delta}{\delta \varphi_b(x_2)} U(\infty, -\infty) \right| 0 \right\rangle \\ &\quad \times \frac{\langle 0 | \varphi_a(x_1) | a \rangle \langle 0 | \varphi_b(x_2) | b \rangle}{\langle 0 | U(\infty, -\infty) | 0 \rangle} . \end{aligned} \quad (11.119)$$

From the reduction formula, the functional derivative is given by

$$(\square_x^2 - m_a^2) T[\varphi_a(x) \dots] = i \frac{\delta}{\delta \varphi_a^\dagger(x)} T[\dots] . \quad (11.120)$$

Denoting the Klein–Gordon operator on the left-hand side by  $K_x^a$  and using (11.120), equation (11.119) becomes

$$\begin{aligned} \langle a', b' | S - 1 | a, b \rangle &= \langle 0 | U(\infty, -\infty) | 0 \rangle^{-1} \\ &\quad \times \int d^4 x'_1 d^4 x'_2 d^4 x_1 d^4 x_2 \langle a' | \varphi_a^\dagger(x'_1) | 0 \rangle \langle b' | \varphi_b^\dagger(x'_2) | 0 \rangle \\ &\quad \times (-i)^4 K_{x'_1}^a K_{x'_2}^b K_{x_1}^a K_{x_2}^b \langle 0 | T[\varphi_a(x'_1) \varphi_b(x'_2) \varphi_a^\dagger(x_1) \varphi_b^\dagger(x_2) U(\infty, -\infty)] | 0 \rangle \\ &\quad \times \langle 0 | \varphi_a(x_1) | a \rangle \langle 0 | \varphi_b(x_2) | b \rangle . \end{aligned} \quad (11.121)$$

Here  $\langle 0 | \varphi_a(x) | a \rangle$ , which we should call a one-body wave function, has the same structure in both the interaction picture and the Heisenberg picture. The only difference would be a proportionality coefficient. Although it is not trivial to separate the whole Lagrangian density into the free part and the interaction part, the expressions are equal if we use the renormalized interaction picture discussed in



the next chapter:

$$\langle 0|\varphi_a(x)|a\rangle = \langle \mathbf{0}|\varphi_a(x)|\mathbf{a}\rangle = \frac{1}{\sqrt{2p_0V}}e^{ip\cdot x}. \quad (11.122)$$

If we use the Gell-Mann–Low relation in this case, (11.121) can be expressed solely in terms of quantities in the Heisenberg picture:

$$\begin{aligned} \langle a', b'|S - 1|a, b\rangle &= \int d^4x'_1 d^4x'_2 d^4x_1 d^4x_2 \langle a'|\varphi_a^\dagger(x'_1)|\mathbf{0}\rangle \langle b'|\varphi_b^\dagger(x'_2)|\mathbf{0}\rangle \\ &\quad \times K_{x'_1}^a K_{x'_2}^b K_{x_1}^a K_{x_2}^b \langle \mathbf{0}|T[\varphi_a(x'_1)\varphi_b(x'_2)\varphi_a^\dagger(x_1)\varphi_b^\dagger(x_2)]|\mathbf{0}\rangle \\ &\quad \times \langle \mathbf{0}|\varphi_a(x_1)|\mathbf{a}\rangle \langle \mathbf{0}|\varphi_b(x_2)|\mathbf{b}\rangle. \end{aligned} \quad (11.123)$$

Unlike Dyson's formula, in the above expression of the  $S$ -matrix element, the explicit form of the interaction does not appear. The problem in the interaction picture of separating the Lagrangian into the free part and the interaction part does not arise. However, in the process of deriving this formula, we have made the assumption (11.7), which is hard to justify. In fact, this result is justified only when we start with the renormalized interaction picture discussed above. Thus, we have to discuss the asymptotic conditions which lead to the above formula in the framework of the Heisenberg picture.

## 11.7 Asymptotic Conditions

The derivation of the  $S$ -matrix element in the Heisenberg picture in the previous section has been based on several assumptions. The question is whether or not we can derive the same result from clearer assumptions. In fact, this was done by Lehmann, Symanzik, and Zimmermann. The assumptions they made are called *asymptotic conditions* [113].

In Sect. 6.3, we introduced two kinds of asymptotic field in connection with the Yang–Feldman formalism. The asymptotic fields  $\varphi^{\text{int}}$  and  $\varphi^{\text{out}}$  for the real scalar field  $\varphi$  satisfy

$$(\square - m^2)\varphi^{\text{in}}(x) = 0 = (\square - m^2)\varphi^{\text{out}}(x), \quad (11.124)$$

$$[\varphi^{\text{in}}(x), \varphi^{\text{in}}(y)] = i\Delta(x - y) = [\varphi^{\text{out}}(x), \varphi^{\text{out}}(y)]. \quad (11.125)$$

Since these two types of scalar field are engendered by the same scalar field  $\varphi$ , we know that  $\varphi^{\text{int}}$  and  $\varphi^{\text{out}}$  are not independent of one another. This implies that they will not commute. Intuitively speaking, as in the case of the Yang–Feldman

formalism,

$$\varphi(x) \longrightarrow \begin{cases} \varphi^{\text{int}}(x), & t \rightarrow -\infty, \\ \varphi^{\text{out}}(x), & t \rightarrow \infty. \end{cases} \quad (11.126)$$

As just described, the reason why the field asymptotes to the free field when  $t \rightarrow \pm\infty$  is that since the particles are then far away from each other and there is no effect from other particles, so they behave like free particles. This fact is closely related to the issue of renormalization discussed in the next chapter, and to make this idea more rigorous we have to express the wave function of a particle, not by a plane wave, but by a wave packet.

We assume that some function  $f(x)$  satisfies the conditions

$$(\square - m^2)f(x) = 0, \quad (11.127)$$

$$-i \int d^3x \left( f \frac{\partial f^*}{\partial x_0} - f^* \frac{\partial f}{\partial x_0} \right) = 1. \quad (11.128)$$

Then corresponding to this  $f$ , we introduce the operators

$$\varphi_f(t) = -i \int d^3x \left[ \varphi(x) \frac{\partial f^*(x)}{\partial x_0} - f^*(x) \frac{\partial \varphi(x)}{\partial x_0} \right], \quad (11.129)$$

$$\varphi_f^\dagger(t) = i \int d^3x \left[ \varphi(x) \frac{\partial f(x)}{\partial x_0} - f(x) \frac{\partial \varphi(x)}{\partial x_0} \right], \quad (11.130)$$

where  $t = x_0$ . We then define the corresponding asymptotic fields  $\varphi_f^{\text{in}}$  and  $\varphi_f^{\text{out}}$  by

$$\lim_{\tau \rightarrow -\infty} (\Phi, \varphi_f(\tau)\Psi) = (\Phi, \varphi_f^{\text{in}}\Psi), \quad (11.131)$$

$$\lim_{\tau \rightarrow \infty} (\Phi, \varphi_f(\tau)\Psi) = (\Phi, \varphi_f^{\text{out}}\Psi). \quad (11.132)$$

For both states  $\Phi$  and  $\Psi$ , we can define the same asymptotic field if we start with the normalized state vector  $\varphi_f^\dagger(\tau)$ . Note that the right-hand sides of (11.131) and (11.132) no longer depend on the time variable. Such a limit of an operator in the sense of the matrix element is called *weak convergence*, in contrast to strong convergence defined in the sense of the norm.

Next we consider an orthogonal system of wave functions. A wave function here is the matrix element with the vacuum of a field operator in a one-particle state. We consider the set of functions  $\{f_\alpha(x)\}$  satisfying (11.127), (11.128), and the condition

$$-i \int d^3x \left( f_\alpha \frac{\partial f_\beta^*}{\partial x_0} - f_\beta^* \frac{\partial f_\alpha}{\partial x_0} \right) = \delta_{\alpha\beta}. \quad (11.133)$$

The completeness condition for this system of orthogonal functions is

$$\sum_{\alpha} f_{\alpha}(x) f_{\alpha}^{*}(y) = i\Delta^{(+)}(x - y). \quad (11.134)$$

We now introduce the complete system of state vectors  $\{\Phi^{\text{in}}\}$ . Assuming that  $\Phi_0$  is the vacuum,

$$\begin{aligned} & \Phi_0, \\ \Phi_{\alpha}^{\text{in}} &= \varphi_{\alpha}^{\dagger\text{in}} \Phi_0, \\ & \vdots \\ \Phi_{\alpha_1 \dots \alpha_k}^{\text{in}} &= (p_{\alpha_1 \dots \alpha_k})^{-1/2} \varphi_{\alpha_1}^{\dagger\text{in}} \dots \varphi_{\alpha_k}^{\dagger\text{in}} \Phi_0, \end{aligned} \quad (11.135)$$

where  $p_{\alpha_1 \dots \alpha_k} = n_1! n_2! \dots n_r!$  and  $n$  stands for the number of particles in the same one-particle state in  $(\alpha_1, \dots, \alpha_k)$ . Replacing  $\varphi^{\dagger\text{in}}$  by  $\varphi^{\dagger\text{out}}$ , we can also construct the complete system  $\{\Phi^{\text{out}}\}$ . The  $S$ -matrix can be defined as the unitary transformation between these two pairs of complete orthonormal systems. It will be shown later that this definition reproduces the  $S$ -matrix elements given in the previous section. In the next chapter, it will be shown that it also coincides with the definition of the  $S$ -matrix in the Lippmann–Schwinger theory, viz.,

$$S_{\beta\alpha} = (\Phi_{\beta}^{\text{out}}, \Phi_{\alpha}^{\text{in}}). \quad (11.136)$$

An equivalent definition is

$$\Phi_{\alpha}^{\text{in}} = S \Phi_{\alpha}^{\text{out}}. \quad (11.137)$$

It is clear from the definition that, for the two asymptotic fields,

$$(\Phi_{\beta}^{\text{in}}, \varphi_f^{\text{in}} \Phi_{\alpha}^{\text{in}}) = (\Phi_{\beta}^{\text{out}}, \varphi_f^{\text{out}} \Phi_{\alpha}^{\text{out}}). \quad (11.138)$$

Combining (11.137) and (11.138),

$$\varphi_f^{\text{out}} = S^{-1} \varphi_f^{\text{in}} S. \quad (11.139)$$

Similarly,

$$\varphi_f^{\dagger\text{out}} = S^{-1} \varphi_f^{\dagger\text{in}} S. \quad (11.140)$$

We introduce  $\varphi^{\text{in}}(x)$  by

$$\varphi^{\text{in}}(x) = \sum_{\alpha} [f_{\alpha}^{*}(x) \varphi_{\alpha}^{\dagger\text{in}} + f_{\alpha}(x) \varphi_{\alpha}^{\text{in}}], \quad (11.141)$$

and define  $\varphi^{\text{out}}$  by the same equation, viz.,

$$\varphi^{\text{out}}(x) = S^{-1}\varphi^{\text{in}}(x)S. \quad (11.142)$$

Since the latter coincides with (6.32), we see that this is the same as the  $S$ -matrix given previously. In relation to (11.122) in the previous section, we mentioned that we have not distinguished whether we take the in-state or the out-state for the one-particle state in the Heisenberg picture. This implies that, *for the stable one-particle state*  $\alpha$ , we must have

$$\Phi_\alpha^{\text{in}} = \Phi_\alpha^{\text{out}}. \quad (11.143)$$

Combining this with (11.122) in the previous section,

$$(\square - m^2)(\Phi_0, \varphi_0 \Phi_\alpha^{\text{in}}) = (\square - m^2)(\Phi_0, \varphi_0 \Phi_\alpha^{\text{out}}), \quad (11.144)$$

$$(\Phi_0, \varphi_0 \Phi_\alpha^{\text{in}}) = (\Phi_0, \varphi_0 \Phi_\alpha^{\text{out}}) = f_\alpha(x). \quad (11.145)$$

We will discuss this requirement in the context of renormalization in the next chapter.

Equations (11.131) and (11.132), together with the assumption of the existence of the asymptotic field, are called *asymptotic conditions*. Starting from these conditions, we will derive the LSZ reduction formula in the Heisenberg picture, or as they called it, the magic formula (Zauberformel) [113]. We introduce a more concise notation:

$$T(x_1, \dots, x_n) = T[\varphi(x_1) \dots \varphi(x_n)], \quad (11.146)$$

$$\tau(x_1, \dots, x_n) = (\Phi_0, T[\varphi(x_1) \dots \varphi(x_n)]\Phi_0), \quad (11.147)$$

$$K_y = \square_y - m^2, \quad f \overleftrightarrow{\partial}_x g = f \frac{\partial g}{\partial x} - \frac{\partial f}{\partial x} g. \quad (11.148)$$

To begin with, we prove the following equation:

$$(\Phi_0, T(x_1, \dots, x_n)\Phi_\alpha^{\text{in}}) = -i \int d^4y f_\alpha(y) K_y \tau(x_1, \dots, x_n, y). \quad (11.149)$$

The left-hand side is

$$\begin{aligned} \text{LHS} &= \lim_{y_0 \rightarrow -\infty} i \int d^3y (\Phi_0, T(x_1, \dots, x_n, y)\Phi_0) \overleftrightarrow{\partial}_{y_0} f_\alpha(y) \\ &= i \lim_{y_0 \rightarrow -\infty} \int d^3y \tau(x_1, \dots, x_n, y) \overleftrightarrow{\partial}_{y_0} f_\alpha(y). \end{aligned} \quad (11.150)$$

Noting that  $\varphi_\alpha^{\text{out}}$  is an annihilation operator, in the limit  $y_0 \rightarrow \infty$ , we have

$$\begin{aligned} i \lim_{y_0 \rightarrow \infty} \int d^3 y \tau(x_1, \dots, x_n, y) \frac{\overleftrightarrow{\partial}}{\partial y_0} f_\alpha(y) &= (\Phi_0, \varphi_\alpha^{\dagger \text{out}} T(x_1, \dots, x_n) \Phi_0) \\ &= (\varphi_\alpha^{\text{out}} \Phi_0, T(x_1, \dots, x_n) \Phi_0) = 0. \end{aligned} \quad (11.151)$$

Taking the difference between the two equations above,

$$\begin{aligned} &(\Phi_0, T(x_1, \dots, x_n) \Phi_0^{\text{in}}) \\ &= i (\lim_{y_0 \rightarrow -\infty} - \lim_{y_0 \rightarrow \infty}) \int d^3 y \tau(x_1, \dots, x_n, y) \frac{\overleftrightarrow{\partial}}{\partial y_0} f_\alpha(y) \\ &= -i \int d^4 y \frac{\partial}{\partial y_0} \left[ \tau(x_1, \dots, x_n, y) \frac{\overleftrightarrow{\partial}}{\partial y_0} f_\alpha(y) \right] \\ &= -i \int d^4 y \left[ \tau(x_1, \dots, x_n, y) \frac{\partial^2 f_\alpha(y)}{\partial y_0^2} - \frac{\partial^2 \tau(x_1, \dots, x_n, y)}{\partial y_0^2} f_\alpha(y) \right]. \end{aligned} \quad (11.152)$$

We combine this with Green's theorem:

$$\begin{aligned} &\int_V d^3 y \left[ \tau(x_1, \dots, x_n, y) \cdot \Delta_y f_\alpha(y) - \Delta_y \tau(x_1, \dots, x_n, y) \cdot f_\alpha(y) \right] \\ &= \int_{S=\partial V} dS \left[ \tau(x_1, \dots, x_n, y) \cdot \frac{\partial}{\partial y_n} f_\alpha(y) - \frac{\partial}{\partial y_n} \tau(x_1, \dots, x_n, y) \cdot f_\alpha(y) \right] \\ &\rightarrow 0, \quad \partial V \rightarrow \infty. \end{aligned} \quad (11.153)$$

Here, we use the fact that, since  $f_\alpha(y)$  is a wave packet and corresponds to a local wave, it vanishes at a long range. Therefore, combining (11.152) with (11.153),

$$\begin{aligned} &(\Phi_0, T(x_1, \dots, x_n) \Phi_\alpha^{\text{in}}) \\ &= i \int d^4 y \left[ \tau(x_1, \dots, x_n, y) \cdot \square_y f_\alpha(y) - \square_y \tau(x_1, \dots, x_n, y) \cdot f_\alpha(y) \right] \\ &= -i \int d^4 y f_\alpha(y) K_y \tau(x_1, \dots, x_n, y), \end{aligned} \quad (11.154)$$

where we have used  $K_y f_\alpha(y) = 0$ . The generalization of this equation is

$$\begin{aligned} (\Phi_0, T(x_1, \dots, x_n) \Phi_{\alpha_1 \dots \alpha_k}^{\text{in}}) &= (-i)^k \int d^4 y_1 \dots d^4 y_k f_{\alpha_1}(y_1) \dots f_{\alpha_k}(y_k) K_{y_1} \dots K_{y_k} \\ &\quad \times \tau(x_1, \dots, x_n, y_1, \dots, y_k) \end{aligned}$$

Generalizing further, we obtain

$$\begin{aligned}
 (\Phi_\alpha^{\text{out}}, T(x_1, \dots, x_n) \Phi_\beta^{\text{in}}) &= -i \int d^4 \eta f_{\beta_l}(\eta) K_\eta (\Phi_\alpha^{\text{out}}, T(x_1, \dots, x_n, \eta) \Phi_{\beta_1 \dots \beta_{l-1}}^{\text{in}}) \\
 &= -i \int d^4 \zeta f_{\alpha_k}^*(\zeta) K_\zeta (\Phi_{\alpha_1 \dots \alpha_{k-1}}^{\text{out}}, T(x_1, \dots, x_n, \zeta) \Phi_\beta^{\text{in}}),
 \end{aligned}
 \tag{11.155}$$

where  $\alpha = \alpha_1 \dots \alpha_k$ ,  $\beta = \beta_1 \dots \beta_l$ . We have assumed that there is no common one-particle state between  $\alpha$  and  $\beta$ . Under a similar assumption, the  $S$ -matrix element becomes

$$\begin{aligned}
 S_{\alpha\beta} &= (\Phi_\alpha^{\text{out}}, \Phi_\beta^{\text{in}}) \\
 &= (-1)^{k+l} \int d^4 \zeta_1 \dots d^4 \zeta_k d^4 \eta_1 \dots d^4 \eta_l f_{\alpha_1}^*(\zeta_1) \dots f_{\alpha_k}^*(\zeta_k) \\
 &\quad \times f_{\beta_1}(\eta_1) \dots f_{\beta_l}(\eta_l) K_{\zeta_1} \dots K_{\zeta_k} K_{\eta_1} \dots K_{\eta_l} \tau(\zeta_1 \dots \zeta_k \eta_1 \dots \eta_l).
 \end{aligned}
 \tag{11.156}$$

What we understand from this is that, when  $k = l = 2$ , the above expression is basically the same as (11.123).

Although we considered the matrix element in the above derivation, it also holds true for the operator, i.e.,

$$-i \int d^4 y f_\alpha(y) K_y T(x_1, \dots, x_n, y) = T(x_1, \dots, x_n) \varphi_\alpha^{\dagger \text{in}} - \varphi_\alpha^{\dagger \text{out}} T(x_1, \dots, x_n),
 \tag{11.157}$$

and

$$i \int d^4 y f_\alpha^*(y) K_y T(x_1, \dots, x_n, y) = T(x_1, \dots, x_n) \varphi_\alpha^{\text{in}} - \varphi_\alpha^{\text{out}} T(x_1, \dots, x_n).
 \tag{11.158}$$

Combining (11.134) and (4.18),

$$\int d^4 y \Delta(y-x) K_y T(x_1, \dots, x_n, y) = T(x_1, \dots, x_n) \varphi^{\text{in}}(x) - \varphi^{\text{out}}(x) T(x_1, \dots, x_n).
 \tag{11.159}$$

This is the operator form of the LSZ reduction formula, which corresponds to (11.120). Putting together

$$S \varphi^{\text{out}}(x) = \varphi^{\text{in}}(x) S
 \tag{11.160}$$

and (11.159),

$$\int d^4y \Delta(y-x) K_y ST(x_1, \dots, x_n, y) = [ST(x_1, \dots, x_n), \varphi^{\text{in}}(x)], \quad (11.161)$$

$$\int d^4y \Delta(x-y) K_y ST(x_1, \dots, x_n, y) = [\varphi^{\text{in}}(x), ST(x_1, \dots, x_n)]. \quad (11.161')$$

Using the above recursively,

$$\begin{aligned} \int d^4y_1 \dots d^4y_l \Delta(z_1 - y_1) \dots \Delta(z_l - y_l) K_{y_1} \dots K_{y_l} ST(x_1, \dots, x_n, y_1, \dots, y_l) \\ = [\varphi^{\text{in}}(z_1), [\varphi^{\text{in}}(z_2), [\dots [\varphi^{\text{in}}(z_l), ST(x_1, \dots, x_n)] \dots]]]. \end{aligned} \quad (11.162)$$

Taking the vacuum expectation value of this and using one of the renormalization conditions mentioned in the next chapter, viz.,

$$S\Phi_0 = \Phi_0, \quad (11.163)$$

we obtain

$$\begin{aligned} \int d^4y_1 \dots d^4y_l \Delta(z_1 - y_1) \dots \Delta(z_l - y_l) K_{y_1} \dots K_{y_l} \tau(x_1, \dots, x_n, y_1, \dots, y_l) \\ = (\Phi_0, [\varphi^{\text{int}}(z_1), [\varphi^{\text{in}}(z_2), [\dots [\varphi^{\text{in}}(z_l), ST(x_1, \dots, x_n)] \dots]] \Phi_0). \end{aligned} \quad (11.164)$$

For  $n = 0$ ,

$$\begin{aligned} \int d^4y_1 \dots d^4y_l \Delta(z_1 - y_1) \dots \Delta(z_l - y_l) K_{y_1} \dots K_{y_l} \tau(y_1, \dots, y_l) \\ = (\Phi_0, [\varphi^{\text{int}}(z_1), [\varphi^{\text{in}}(z_2), [\dots [\varphi^{\text{in}}(z_l), S] \dots]] \Phi_0). \end{aligned} \quad (11.165)$$

The operator form of the  $S$ -matrix is determined by (11.165). Expanding the  $S$ -matrix in the normal product form based on Wick's theorem, we have

$$S = \sum_{l=0}^{\infty} \frac{1}{l!} \int d^4y_1 \dots d^4y_l c(y_1, \dots, y_l) : \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_l) :, \quad (11.166)$$

where we have assumed that  $c$  is symmetric with respect to  $y_1, y_2, \dots, y_l$ . When we insert (11.166) into (11.165), what is left on the right-hand side is only the term

including the normal ordered product of  $l$  operators, whence

$$\begin{aligned}
 & (\Phi_0, [\varphi^{\text{in}}(z_1), [\dots [\varphi^{\text{in}}(z_l), S] \dots] \Phi_0) \\
 &= i^l \int d^4 y_1 \dots d^4 y_l \Delta(z_1 - y_1) \dots \Delta(z_l - y_l) c(y_1, \dots, y_l) \\
 &= \int d^4 y_1 \dots d^4 y_l \Delta(z_1 - y_1) \dots \Delta(z_l - y_l) K_{y_1} \dots K_{y_l} \tau(y_1, \dots, y_l) .
 \end{aligned}
 \tag{11.167}$$

Then  $c$  can be determined uniquely from (11.167), at least on the mass shell, i.e., for the Fourier components satisfying the Einstein energy–momentum dispersion relation. Moreover, since only the value of  $c$  on the mass shell contributes to (11.166),

$$c(y_1, \dots, y_l) = (-i)^l K_{y_1} \dots K_{y_l} \tau(y_1, \dots, y_l) .
 \tag{11.168}$$

Substituting this into (11.166),

$$S = \sum_{l=0}^{\infty} \int d^4 y_1 \dots d^4 y_l K_{y_1} \dots K_{y_l} \tau(y_1, \dots, y_l) : \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_l) : ,
 \tag{11.169}$$

where we have assumed that the term corresponding to  $l = 0$  is equal to unity. This is the operator form of the  $S$ -matrix. In addition, going back to (11.164), we have

$$\begin{aligned}
 ST(x_1, \dots, x_n) &= \sum_{l=0}^{\infty} \frac{(-i)^l}{l!} \int d^4 y_1 \dots d^4 y_l K_{y_1} \dots K_{y_l} \tau(x_1, \dots, x_n, y_1, \dots, y_l) \\
 &\quad \times : \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_l) : .
 \end{aligned}
 \tag{11.170}$$

As just described, many reduction formulae can be obtained from the asymptotic conditions. Indeed, the last formula effectively defines the quantization method for fields.

## 11.8 Unitarity Condition on the Green's Function

The unitarity of the  $S$ -matrix is obvious as long as the  $S$ -matrix element is defined by (11.136) as a transition matrix between two complete orthonormal systems  $\{\Phi^{\text{in}}\}$  and  $\{\Phi^{\text{out}}\}$ . When the asymptotic states form complete systems like this, we speak of *asymptotic completeness*. In this section, we extend the unitarity of the  $S$ -matrix from unitarity on the mass shell to unitarity off the mass shell. This can be expressed by the unitarity condition for the Green's functions.



To begin with, we consider the operator

$$T \exp \left[ -i \exp \int d^4x J(x) \varphi(x) \right]. \quad (11.171)$$

This operator is unitary, and denoting the operator for inverse time-ordering by  $\tilde{T}$ ,

$$T \exp \left[ -i \int d^4x J(x) \varphi(x) \right] \tilde{T} \exp \left[ i \int d^4y J(y) \varphi(y) \right] = 1. \quad (11.172)$$

Functionally differentiating this equation  $n$  times with respect to  $J$  and subsequently setting  $J = 0$ ,

$$\sum_{\text{comb}} (-i)^k i^{n-k} T(x'_1, \dots, x'_k) \tilde{T}(x'_{k+1}, \dots, x'_n) = 0, \quad (11.173)$$

where  $(x'_1, \dots, x'_n)$  is a permutation of  $(x_1, \dots, x_n)$  and we sum over all ways of dividing a set of  $n$  variables into two complementary subsets. We take the vacuum expectation value of this equation. Inserting the complete system  $\{\Phi^{\text{in}}\}$  between  $T$  and  $\tilde{T}$ , we use the equation

$$\begin{aligned} (\Phi_0, T(x_1, \dots, x_k) \Phi_{\alpha_1 \dots \alpha_l}^{\text{in}}) &= (-i)^l \int d^4u_1 \dots d^4u_l f_{\alpha_1}(u_1) \dots f_{\alpha_l}(u_l) \\ &\quad \times K_{u_1} \dots K_{u_l} \tau(x_1, \dots, x_k, u_1, \dots, u_l) \end{aligned} \quad (11.174)$$

and its complex conjugate

$$\begin{aligned} (\Phi_{\alpha_1 \dots \alpha_l}^{\text{in}}, \tilde{T}(x_1, \dots, x_k) \Phi_0) &= i^l \int d^4v_1 \dots d^4v_l f_{\alpha_1}^*(v_1) \dots f_{\alpha_l}^*(v_l) \\ &\quad \times K_{v_1} \dots K_{v_l} \tau(x_{k+1}, \dots, x_n, v_1, \dots, v_l). \end{aligned} \quad (11.175)$$

Then using (11.134), we sum over intermediate states. We use the notation

$$\bar{\tau}(x_1, \dots, x_n) = (-i)^n K_{x_1} \dots K_{x_n} \tau(x_1, \dots, x_n). \quad (11.176)$$

The Fourier transformation is the  $S$ -matrix element itself if all momenta are on the mass shell. Rewriting the expectation value of (11.173),

$$0 = \bar{\tau}(x_1, \dots, x_n) + \bar{\tau}^*(x_1, \dots, x_n) \quad (11.177)$$

$$+ \sum_{\text{comb}} \sum_{l=0}^{\infty} \frac{i^l}{l!} \int d^4 u_1 \dots d^4 u_l d^4 v_1 \dots d^4 v_l \bar{\tau}(x'_1, \dots, x'_k, u_1, \dots, u_l)$$

$$\times \Delta^{(+)}(u_1 - v_1) \dots \Delta^{(+)}(u_l - v_l) \bar{\tau}^*(x'_{k+1}, \dots, x'_n, v_1, \dots, v_l),$$

where  $l!$  in the denominator is a factor introduced to ensure that we do not count the same state more than once, and the prime on  $\sum'$  indicates that we neglect  $k = 0$  and  $k = n$ . Restricting all momenta to the mass shell in the Fourier transformation of this equation, it becomes the condition for unitarity. Hence, when the momenta lie outside the mass shell, the Fourier transformation can be taken as its generalization. We call (11.177) the *generalized unitarity condition*. In fact, it should be obvious from the following discussion that (11.177) yields the unitarity condition for the  $S$ -matrix on the mass shell. Using

$$S = 1 + \sum_{l=1}^{\infty} \frac{1}{l!} \int d^4 x_1 \dots d^4 x_l \bar{\tau}(x_1, \dots, x_l) : \varphi^{\text{in}}(x_1) \dots \varphi^{\text{in}}(x_l) : , \quad (11.178)$$

$$S^\dagger = 1 + \sum_{l=1}^{\infty} \frac{1}{l!} \int d^4 x_1 \dots d^4 x_l \bar{\tau}^*(x_1, \dots, x_l) : \varphi^{\text{in}}(x_1) \dots \varphi^{\text{in}}(x_l) : , \quad (11.179)$$

we expand  $SS^\dagger$  as a sum of normal products:

$$: \varphi^{\text{in}}(x_1) \dots \varphi^{\text{in}}(x_l) : : \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_m) :$$

$$= : \varphi^{\text{in}}(x_1) \dots \varphi^{\text{in}}(x_l) \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_m) :$$

$$+ \sum_{\text{comb}} i \Delta^{(+)}(x'_1 - y'_1) : \varphi^{\text{in}}(x'_2) \dots \varphi^{\text{in}}(x'_l) \varphi^{\text{in}}(y'_2) \dots \varphi^{\text{in}}(y'_m) :$$

$$+ (\text{terms with two or more contractions}). \quad (11.180)$$

Therefore, looking at the coefficients of each of the normal products, we see that

$$: \varphi^{\text{in}}(x_1) \dots \varphi^{\text{in}}(x_n) :$$

is equal to the right-hand side of (11.177). Thus,

$$SS^\dagger = 1 . \quad (11.181)$$

Similarly,

$$S^\dagger S = 1, \quad (11.182)$$

using (11.177) with  $\tau$  replaced by  $\tau^*$ . As claimed, the unitarity for the  $S$ -matrix and for Green's function are consequences of asymptotic completeness.

In the above, we considered the  $T$ -product of the Heisenberg operator. Next, we introduce the Green's functions based on the retarded product, introduced in Sect. 6.2.

## 11.9 Retarded Green's Functions

If  $A(x)$  is a local field, its retarded product is defined by

$$\begin{aligned} R[A(x) : \varphi(x_1) \dots \varphi(x_n)] & \quad (11.183) \\ & = (-i)^n \sum_p \theta(x - x'_1) \dots \theta(x'_{n-1} - x'_n) [\dots [A(x), \varphi(x'_1)] \dots \varphi(x'_n)], \end{aligned}$$

where  $\theta(x)$  stands for  $\theta(x_0)$ , with  $x'_1, \dots, x'_n$  a permutation of  $x_1, \dots, x_n$  and summation over all permutations. The only permutations to contribute are those satisfying  $x'_1 > x'_2 > \dots > x'_n$  for the time variables. We introduce the unitary operator (11.171), denoting it by  $U$ :

$$U = T \exp \left[ -i \int d^4x J(x) \varphi(x) \right]. \quad (11.184)$$

Therefore,

$$U^{-1} = U^\dagger = \tilde{T} \exp \left[ i \int d^4x J(x) \varphi(x) \right]. \quad (11.185)$$

We now introduce the generating functional

$$A_R[x, J] = U^\dagger T[UA(x)]. \quad (11.186)$$

Therefore, it is easy to check that the R-product above can be expressed by

$$R[A(x) : \varphi(x_1) \dots \varphi(x_n)] = \frac{\delta^n A_R[x, J]}{\delta J(x_1) \dots \delta J(x_n)} \Big|_{J=0}. \quad (11.187)$$

Directly from the definition,

$$\frac{\delta}{\delta J(y)} A_R[x, J] = -i\theta(x-y)[A_R[x, J], \varphi_R[y, J]]. \quad (11.188)$$

In particular, taking  $A = \varphi$ , we have

$$\frac{\delta}{\delta J(y)} \varphi_R[x, J] - \frac{\delta}{\delta J(x)} \varphi_R[y, J] + i[\varphi_R[x, J], \varphi_R[y, J]] = 0. \quad (11.189)$$

This is called a *unitarity condition*. It corresponds to (11.173) in the case of the T-product. In addition, functionally differentiating  $A_R$  a total of  $n$  times with respect to  $J$ ,

$$\begin{aligned} R[A(x) : \varphi(x_1) \dots \varphi(x_n)] & \quad (11.190) \\ &= \sum_{\text{comb}} i^k \tilde{T}[\varphi(x'_1) \dots \varphi(x'_k)] (-i)^{n-k} T[\varphi(x'_{k+1}) \dots \varphi(x'_n) A(x)]. \end{aligned}$$

Then taking the Hermitian conjugate of the reduction formula for the T-product, viz.,

$$\int d^4y \Delta(x-y) K_y ST(x_1, \dots, x_n, y) = [\varphi^{\text{int}}(x), ST(x_1, \dots, x_n)], \quad (11.191)$$

we obtain

$$\int d^4y \Delta(x-y) K_y \tilde{T}(x_1, \dots, x_n, y) S^\dagger = -[\varphi^{\text{int}}(x), \tilde{T}(x_1, \dots, x_n) S^\dagger]. \quad (11.192)$$

Combining the three equations above,

$$\int d^4y \Delta(x-y) K_y R(w : x_1, \dots, x_n, y) = -i[\varphi^{\text{int}}(x), R(w : x_1, \dots, x_n)], \quad (11.193)$$

where

$$R(w : x_1, \dots, x_n, y) = R[A(w) : \varphi(x_1) \dots \varphi(x_n)]. \quad (11.194)$$

Then using (11.193) iteratively,

$$\begin{aligned} & \int d^4 y_1 \dots d^4 y_l \Delta(z_1 - y_1) \dots \Delta(z_l - y_l) K_{y_1} \dots K_{y_l} R(w : x_1, \dots, x_n, y_1, \dots, y_l) \\ & = (-i)^l [\varphi^{\text{in}}(z_1), [\dots [\varphi^{\text{in}}(z_l), R(w : x_1, \dots, x_n)] \dots]] . \end{aligned} \quad (11.195)$$

We now expand  $R$  as a sum of normal products:

$$R(w : x_1, \dots, x_n) = \sum_{l=0}^{\infty} \frac{1}{l!} \int d^4 y_1 \dots d^4 y_l f(w : x_1, \dots, x_n, y_1, \dots, y_l) : \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_l) : , \quad (11.196)$$

where we have assumed that  $f$  is a symmetric function with respect to  $y_1, \dots, y_l$ . Inserting this into (11.195) and taking the vacuum expectation value,

$$\begin{aligned} & (-i)^l (\Phi_0, [\varphi^{\text{in}}(z_1), [\dots [\varphi^{\text{in}}(z_l), R(w : x_1, \dots, x_n)] \dots]] \Phi_0) \\ & = \int d^4 y_1 \dots d^4 y_l \Delta(z_1 - y_1) \dots \Delta(z_l - y_l) f(w : x_1, \dots, x_n, y_1, \dots, y_l) . \end{aligned} \quad (11.197)$$

However, from (11.195), the right-hand side is equivalent to

$$\int d^4 y_1 \dots d^4 y_l \Delta(z_1 - y_1) \dots \Delta(z_l - y_l) r(w : x_1, \dots, x_n, y_1, \dots, y_l) , \quad (11.198)$$

where

$$r(w : x_1, \dots, x_n, y_1, \dots, y_l) = K_{x_1} \dots K_{x_n} (\Phi_0, R(w : x_1, \dots, x_n) \Phi_0) . \quad (11.199)$$

Thus, if the momenta corresponding to  $y_1, \dots, y_l$  are on the mass shell,

$$f(w : x_1, \dots, x_n, y_1, \dots, y_l) = r(w : x_1, \dots, x_n, y_1, \dots, y_l) . \quad (11.200)$$

Since only those on the mass shell exert any influence,

$$\begin{aligned} & R(w : x_1, \dots, x_n) \\ & = \sum_{l=0}^{\infty} \frac{1}{l!} \int d^4 y_1 \dots d^4 y_l r(w : x_1, \dots, x_n, y_1, \dots, y_l) : \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_l) : . \end{aligned} \quad (11.201)$$

In particular, for  $n = 0$ ,

$$A(w) = \sum_{l=0}^{\infty} \frac{1}{l!} \int d^4y_1 \dots d^4y_l r(w : y_1, \dots, y_l) : \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_l) : . \tag{11.202}$$

Moreover, if we take  $A = \varphi$ , then when  $(\Phi_0, \varphi(x)\Phi_0) = 0$ , we have

$$\varphi(x) = \varphi^{\text{in}}(x) + \sum_{l=2}^{\infty} \frac{1}{l!} \int d^4y_1 \dots d^4y_l r(x : y_1, \dots, y_l) : \varphi^{\text{in}}(y_1) \dots \varphi^{\text{in}}(y_l) : . \tag{11.203}$$

This gives the formal solution to the Yang–Feldman equation introduced in Sect. 6.3. In addition, from (11.189),

$$R(x : y, x_1, \dots, x_n) - R(y : x, x_1, \dots, x_n) + i \sum_{\text{comb}} [R(x : x'_1, \dots, x'_k), R(y : x'_{k+1}, \dots, x'_n)] = 0 . \tag{11.204}$$

Taking the vacuum expectation value of this equation and using the reduction formula obtained by inserting the complete system  $\{\Phi^{\text{in}}\}$ , we obtain a non-linear equation for the system  $(\Phi_0, R(x : x_1, \dots, x_n)\Phi_0)$ . This is also one of the generalized unitarity conditions.

Both the in- and the out-states appear in the reduction formula for the T-product, while only the in-states appear in the R-products.

# Chapter 12

## Renormalization Theory



So far we have shown only the lowest order calculations, but when computing higher order corrections, divergences must show up. A method for deducing finite consequences by a suitable interpretation is called a *renormalization theory*. As mentioned once or twice before, the basic idea of such a formalism is to specify a way of separating the Lagrangian density into the free part and the interaction part. This grouping is related to the definition of the interaction picture. We call the interaction picture defined by the correct grouping a *renormalized interaction picture*. Several properties of the Green's functions discussed in the previous chapter hold true in the renormalized interaction picture, while they may not hold true in other pictures.

In this chapter, in order to show that such a separation of the Lagrangian or the Hamiltonian is not necessarily trivial, we first review the scattering theory in non-relativistic quantum mechanics. The formal system discussed here has many similarities with the  $S$ -matrix theory based on the reduction formula given in the last chapter.

### 12.1 Lippmann–Schwinger Equation

In this section, we introduce the formal logic for the standard quantum mechanical system. This theory has a lot in common with the theory of Green's functions. What is important in the scattering problem is the way we formulate the boundary conditions. This issue is discussed in detail in my book, "Relativistic Quantum Mechanics" [78], but here we shall present it in a slightly different order.

In this section, we formulate the scattering problem using the notion of Stueckelberg causality:

Assuming that the potential  $V(t)$  is a function of time, if  $V(t) = 0$  in the region  $t < T$ , then there exists no scattered wave for  $t < T$ .

The principle above has already been used for the quantization of free fields in Sect. 4.3 and for the derivation of the Yang–Feldman equation in Sect. 6.3. Here we apply it to the non-relativistic formulation. We consider the Schrödinger equation

$$i \frac{\partial}{\partial t} \psi(t) = [H_0 + V(t)] \psi(t) , \quad (12.1)$$

where we have omitted the spatial coordinate for simplicity, and  $V(t)$  is defined by

$$V(t) = \begin{cases} V , & t > T , \\ 0 , & t < T . \end{cases} \quad (12.2)$$

For  $t < T$ ,

$$i \frac{\partial}{\partial t} \psi(t) = H_0 \psi(t) . \quad (12.3)$$

Since this is the equation for the incident wave, we write its solution as

$$\psi(t) = \psi_{\text{in}}(t) , \quad H_0 \psi_{\text{in}}(t) = E \psi_{\text{in}}(t) . \quad (12.4)$$

Therefore, the scattered wave appears at a generic time  $t$  and its equation is

$$\psi(t) = \psi_{\text{in}}(t) + \psi_{\text{scatt}}(t) , \quad (12.5)$$

$$\left( i \frac{\partial}{\partial t} - H_0 \right) \psi_{\text{scatt}}(t) = V(t) \psi(t) . \quad (12.6)$$

To solve this equation, we introduce the Green's function, which is the solution of the equation

$$\left( i \frac{\partial}{\partial t} - H_0 \right) K_{\text{ret}}(t, \mathbf{x} : t', \mathbf{x}') = \delta(t - t') \delta^3(\mathbf{x} - \mathbf{x}') , \quad (12.7)$$

$$K_{\text{ret}}(t, \mathbf{x} : t', \mathbf{x}') = 0 , \quad t < t' . \quad (12.8)$$

Therefore, the solution of (12.6) satisfying the causality condition is

$$\psi_{\text{scatt}}(t, \mathbf{x}) = \int dt' d^3 x' K_{\text{ret}}(t, \mathbf{x} : t', \mathbf{x}') V(t', \mathbf{x}') \psi(t', \mathbf{x}') . \quad (12.9)$$

Here, taking the limit  $T \rightarrow -\infty$ , the  $t$ -dependence of the potential disappears, whence

$$\psi_{\text{scatt}}(t, \mathbf{x}) = \int dt' d^3 x' K_{\text{ret}}(t, \mathbf{x} : t', \mathbf{x}') V(\mathbf{x}') \psi(t', \mathbf{x}') . \quad (12.10)$$



Then, if  $H_0$  does not depend on the space-time coordinate,

$$K_{\text{ret}}(t, \mathbf{x} : t', \mathbf{x}') = K_{\text{ret}}(t - t', \mathbf{x} - \mathbf{x}') . \quad (12.11)$$

Thus,

$$\psi(t, \mathbf{x}) = \psi_{\text{in}}(t, \mathbf{x}) + \int dt' d^3x' K_{\text{ret}}(t - t', \mathbf{x} - \mathbf{x}') V(\mathbf{x}') \psi(t', \mathbf{x}') . \quad (12.12)$$

This is the integral equation governing the scattering, and corresponds to the Yang–Feldman equation. To remove the time variable, we take

$$\psi_{\text{in}}(t, \mathbf{x}) = e^{-iEt} \psi_{\text{in}}(\mathbf{x}) , \quad \psi(t, \mathbf{x}) = e^{-iEt} \psi(\mathbf{x}) , \quad (12.13)$$

leading to

$$\psi(\mathbf{x}) = \psi_{\text{in}}(\mathbf{x}) + \int d^3x' G(\mathbf{x} - \mathbf{x}' : E) V(\mathbf{x}') \psi(\mathbf{x}') , \quad (12.14)$$

where

$$G(\mathbf{x} - \mathbf{x}' : E) = \int dt' e^{iE(t-t')} K_{\text{ret}}(t - t', \mathbf{x} - \mathbf{x}') . \quad (12.15)$$

We now introduce the Fourier representation of the retarded Green's function:

$$K_{\text{ret}}(t, \mathbf{x}) = \frac{1}{(2\pi)^4} \int dE d^3p e^{ipx - iEt} K(E, \mathbf{p}) . \quad (12.16)$$

Furthermore, we restrict  $H_0$  to the form

$$H_0 = \frac{\mathbf{p}^2}{2m} . \quad (12.17)$$

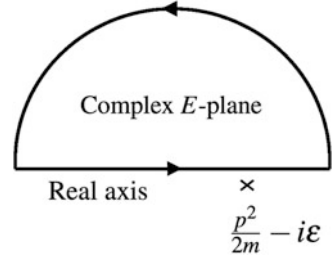
Thus, the equation for  $K(E, \mathbf{p})$  obtained from (12.7) can be written in the form

$$\left( E - \frac{\mathbf{p}^2}{2m} \right) K(E, \mathbf{p}) = 1 . \quad (12.18)$$

As will be shown later, the solution of this equation satisfying the boundary condition (12.8) is

$$K(E - \mathbf{p}) = \left( E - \frac{\mathbf{p}^2}{2m} + i\epsilon \right)^{-1} , \quad (12.19)$$

**Fig. 12.1** Path of  $E$ -integration in (12.22)



where  $\epsilon$  is a positive infinitesimal. Taking  $E = k^2/2m$ , Eqs. (12.15) and (12.16) imply

$$G(\mathbf{x} - \mathbf{x}' : E) = -\frac{m}{2\pi} \frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} . \quad (12.20)$$

Thus, (12.14) can be written in the form

$$\psi(\mathbf{x}) = \psi_{\text{in}}(\mathbf{x}) - \frac{m}{2\pi} \int d^3x' \frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} V(\mathbf{x}') \psi(\mathbf{x}') . \quad (12.21)$$

This equation indicates that the scattered wave becomes an outward-directed spherical wave as a consequence of causality. When  $t - t' < 0$ ,

$$K_{\text{ret}}(t - t', \mathbf{x} - \mathbf{x}') = \frac{1}{(2\pi)^4} \int d^3p e^{ip(\mathbf{x}-\mathbf{x}')} \int dE \frac{e^{-iE(t-t')}}{E - \frac{p^2}{2m} + i\epsilon} . \quad (12.22)$$

Then, since  $-iE(t - t') = iE|t - t'|$ , if  $\text{Im } E > 0$ , this exponential decreases exponentially with  $\text{Im } E$ . Hence, selecting the integration path as in Fig. 12.1 and taking into account the fact that the pole is located outside the semicircle, it vanishes by Cauchy's theorem. Thus, if  $t - t' < 0$ , then  $K_{\text{ret}} = 0$  and we see that (12.8) is satisfied.

Clearly, using (12.22) and taking  $\Psi = \psi$  and  $\Phi = \psi_{\text{in}}$ , Eq. (12.21) can be written formally as

$$\Psi = \Phi + \frac{1}{E - H_0 + i\epsilon} V\Psi \quad (E > 0) , \quad (12.23)$$

In fact, looking at the Fourier representation of the retarded Green's function (12.22), we see that this corresponds to the operator which gives the energy denominator in the equation above. From (12.23) and (12.3),

$$(E - H_0)\Psi = V\Psi , \quad (E - H_0)\Phi = 0 . \quad (12.24)$$

Moreover, for the bound state ( $E < 0$ ), the inverse of  $(E - H_0)$  is uniquely defined, so we obtain the homogenous equation

$$\Psi = \frac{1}{E - H_0} V \Phi \quad (E < 0). \quad (12.25)$$

The pair of Eqs.(12.23) and (12.25) is known collectively as the *Lippmann–Schwinger equation* [114]. The solution of (12.23) consists of the incident wave and the outward-going spherical wave, and it can be written as

$$\Psi_a^{(+)} = \Phi_a + \frac{1}{E_a - H_0 + i\epsilon} V \Psi_a^{(+)} . \quad (12.26)$$

Mathematically, it is useful to consider the solution which consists of the incident wave and the inward-going spherical wave. The equation for such a state can be obtained by replacing  $i\epsilon$  with  $-i\epsilon$  :

$$\Psi_a^{(-)} = \Phi_a + \frac{1}{E_a - H_0 - i\epsilon} V \Psi_a^{(-)} . \quad (12.27)$$

The two cases are solutions of the same Schrödinger equation under different boundary conditions:

$$(E_a - H_0)\Psi_a^{(\pm)} = V\Psi_a^{(\pm)} . \quad (12.28)$$

Thus,  $\{\Psi_a^{(+)}\}$  and  $\{\Psi_a^{(-)}\}$  form complete systems separately. Solving (12.26) sequentially,

$$\begin{aligned} \Psi_a^{(+)} &= \Phi_a + \frac{1}{E_a - H_0 + i\epsilon} \left( 1 + V \frac{1}{E_a - H_0 + i\epsilon} + \dots \right) V \Phi_a \\ &= \Phi_a + \frac{1}{E_a - H_0 + i\epsilon} \left( 1 - V \frac{1}{E_a - H_0 + i\epsilon} \right)^{-1} V \Phi_a . \end{aligned} \quad (12.29)$$

Then using  $A^{-1}B^{-1} = (BA)^{-1}$ ,

$$\Psi_a^{(+)} = \Phi_a + \frac{1}{E_a - H + i\epsilon} V \Phi_a \quad (H = H_0 + V) . \quad (12.30)$$

We call this the *Chew–Goldberger formal solution* [115]. Another solution is

$$\Psi_a^{(-)} = \Phi_a + \frac{1}{E_a - H - i\epsilon} V \Phi_a . \quad (12.31)$$

Although this formal solution is in fact not useful for solving the problem, it is very useful for deriving the general properties of the  $S$ -matrix, and we will show this application below.

We assume that the state of the incident wave is suitably normalized, i.e.,

$$(\Phi_b, \Phi_a) = \delta_{ba} . \quad (12.32)$$

Therefore, combining the LS equation with the CG formal solution,

$$\begin{aligned} (\Psi_b^{(+)}, \Psi_a^{(+)}) &= \left( \Phi_b + \frac{1}{E_b - H + i\epsilon} V \Phi_b, \Psi_a^{(+)} \right) \\ &= (\Phi_b, \Psi_a^{(+)}) + \left( \Phi_b, V \frac{1}{E_b - H - i\epsilon} \Psi_a^{(+)} \right) \\ &= \left( \Phi_b, \Phi_a + \frac{1}{E_a - H_0 + i\epsilon} V \Psi_a^{(+)} \right) + \left( \Phi_b, V \frac{1}{E_b - H - i\epsilon} \Psi_a^{(+)} \right) \\ &= (\Phi_b, \Phi_a) + \left( \frac{1}{E_a - E_b + i\epsilon} + \frac{1}{E_b - E_a - i\epsilon} \right) (\Phi_b, V \Psi_a^{(+)}) \\ &= (\Phi_b, \Phi_a) , \end{aligned} \quad (12.33)$$

so we see that  $\{\Psi^{(+)}\}$  forms an orthonormal system just as  $\{\Phi\}$  does. The same is true of  $\{\Psi^{(-)}\}$ . Thus, the transformation matrix between these two pairs of complete orthonormal systems, viz.,

$$S_{ab} = (\Psi_b^{(-)}, \Psi_a^{(+)}) , \quad (12.34)$$

is unitary. Comparing with (11.136),  $\Psi^{(+)}$  and  $\Psi^{-}$  correspond to  $\Phi^{\text{in}}$  and  $\Phi^{\text{out}}$ , respectively. The unitarity condition is

$$S^\dagger S = S S^\dagger = 1 . \quad (12.35)$$

Starting with the definition and modifying it suitably, the  $S$ -matrix above can be written as

$$\begin{aligned} S_{ba} &= \left( \Phi_b + \frac{1}{E_b - H - i\epsilon} V \Phi_b, \Psi^{(+)}_a \right) \\ &= (\Phi_b, \Psi^{(+)}_a) + \left( \Phi_b, V \frac{1}{E_b - H + i\epsilon} \Psi^{(+)}_a \right) \\ &= (\Phi_b, \Phi_a) + \left( \Phi_b, \frac{1}{E_a - H_0 + i\epsilon} V \Psi^{(+)}_a \right) + \left( \Phi_b, V \frac{1}{E_b - H + i\epsilon} \Psi^{(+)}_a \right) \\ &= \delta_{ba} + \left( \frac{1}{E_a - E_b + i\epsilon} + \frac{1}{E_b - E_a + i\epsilon} \right) (\Phi_b, V \Psi^{(+)}_a) \\ &= \delta_{ba} - 2\pi i \delta(E_b - E_a) (\Phi_b, V \Psi^{(+)}_a) . \end{aligned} \quad (12.36)$$

Although in the transformation above we have expressed  $\Psi^{(-)}$  in terms of  $\Phi$ , expressing  $\Psi^{(+)}$  in terms of  $\Phi$ , we obtain

$$S_{ba} = \delta_{ba} - 2\pi i \delta(E_b - E_a) (\Psi_b^{(-)}, V\Phi_a). \quad (12.37)$$

Therefore, if  $E_b = E_a$ , the transition amplitude  $T_{ba}$  can be written in a symmetric form:

$$T_{ba} = (\Phi_b, V\Psi_a^{(+)}) = (\Psi_b^{(-)}, V\Phi_a). \quad (12.38)$$

We now express the unitarity of the  $S$ -matrix in terms of  $T$ . Using  $T$  as above,

$$\begin{aligned} T_{ba}^\dagger &= T_{ab}^* = (\Phi_a, V\Psi_b^{(+)})^* = (\Psi_b^{(+)}, V\Phi_a), \\ T_{ba}^\dagger - T_{ba} &= (\Psi_b^{(+)}, V\Phi_a) - (\Phi_b, V\Psi_a^{(+)}) \\ &= (\Phi_b, V\Phi_a) + \left( \frac{1}{E_b - H + i\epsilon} V\Phi_b, V\Phi_a \right) \\ &\quad - (\Phi_b, V\Phi_a) - \left( \Phi_b, V \frac{1}{E_a - H + i\epsilon} V\Phi_a \right) \\ &= \left( V\Phi_b, \left( \frac{1}{E_b - H - i\epsilon} - \frac{1}{E_b - H + i\epsilon} \right) V\Phi_a \right) \\ &= 2\pi i (V\Phi_b, \delta(E_b - H) V\Phi_a), \end{aligned} \quad (12.39)$$

then inserting the complete system  $\{\Psi^{(-)}\}$ , we obtain

$$\begin{aligned} T_{ba}^\dagger - T_{ba} &= 2\pi i \sum_n (V\Phi_b, \Psi_n^{(-)}) \delta(E_b - E_n) (\Psi_n^{(-)}, V\Phi_a) \\ &= 2\pi i \sum_n T_{bn}^\dagger \delta(E_b - E_n) T_{na}, \end{aligned} \quad (12.40)$$

where

$$T_{bn}^\dagger = T_{nb}^* = (\Psi_n^{(-)}, V\Phi_b)^* = (V\Phi_b, \Psi_n^{(-)}).$$

If we inserted the other complete system  $\{\Psi^{(+)}\}$  instead of  $\{\Psi^{(-)}\}$ , then instead of (12.40) we would get

$$2\pi i \sum_n T_{bn} \delta(E_b - E_n) T_{na}^\dagger. \quad (12.41)$$

Thus, the unitarity condition can be written as

$$\begin{aligned} T_{ba}^\dagger - T_{ba} &= 2\pi i \sum_n T_{bn}^\dagger \delta(E_b - E_n) T_{na} \\ &= 2\pi i \sum_n T_{bn} \delta(E_b - E_n) T_{na}^\dagger . \end{aligned} \quad (12.42)$$

Using this unitarity condition, we can reproduce the optical theorem already discussed in Sect. 9.4. Although the content is exactly the same, let us express it in terms of the notation used in this section.

The probability per unit time for the transition  $a \rightarrow b$  is

$$w_{ba} = 2\pi \delta(E_b - E_a) |T_{ba}|^2 . \quad (12.43)$$

If we take the sum over all probable final states, then from (12.42),

$$w_a = \sum_b w_{ba} = 2\pi \sum_b \delta(E_b - E_a) |T_{ba}|^2 = \frac{1}{i} (T_{aa}^\dagger - T_{aa}) = -2\text{Im}T_{aa} . \quad (12.44)$$

Starting with the two-particle state  $a$ , the total cross-section can be obtained as

$$\sigma_a = \frac{\Omega}{v_{\text{rel}}} w_a , \quad (12.45)$$

where  $\Omega$  is the volume of quantization and  $v_{\text{rel}}$  is the relative speed of the two particles. Therefore, from (12.44),

$$\sigma_a = -\frac{2\Omega}{v_{\text{rel}}} \text{Im}T_{aa} . \quad (12.46)$$

Then, since the final result does not depend on  $\Omega$ , we take  $\Omega = 1$ . The scattering amplitude  $f(\theta)$  in the scattering potential is given by

$$\begin{aligned} f(\theta) &= -\frac{m}{2\pi} \int d^3x e^{-ik_f x} V(x) \psi(x) \\ &= -\frac{m}{2\pi} (\Phi_f, V \Psi_i^{(+)}). \end{aligned} \quad (12.47)$$

Computing  $w_{fi}/v_{\text{rel}}$  for scattering into a constant solid angle, the cross-section becomes

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{2\pi}{v} \frac{1}{(2\pi)^3} \int k_f^2 dk_f \delta\left(\frac{k_f^2}{2m} - \frac{k_i^2}{2m}\right) |T_{fi}|^2 \quad \left(v = \frac{k}{m}\right) \\ &= \left(\frac{m}{2\pi}\right)^2 |T_{fi}|^2 \\ &= |f(\theta)|^2, \end{aligned} \quad (12.48)$$

which reproduces the well known result. In this regard, however,  $T_{fi} = (\Phi_f, V\Psi_i^{(+)})$ . In addition, from (12.46), the total cross-section becomes

$$\sigma = -\frac{2}{v} \text{Im}T_{aa} = -\frac{2}{v} \left(-\frac{2\pi}{m}\right) \text{Im}f(0) = \frac{4\pi}{k} \text{Im}f(0), \quad (12.49)$$

which also reproduces the optical theorem in its well-known form.

Although in the scattering potential the asymptotic form of the wave function has been obtained easily, it turns out that the asymptotic form of the abstract Lippmann–Schwinger state vector  $\Psi^{(+)}$  is given by

$$\Psi_a^{(+)} \sim S\Phi_a = \Phi_a - 2\pi i\delta(E_a - H_0)V\Psi_a^{(+)}. \quad (12.50)$$

Let us compare this asymptotic form with  $\Psi_a^{(+)}$  itself:

$$\Psi_a^{(+)} = \Phi_a + \frac{1}{E_a - H_0 + i\epsilon} V\Psi_a^{(+)}. \quad (12.51)$$

From this we understand that the asymptotic form can be derived if we make the following replacement for the scattered wave:

$$\frac{1}{E_a - H_0 + i\epsilon} \rightarrow -2\pi i\delta(E_a - H_0), \quad (12.51)$$

or

$$\Psi_a^{(+)} \sim \Phi_a - 2\pi i\delta(E_a - H_0)[(E_a - H_0)\Psi_a^{(+)}], \quad (12.52)$$

where we multiply by  $\delta(E_a - H_0)$  after multiplying by  $(E_a - H_0)$ . This operation reminds us of the Lehmann–Symanzik–Zimmermann (LSZ) asymptotic condition in Sect. 11.7. In (11.159),  $K_y$  corresponds to  $(E_a - H_0)$ , and  $\Delta(y - x)$  corresponds to  $\delta(E_a - H_0)$ . Moreover, the asymptotic form satisfies the equation for the free

particle, viz.,

$$(E_a - H_0)S\Phi_a = 0 . \quad (12.53)$$

This corresponds to (11.124).

Equation (12.52) is the reduction formula. It plays an important role in the discussion about recombination reactions. So far we have assumed that the separation of the Hamiltonian into the free part and the interaction part is unique, but in general, if a bound state appears, this uniqueness is lost. It is due to recombination reactions that this grouping in the initial state differs from that in the final state. For instance, consider a reaction such as

$$n + d \rightarrow n + n' + p , \quad (12.54)$$

where  $d$  stands for the deuteron and  $n'$  has been labeled by the prime in order to distinguish it from the other meson  $n$ . The total Hamiltonian is

$$H = T_p + T_n + T_{n'} + V_{np} + V_{n'p} + V_{nn'} , \quad (12.55)$$

where  $T$  and  $V$  denote the kinetic energies and the potentials for the two-body forces, respectively. The decomposition of the Hamiltonian corresponding to the initial state is

$$H_0 = T_p + T_n + T_{n'} + V_{n'p} , \quad V = V_{np} + V_{nn'} . \quad (12.56)$$

This is because, if we do not insert  $V_{n'p}$  into  $H_0$ , there is no way to make  $d$ . On the other hand, the decomposition in the final state is

$$H'_0 = T_p + T_n + T_{n'} , \quad V' = V_{np} + V_{n'p} + V_{nn'} . \quad (12.57)$$

In general, we introduce two decompositions, one for the initial state and one for the final state:

$$H = H_a + V_a = H_b + V_b . \quad (12.58)$$

The free state vectors corresponding to each decomposition are  $\Phi_a$  and  $\Phi_b$  satisfying

$$(E_a - H_a)\Phi_a = 0 , \quad (E_b - H_b)\Phi_b = 0 . \quad (12.59)$$

Of course, in order for the transition  $a \rightarrow b$  to occur, we must have  $E_b = E_a$ . Let us derive the transition amplitude  $T_{ba}$  in this case. Corresponding to the initial state,



we have

$$\Psi_a^{(+)} = \Phi_a + \frac{1}{E_a - H + i\epsilon} V_a \Phi_a . \quad (12.60)$$

Then, to construct the asymptotic form corresponding to the final state, we use the formula

$$\frac{1}{A} - \frac{1}{B} = \frac{1}{B} (B - A) \frac{1}{A} . \quad (12.61)$$

Therefore,

$$\frac{1}{E_a - H + i\epsilon} - \frac{1}{E_a - H_b + i\epsilon} = \frac{1}{E_a - H_b + i\epsilon} V_b \frac{1}{E_a - H + i\epsilon} , \quad (12.62)$$

which yields

$$\Psi_a^{(+)} = \Phi_a + \frac{1}{E_a - H_b + i\epsilon} \left( 1 + V_b \frac{1}{E_a - H + i\epsilon} \right) V_a \Phi_a . \quad (12.63)$$

In the following, we write  $E_a = E_b = E$  and make the replacement

$$\frac{1}{E - H_b + i\epsilon} \longrightarrow -2\pi i \delta(E - H_b) , \quad (12.64)$$

corresponding to (12.51). Then the asymptotic form corresponding to the final state is

$$\Psi_a^{(+)} \sim -2\pi i \left( 1 + V_b \frac{1}{E - H + i\epsilon} \right) V_a \Phi_a . \quad (12.65)$$

Thus,  $T_{ba}$  can be given by the inner product of this asymptotic form and  $\Phi_b$ :

$$\begin{aligned} T_{ba} &= \left( \Phi_b, \left( 1 + V_b \frac{1}{E - H + i\epsilon} \right) V_a \Phi_a \right) \\ &= \left( \left( 1 + \frac{1}{E - H - i\epsilon} V_b \right) \Phi_b, V_a \Phi_a \right) \\ &= (\Psi_b^{(-)}, V_a \Phi_a) . \end{aligned} \quad (12.66)$$

Note also that, corresponding to (12.38), there are also two ways to represent  $T_{ba}$ . Using  $\Psi_a^{(+)}$  instead of  $\Psi_b^{(-)}$ ,

$$T_{ba} = (\Psi_b^{(-)}, V_a \Phi_a) = (\Phi_b, V_b \Psi_a^{(+)}) . \quad (12.67)$$

This is because, subtracting one from the other,

$$\begin{aligned}
 (\Psi_b^{(-)}, V_a \Phi_a) - (\Phi_b, V_b \Psi_a^{(+)}) &= (\Phi_b, V_a \Phi_a) - (\Phi_b, V_b \Phi_a) \\
 &\quad + (\Phi_b, V_a \Phi_a) - \left( \Phi_b, V_b \frac{1}{E - H + i\epsilon} V_a \Phi_a \right) \\
 &= (\Phi_b, (V_a - V_b) \Phi_a) \\
 &= (\Phi_b, (H_b - H_a) \Phi_a) \\
 &= (E_b - E_a) (\Phi_b, \Phi_a) = 0 .
 \end{aligned}$$

This shows that the two expressions in (12.67) are equal.

## 12.2 Renormalized Interaction Picture

In Chap. 8, we described the computational method for obtaining the  $S$ -matrix in the interaction picture based on a covariant perturbation theory. To define the interaction picture, we have to decompose the Hamiltonian or the Lagrangian into the free part and the interaction part. We took this decomposition to be trivial, but it is clear from the discussion about the recombination reaction in the previous section that this decomposition is not unique. Although at the lowest level of the perturbation it has not posed serious problems, it will turn out that this difference between decomposition methods has an important implication when computing higher order corrections.

We thus set down several conditions to determine the decomposition method. These conditions are called *renormalization conditions*. The interaction picture defined by the decomposition satisfying these conditions is called the *renormalized interaction picture*. In fact, it is in the renormalized interaction picture that the Gell-Mann–Low relation derived in Chap. 11, the related asymptotic conditions, and so on, all hold true, although we have not stated this clearly up to now. Another aspect of renormalization, and in general only this aspect is emphasized, is that it can remove the divergences appearing in higher order corrections. Indeed, it was through this that, in the period after World War II, there was a major development of QED. A theory in which divergences can be removed in this way is said to be *renormalizable*, and renormalizability has been promoted as one of the guiding principles. It was also an important motivation for the more recent development of gauge theories.

Let us now go back to the problem of the decomposition of the Hamiltonian:

1. We have two complete orthonormal systems of eigenstates of the Hamiltonian:

$$\{ \Psi_a^{(+)} \} \quad \text{and} \quad \{ \Psi_a^{(-)} \} . \tag{12.68}$$

Furthermore, in relativistic quantum mechanics, these are eigenstates of the four-momentum:

$$P_\mu \Psi_a^{(\pm)} = (p_\mu)_a \Psi_a^{(\pm)}. \quad (12.69)$$

The  $S$ -matrix element is given by

$$S_{ab} = (\Psi_b^{(-)}, \Psi_a^{(+)}). \quad (12.70)$$

It is clear from this expression for the  $S$ -matrix element that these two complete systems can be considered to be the same as

$$\{\Phi_a^{\text{in}}\} \text{ and } \{\Phi_a^{\text{out}}\}. \quad (12.71)$$

2. In field theory, the vacuum state  $\Phi_0$  and the stable one-particle state  $\Phi_\alpha$  satisfy the conditions

$$\Phi_0^{\text{in}} = \Phi_0^{\text{out}}, \quad \Phi_\alpha^{\text{in}} = \Phi_\alpha^{\text{out}}. \quad (12.72)$$

If we express these conditions in terms of the  $S$ -matrix, then

$$S\Phi_0 = \Phi_0, \quad S\Phi_\alpha = \Phi_\alpha, \quad (12.73)$$

where (12.72) implies that  $\Phi_0$  and  $\Phi_\alpha$  have been written without distinguishing between the in-state and the out-state. Equations (12.72) and (12.73) are called *renormalization conditions*.

Then, since the condition on the vacuum is satisfied by (8.70), this means that the  $S$ -matrix is defined by dropping all bubble diagrams. It turns out that we define the interaction picture in such a way as to satisfy these conditions. This requires as a consequence reintroducing several kinds of physical observable using the following two kinds of renormalization

- **Mass Renormalization.** In field theory, no elementary particles are in the bare state, because they have self-interactions. For example, even within the framework of classical theory, a charged particle carries the Coulomb field. Thus, considering an electron, the conditions (12.73) should hold true for the electron carrying its own field and not for the bare electron. Due to the existence of the self-interaction, the mass of an electron increases by what we call the *self-energy*  $\delta m$ . Thus, the observable mass of an electron is not just the originally given mass  $m$ , but changes to  $m + \delta m$ . We call  $m$  and  $m + \delta m = m_{\text{obs}}$  the *bare mass* and the *observed mass*, respectively. We interpret this by saying that the one-electron state is the state with mass  $m_{\text{obs}}$ . We determine the self-energy  $\delta m$  from (12.73), i.e., we consider  $m_{\text{obs}}$  as the given mass rather than  $m$ . In the sense of absorbing  $\delta m$  into the mass, we call this reinterpretation a *renormalization of the mass*.

- **Charge Renormalization.** As discussed in Chap. 8, due to the phenomena of vacuum polarization, the vacuum behaves like a dielectric medium in field theory. Thus, taking the permittivity of the vacuum as  $\epsilon$ , when the distance  $r$  between two charges  $e_1$  and  $e_2$  is large enough, the Coulomb potential between them is

$$V = \frac{e_1 e_2}{\epsilon} \frac{1}{4\pi r}. \quad (12.74)$$

However, since the permittivity of the vacuum is normalized to unity,

$$V = (e_1)_{\text{obs}}(e_2)_{\text{obs}} \frac{1}{4\pi r}. \quad (12.75)$$

Comparing (12.74) and (12.75),

$$(e_1)_{\text{obs}} = \frac{e_1}{\sqrt{\epsilon}}, \quad (e_2)_{\text{obs}} = \frac{e_2}{\sqrt{\epsilon}}, \quad (12.76)$$

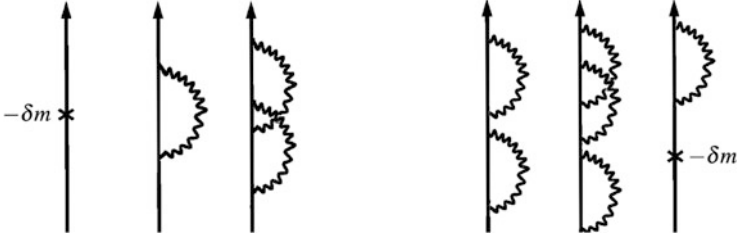
where  $(e)_{\text{obs}}$  is the experimental value. Thus, we must reinterpret  $(e)_{\text{obs}}$  as the given charge, and not the bare charge  $e$ . Since the permittivity of the vacuum  $\epsilon$  is absorbed into the definition of the electric charge, this is called a *renormalization of the electric charge*.

Renormalization can thus be additive, as for the mass, or multiplicative, as for the electric charge. Moreover, when we compute the self-energy  $\delta m$  or the permittivity  $\epsilon$  using QED, we find that they diverge.

What we have discussed above is the renormalization of specific quantities due to the self-energy and the permittivity. There exists a procedure called *renormalization of the field operators*, which is a slightly abstract renormalization condition. This has already been required in Chap. 11 and expressed in (11.145). There exists a condition whereby the wave function of a given body is not varied by introducing interactions, i.e., the normalization is unchanged. What this condition means will be explained in detail later.

### 12.3 Mass Renormalization

We call a Feynman diagram which starts with and ends in a one-electron state a *self-energy diagram*. Any diagram which cannot be separated into two disconnected diagrams by cutting a single electron line is called an *irreducible self-energy diagram* (see Fig. 12.2 left). Others are said to be *reducible* (see Fig. 12.2 right).



**Fig. 12.2** *Left:* Irreducible self-energy diagram. *Right:* Reducible self-energy diagram

Considering the  $S$ -matrix element for the case where an electron with the four-momentum  $p$  enters and electron in the same state comes out, we have

$$\langle p|S|p\rangle = \langle p|p\rangle + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \dots d^4x_n \langle p|T[\mathcal{H}_{\text{int}}(x_1) \dots \mathcal{H}_{\text{int}}(x_n)]|p\rangle_{\text{conn}}. \tag{12.77}$$

Here, in order to drop contributions from bubble diagrams, after the second term we keep only contributions from the connected parts in which bubble diagrams are omitted. The subscript ‘conn’ stands for dropping bubble diagrams which are not connected to any points in  $x_1, \dots, x_n$ . The condition (12.73) requires the terms after the second term on the right-hand side of (12.77) to vanish. Taking  $m_{\text{obs}}$  as the mass in the free part, the difference from  $m$ , viz.,  $\delta m$ , is included in the interaction part, so that in a first approximation where we consider only the renormalization of the mass, we can write

$$\mathcal{H}_{\text{int}} = -ie\bar{\psi}\gamma_{\mu}\psi A_{\mu} - \delta m\bar{\psi}\psi. \tag{12.78}$$

Computing the  $S$ -matrix element to order  $e^2$ , if we consider  $\delta m$  to be of order  $e^2$ , then from the Feynman–Dyson rule,

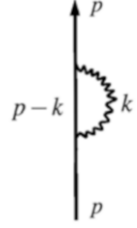
$$\langle p'|S^{(2)}|p\rangle = \int d^4x \langle p'|\bar{\psi}(x)|0\rangle \left[ i\delta m^{(2)} + e^2 \int d^4y \gamma_{\mu} S_{\text{F}}(y) \gamma_{\mu} D_{\text{F}}(y) e^{-ip\cdot y} \right] \langle 0|\psi(x)|p\rangle, \tag{12.79}$$

where we have taken into account only the first two irreducible self-energy diagrams shown above. This is because the third is of order  $e^4$ . We now introduce

$$\Sigma^{*(2)}(p) = ie^2 \int d^4y \gamma_{\mu} S_{\text{F}}(y) \gamma_{\mu} D_{\text{F}}(y) e^{-ip\cdot y}, \tag{12.80}$$

where the computation above is carried out in the Fermi–Feynman gauge corresponding to  $\alpha = 1$ . In general, the sum over contributions from all irreducible self-energy diagrams is called a *proper self-energy operator* or a *mass operator*.

**Fig. 12.3** Self-energy diagram in momentum space



Hence,

$$\langle p' | S^{(2)} | p \rangle = i \int d^4x \langle p' | \bar{\psi}(x) | 0 \rangle [\delta m^{(2)} - \Sigma^{*(2)}(p)] \langle 0 | \psi(x) | p \rangle. \quad (12.81)$$

The renormalization condition requires the expression above to vanish. Writing  $m_{\text{obs}}$  simply as  $m$ , we now have (Fig. 12.3)

$$\begin{aligned} \Sigma^{*(2)}(p) &= \frac{ie^2}{(2\pi)^4} \int d^4k \gamma_\mu \frac{i(p-k) \cdot \gamma - m}{(p-k)^2 + m^2 - i\epsilon} \gamma_\mu \frac{1}{k^2 - i\epsilon} \\ &= \frac{ie^2}{(2\pi)^4} \int d^4k \frac{-2i(p-k) \cdot \gamma - 4m}{(p-k)^2 + m^2 - i\epsilon} \frac{1}{k^2 - i\epsilon} \\ &= \frac{ie^2}{(2\pi)^4} \int d^4k \int dx \frac{-2i(p-k) \cdot \gamma - 4m}{(k^2 + xp^2 - 2xp \cdot k + xm^2 - i\epsilon)^2}, \end{aligned} \quad (12.82)$$

where we have used the fact that  $\gamma_\mu \gamma_\mu = 4$  and  $\gamma_\mu \gamma_\lambda \gamma_\mu = -2\gamma_\lambda$ . Making the change of variable

$$k \rightarrow k' = k - xp, \quad (12.83)$$

rewriting (12.82), and dropping odd-order terms in  $k'$ , we obtain

$$\Sigma^{*(2)}(p) = \frac{-ie^2}{(2\pi)^4} \int_0^1 dx \int d^4k' \frac{2i(1-x)p \cdot \gamma + 4m}{[k'^2 + x(1-x)p^2 + xm^2 - i\epsilon]^2}. \quad (12.84)$$

Since  $p^2 = -(ip \cdot \gamma)^2$ ,  $\Sigma^*(p)$  can be identified with a function of  $ip \cdot \gamma$ , and expanding this as a power series in  $(ip \cdot \gamma + m)$ , we have

$$\Sigma^{*(2)}(p) = A + B(ip \cdot \gamma + m) + C(p), \quad (12.85)$$

where  $C(p)$  is the sum over all terms higher than the second order in  $(ip \cdot \gamma + m)$ . Therefore,  $A$  can be obtained if we set  $ip \cdot \gamma = -m$  in  $\Sigma^{*(2)}(p)$ . This yields

$$A = \frac{-ie^2}{(2\pi)^4} \int_0^1 dx \int d^4k' \frac{1}{(k'^2 + x^2 m^2 - i\epsilon)^2}. \quad (12.86)$$

To carry out this  $k'$ -integral, we use

$$\frac{1}{(k'^2 + x^2 m^2 - i\epsilon)^2} = 2 \lim_{\Lambda \rightarrow \infty} \int_{m^2}^{\Lambda^2} \frac{x^2 dM^2}{(k'^2 + x^2 M^2 - i\epsilon)^3} . \quad (12.87)$$

Plugging this into (12.86) and carrying out the  $k'$ -integral, we obtain

$$\begin{aligned} A &= \frac{2\pi^2 e^2 m}{(2\pi)^4} \int_0^1 dx (1+x) \int_{m^2}^{\Lambda^2} \frac{dM^2}{M^2} \\ &= \frac{3}{2\pi} \alpha m \ln \frac{\Lambda}{m} , \quad \text{where } \alpha = \frac{e^2}{4\pi} . \end{aligned} \quad (12.88)$$

Going back to (12.81) and using the wave functions (3.191a), (3.191b), (3.192a) and (3.192b),

$$\langle p' | S^{(2)} | p \rangle = \frac{i(2\pi)^4}{V} \delta^4(p' - p) \bar{u}(p') [\delta m^{(2)} - \Sigma^{*(2)}(p)] u(p) . \quad (12.89)$$

Using Dirac's equation for  $u(p)$ , we can replace  $-p \cdot \gamma$  by  $-m$  in the equation above. Thus, the condition that (12.89) should vanish can be written as

$$\delta m^{(2)} = A . \quad (12.90)$$

This stands for the self-energy of the electron up to order  $e^2$ . It is clear from (12.88) that this diverges logarithmically in the limit  $\Lambda \rightarrow \infty$ . However, what we should emphasize here is that the renormalization condition (12.73) requires a choice of interaction part of the form (12.78).

## 12.4 Renormalization of Field Operators

We have used (12.90) to understand the meaning of  $A$  in the expansion (12.85), but what about  $B$ ? To answer this, we write down the condition for the normalization of the one-body wave function to remain unchanged when interactions are introduced:

$$\langle 0 | T[\psi(x) U(\infty, -\infty)]_{\text{conn}} | p \rangle = \langle 0 | \psi(x) | p \rangle . \quad (12.91)$$

The Feynman diagram corresponding to the left-hand side is shown in Fig. 12.4.

Here,  $\Sigma^*$  is the contribution from all the irreducible self-energy diagrams except for  $\delta m$ . It should be clear from the diagram above that the contribution from the



Fig. 12.4 Feynman diagram corresponding to the left-hand side of (12.91)

left-hand side of (12.91) is

$$\frac{i p \cdot \gamma + m}{i p \cdot \gamma + m + \Sigma^*(p) - \delta m} u(p) = \left\{ 1 - [\Sigma^*(p) - \delta m] \frac{1}{i p \cdot \gamma + m} \right. \\ \left. + [\Sigma^*(p) - \delta m] \frac{1}{i p \cdot \gamma + m} [\Sigma^*(p) - \delta m] \frac{1}{i p \cdot \gamma + m} + \dots \right\} u(p). \quad (12.92)$$

The discussion so far is based on the assumption that the interaction part is given by (12.78). Expanding  $\Sigma^*$  as suggested by (12.78), we can take  $i p \cdot \gamma = -m$  because  $u(p)$  appears in (12.92). This yields

$$\lim_{i p \cdot \gamma + m \rightarrow 0} \frac{i p \cdot \gamma + m}{i p \cdot \gamma + m + \Sigma^*(p) - \delta m} = \frac{1}{1 + B}. \quad (12.93)$$

This is not equal to unity unless  $B = 0$ , so (12.91) cannot be satisfied. Recall the origin of the condition that the normalization is unchanged. In fact, it originally arose from the asymptotic condition (11.126) for the operator in the Heisenberg picture. It was not assumed that the operator used in this case was the same as the original one. Hence, we may consider that the Heisenberg operator satisfying the asymptotic condition has a different normalization from the original Heisenberg operator appearing in the Lagrangian. We call operators satisfying the asymptotic condition *renormalized field operators*. They carry the subscript  $r$  and we assume the following multiplicative renormalization:

$$\psi(x) = Z_2^{1/2} \psi(x)_r, \quad \bar{\psi}(x) = Z_2^{1/2} \bar{\psi}(x)_r, \quad A_\mu(x) = Z_3^{1/2} A_\mu(x)_r. \quad (12.94)$$

The multiplicative renormalization has already appeared in the renormalization of the electric charge, and we shall see that this is in fact closely related to the renormalization of operators mentioned above.

Going back to the general gauge, let us express the whole Lagrangian in terms of renormalized operators. Since in this case the operators are originally the Heisenberg operators, we should use bold face for them, but there should be no



confusion if the usual type face is used:

$$\mathcal{L} = -Z_2 \bar{\psi}_r \left[ \gamma_\mu (\partial_\mu - ie Z_3^{1/2} A_{\mu r}) + m - \delta m \right] \psi_r - \frac{1}{4} Z_3 F_{\mu\nu r} F_{\mu\nu r} - \frac{1}{2\alpha} Z_3 (\partial_\mu A_{\mu r})^2 . \quad (12.95)$$

Although  $m - \delta m$  should be written as  $m_{\text{obs}} - \delta m$ , as already mentioned, for simplicity we have written  $m_{\text{obs}}$  as  $m$ . We decompose (12.95) into the free part and the interaction part as follows:

$$\mathcal{L}_f = -\bar{\psi}_r (\gamma_\mu \partial_\mu + m) \psi_r - \frac{1}{4} F_{\mu\nu r} F_{\mu\nu r} - \frac{1}{2\alpha_r} (\partial_\mu A_{\mu r})^2 , \quad (12.96)$$

$$\begin{aligned} \mathcal{L}_{\text{int}} = (1 - Z_2) \bar{\psi}_r (\gamma_\mu \partial_\mu + m) \psi_r + (1 - Z_3) \frac{1}{4} F_{\mu\nu r} F_{\mu\nu r} \\ + ie Z_2 Z_3^{1/2} A_{\mu r} \bar{\psi}_r \gamma_\mu \psi_r + Z_2 \delta m \bar{\psi}_r \psi_r , \end{aligned} \quad (12.97)$$

where  $\alpha_r$  is the renormalized gauge parameter. Hence, the gauge parameter changes under renormalization. The interaction picture corresponding to the partition above is called the *renormalized interaction picture*. The renormalized gauge parameter  $\alpha_r$  is defined by

$$\alpha_r = \alpha Z_3^{-1} . \quad (12.98)$$

Since we will only use the renormalized interaction picture in the following discussions, for simplicity we will drop the subscript  $r$ . In perturbation theory, several kinds of renormalization constant can be expanded as power series in  $e^2$ . We thus assume the expansions

$$\delta m = \delta m^{(2)} + \delta m^{(4)} + \dots , \quad (12.99)$$

$$Z_2 = 1 + Z_2^{(2)} + \dots , \quad (12.100)$$

$$Z_3 = 1 + Z_3^{(2)} + \dots . \quad (12.101)$$

In the interaction picture, the interaction includes derivatives of field operators, so we use Matthew's theorem to express the  $S$ -matrix in the form

$$S = 1 + \sum_{n=1}^{\infty} \frac{i^n}{n!} \int d^4 x_1 \dots d^4 x_n T^* [\mathcal{L}_{\text{int}}(x_1) \dots \mathcal{L}_{\text{int}}(x_n)]_{\text{conn}} . \quad (12.102)$$

We now repeat the discussion in the last section in this new interaction picture:

$$\langle p' | S^{(2)} | p \rangle = \int d^4 x \langle p' | \bar{\psi}(x) | 0 \rangle \mathcal{S}(p) \langle 0 | \psi(x) | p \rangle . \quad (12.103)$$

The new quantity  $\mathcal{S}(p)$  is given by the following equation, corresponding to  $i[\delta m^{(2)} - \Sigma^{*(2)}(p)]$  in (12.81):

$$\mathcal{S}(p) = iZ_2\delta m + i(1 - Z_2)(ip \cdot \gamma + m) - i\Sigma^*(p) . \quad (12.104)$$

In the lowest order approximation,  $\Sigma^*$  is given by

$$\Sigma^*(p) = ie^2 Z_2^2 Z_3 \int d^4 y \gamma_\mu S_F(y) \gamma_\mu D_F(y) e^{-ip \cdot y} . \quad (12.105)$$

Replacing  $Z_2$  and  $Z_3$  by 1, this coincides with the expression in the last section. Then in general, expanding as

$$\Sigma^*(p) = A + B(ip \cdot \gamma + m) + C(p) , \quad (12.106)$$

the expression (12.103) vanishes. Using the fact that the normalization of the one-particle wave function is unchanged, we have

$$Z_2\delta m = A , \quad 1 - Z_2 = B . \quad (12.107)$$

It turns out that  $\delta m$  and  $Z_2$  can be determined from this renormalization condition. Hence,

$$\mathcal{S}(p) = -iC(p) \equiv -i\Sigma_{\text{ren}}^*(p) . \quad (12.108)$$

This equation defines the renormalized mass operator  $\Sigma_{\text{ren}}^*$ . Although  $A^{(2)}$  and  $B^{(2)}$  diverge logarithmically,  $C^{(2)}(p)$  is finite.

## 12.5 Renormalized Propagators

In the renormalized interaction picture, the electron propagator is defined by

$$S_F(x - y) = \langle 0 | T[\psi(x) \bar{\psi}(y)] | 0 \rangle = \frac{-i}{(2\pi)^4} \int d^4 p e^{ip \cdot (x-y)} S_F(p) ,$$

with

$$S_F(p) = \frac{1}{ip \cdot \gamma + m} . \quad (12.109)$$

A propagator including higher order corrections can be expressed in the Heisenberg picture by

$$S'_F = \langle \mathbf{0} | T[\psi(x) \bar{\psi}(y)] | \mathbf{0} \rangle = \frac{\langle \mathbf{0} | T^*[\psi(x), \bar{\psi}(y), U(\infty, -\infty)] | \mathbf{0} \rangle}{\langle \mathbf{0} | U(\infty, -\infty) | \mathbf{0} \rangle}. \quad (12.110)$$

In the renormalized interaction picture,

$$U(\infty, -\infty) = 1 + \sum_{n=1}^{\infty} \frac{i^n}{n!} \int d^4x_1 \dots d^4x_n T^*[\mathcal{L}_{\text{int}}(x_1) \dots \mathcal{L}_{\text{int}}(x_n)]. \quad (12.111)$$

From the discussion about the Feynman diagram, the Fourier transform of  $S'_F$  is

$$\begin{aligned} S'_F(p) &= S_F(p) - S_F(p) \Sigma_{\text{ren}}^*(p) S_F(p) + \dots \\ &= S_F(p) - S_F(p) \Sigma_{\text{ren}}^*(p) S'_F(p). \end{aligned} \quad (12.112)$$

This is called *Dyson's equation*. Its solution is

$$S'_F(p) = S_F(p) [1 + \Sigma_{\text{ren}}^*(p) S_F(p)]^{-1} = [ip \cdot \gamma + m + \Sigma_{\text{ren}}^*(p)]^{-1}. \quad (12.113)$$

Since  $\Sigma_{\text{ren}}^*(p)$  is a sum over terms higher than second order in  $ip \cdot \gamma + m$ ,

$$\lim_{ip \cdot \gamma + m \rightarrow 0} (ip \cdot \gamma + m) S'_F(p) = 1. \quad (12.114)$$

This is an important property of the renormalized propagator. Then in the computation in Sect. 12.3, and in particular in (12.84), we make the change of variables from the Feynman parameter  $x$  to  $M$ , where

$$m^2 = (1-x)M^2, \quad (12.115)$$

whence  $S'_F(p)$  can be written to order  $e^2$  as

$$\begin{aligned} S'_F(p) &= S_F(p) - S_F(p) \Sigma_{\text{ren}}^{*(2)}(p) S_F(p) \\ &= \frac{1}{ip \cdot \gamma + m} + \frac{e^2}{16\pi^2} \int_m^\infty \frac{dM}{M^3(M^2 - m^2)} \left[ \frac{(M+m)^2(M^2 + m^2 - 4mM)}{ip \cdot \gamma + M - i\epsilon} \right. \\ &\quad \left. + \frac{(M-m)^2(M^2 + m^2 + 4mM)}{ip \cdot \gamma - M + i\epsilon} \right]. \end{aligned} \quad (12.116)$$

This integral diverges at  $M = m$ . In this case, we need to improve the approximation near the mass shell using some suitable method. We will return to this when discussing the renormalization group method in Chap. 20. However, what we call the ultraviolet divergence disappears completely.

Next, let us study the propagator of the electromagnetic field. Here we consider the renormalized Fermi–Feynman gauge  $\alpha_r = 1$ . In the interaction picture,

$$\delta_{\mu\nu} D_F(x - y) = \langle 0 | T[A_\mu(x) A_\nu(y)] | 0 \rangle = \frac{-i}{(2\pi)^4} \delta_{\mu\nu} \int d^4k e^{ik \cdot (x-y)} D_F(k) . \quad (12.117)$$

In the renormalized Heisenberg picture including higher order corrections, this is not proportional to  $\delta_{\mu\nu}$ :

$$D'_{F_{\mu\nu}}(x - y) = \langle 0 | T[A_\mu(x) A_\nu(y)] | 0 \rangle = \frac{\langle 0 | T^*[A_\mu(x) A_\nu(y) U(\infty, -\infty)] | 0 \rangle}{\langle 0 | U(\infty, -\infty) | 0 \rangle} . \quad (12.118)$$

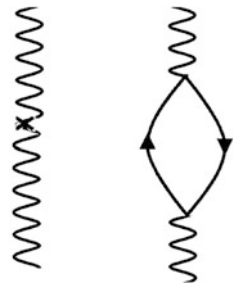
Let us compute corrections up to order  $e^2$ . In this case, the diagrams which should be taken into account are shown in Fig. 12.5.

The terms in the interaction Lagrangian needed for this calculation are

$$\frac{1}{4}(1 - Z_3)F_{\mu\nu}F_{\mu\nu} + ieZ_2Z_3^{1/2}A_\mu\bar{\psi}\gamma_\mu\psi .$$

The first term and iterations of the second term correspond to the first diagram and the second diagram in Fig. 12.5, respectively. However, to order  $e^2$ ,  $Z_2$  and  $Z_3$  in

**Fig. 12.5** Feynman diagrams for the electromagnetic field propagator up to order  $e^2$



the second term can be set equal to 1:

$$\begin{aligned}
 D'_{F_{\mu\nu}}(x-y) &= \delta_{\mu\nu} D_F(x-y) \\
 &\quad - e^2 \int d^4x' d^4x'' D_F(x-x') D_F(x''-y) \text{Tr}[\gamma_\mu S_F(x'-x'') \gamma_\nu S_F(x''-x')] \\
 &\quad + \frac{i}{2} (1-Z_3) \int d^4x' \left( \delta_{\mu\nu} \frac{\partial}{\partial x'_\rho} - \delta_{\mu\rho} \frac{\partial}{\partial x'_\sigma} \right) D_F(x-x'). \quad (12.119)
 \end{aligned}$$

Taking the Fourier transform of this equation yields

$$\begin{aligned}
 D'_{F_{\mu\nu}}(k) &= \frac{\delta_{\mu\nu}}{k^2 - i\epsilon} \\
 &\quad - \frac{ie^2}{(2\pi)^4} \frac{1}{(k^2 - i\epsilon)^2} \int d^4p \text{Tr} \left[ \gamma_\mu \frac{1}{ip \cdot \gamma + m - i\epsilon} \gamma_\nu \frac{1}{i(p-k) \cdot \gamma + m - i\epsilon} \right] \\
 &\quad + (1-Z_3) \frac{1}{k^2 - i\epsilon} \left( \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2 - i\epsilon} \right). \quad (12.120)
 \end{aligned}$$

The second term is the Fourier transform of the expression

$$\langle 0 | T^* [j_\mu(x), j_\nu(y)] | 0 \rangle.$$

And formally, this satisfies the condition

$$\frac{\partial}{\partial x_\mu} \langle 0 | T^* [j_\mu(x), j_\nu(y)] | 0 \rangle = 0. \quad (12.121)$$

In momentum space, this condition becomes

$$k_\mu \int d^4p \text{Tr}[\dots] = 0. \quad (12.122)$$

The integral in the equation above, denoted by  $f_{\mu\nu}(k)$ , has the general form

$$f_{\mu\nu}(k) = \delta_{\mu\nu} f(k^2) - k_\mu k_\nu g(k^2). \quad (12.123)$$

From the condition (12.122),

$$k_\mu f_{\mu\nu}(k) = k_\nu [f(k^2) - k^2 g(k^2)] = 0. \quad (12.124)$$

Thus,

$$f_{\mu\nu}(k) = (k^2 \delta_{\mu\nu} - k_\mu k_\nu) g(k^2). \quad (12.125)$$

Computing the trace in the integral in (12.122),

$$\text{Tr} = \frac{4[\delta_{\mu\nu}(p^2 - p \cdot k + m^2) - 2p_\mu p_\nu + p_\mu k_\nu + p_\nu k_\mu]}{(p^2 + m^2 - i\epsilon)[(p - k)^2 + m^2 - i\epsilon]}.$$

Therefore,

$$\begin{aligned} f_{\mu\nu}(k) &= 4 \int d^4 p \frac{\delta_{\mu\nu}(p^2 - p \cdot k + m^2) - 2p_\mu p_\nu + p_\mu k_\nu + p_\nu k_\mu}{(p^2 + m^2 - i\epsilon)[(p - k)^2 + m^2 - i\epsilon]} \\ &= 4 \int_0^1 dx \int d^4 p \frac{\delta_{\mu\nu}(p^2 - p \cdot k + m^2) - 2p_\mu p_\nu + p_\mu k_\nu + p_\nu k_\mu}{[p^2 + m^2 + x(k^2 - 2p \cdot k) - i\epsilon]^2}. \end{aligned}$$

We make the change of variables

$$p \rightarrow p' = p - xk. \quad (12.126)$$

Thus,

$$f_{\mu\nu}(k) = 4 \int_0^1 dx \int d^4 p' \frac{N}{[p'^2 + m^2 + x(1-x)k^2 - i\epsilon]^2}, \quad (12.127)$$

where

$$N = \delta_{\mu\nu} p'^2 - 2p'_\mu p'_\nu + \delta_{\mu\nu}[m^2 + x(1-x)k^2] - 2x(1-x)(\delta_{\mu\nu}k^2 - k_\mu k_\nu). \quad (12.128)$$

All terms except for the last produce terms proportional to  $\delta_{\mu\nu}$ . Hence, for this to coincide with (12.125), only the last term can survive. Thus,

$$f_{\mu\nu}(k) = -8(\delta_{\mu\nu}k^2 - k_\mu k_\nu) \int_0^1 dx x(1-x) \int \frac{d^4 p}{[p^2 + m^2 + x(1-x)k^2 - i\epsilon]^2}, \quad (12.129)$$

and we can write

$$\begin{aligned} \frac{1}{[p^2 + m^2 + x(1-x)k^2 - i\epsilon]^2} &= \frac{1}{(p^2 + m^2 - i\epsilon)^2} \\ &+ \left\{ \frac{1}{[p^2 + m^2 + x(1-x)k^2 - i\epsilon]^2} - \frac{1}{(p^2 + m^2 - i\epsilon)^2} \right\}. \end{aligned}$$

Therefore, the expression for  $D'_{F_{\mu\nu}}$  becomes

$$\begin{aligned}
 D'_{F_{\mu\nu}}(k) &= \frac{1}{k^2 - i\epsilon} + \frac{8ie^2}{(2\pi)^4} \frac{k^2 \delta_{\mu\nu} - k_\mu k_\nu}{(k^2 - i\epsilon)^2} \int_0^1 x(1-x) dx \\
 &\quad \times \int d^4 p \left\{ \frac{1}{[p^2 + m^2 + x(1-x)k^2 - i\epsilon]^2} - \frac{1}{(p^2 + m^2 - i\epsilon)^2} \right\} \\
 &\quad + \frac{k^2 \delta_{\mu\nu} - k_\mu k_\nu}{(k^2 - i\epsilon)^2} \left[ 1 - Z_3 + \frac{8ie^2}{(2\pi)^4} \int_0^1 x(1-x) dx \int \frac{d^4 p}{(p^2 + m^2 - i\epsilon)^2} \right].
 \end{aligned} \tag{12.130}$$

Then to determine  $Z_3$ , corresponding to (12.114) in the case of the electron, or the equation

$$S_F(p) \Sigma_{\text{ren}}^*(p) u(p) = 0, \tag{12.131}$$

we adopt the condition

$$D_F(k) \Pi_{\text{ren}}^*(k)_{\mu\nu} e_\nu = 0, \tag{12.132}$$

where  $e_\nu$  is the polarization vector of transverse photons and  $\Pi^*$  is called the *proper self-energy operator of the photon* or the *polarization operator*, which corresponds to  $\Sigma^*$  in the electron case. Thus, similarly to Dyson's equation (12.112) for  $\Sigma^*$  in the electron case,  $\Pi^*$  is defined by

$$\begin{aligned}
 D'_{F_{\mu\nu}} &= \delta_{\mu\nu} D_F(k) - D_F(k) \Pi_{\text{ren}}^*(k)_{\mu\nu} D_F(k) + \dots \\
 &= \delta_{\mu\nu} D_F(k) - D_F(k) \Pi_{\text{ren}}^*(k)_{\mu\lambda} D'_{F_{\lambda\nu}}(k).
 \end{aligned} \tag{12.133}$$

To order  $e^2$ ,

$$\begin{aligned}
 \Pi_{\text{ren}}^*(k)_{\mu\nu} &= -\frac{8ie^2}{(2\pi)^4} (k^2 \delta_{\mu\nu} - k_\mu k_\nu) \int_0^1 x(1-x) dx \\
 &\quad \times \int d^4 p \left[ \frac{1}{[p^2 + m^2 + x(1-x)k^2 - i\epsilon]^2} - \frac{1}{(p^2 + m^2 - i\epsilon)^2} \right] \\
 &\quad + \frac{8ie^2}{(2\pi)^4} (\delta_{\mu\nu} k^2 - k_\mu k_\nu) \left[ 1 - Z_3 + \frac{8ie^2}{(2\pi)^4} \int_0^1 x(1-x) dx \int \frac{d^4 p}{(p^2 + m^2 - i\epsilon)^2} \right].
 \end{aligned} \tag{12.134}$$

Then, in the one-photon state, taking into account  $k^2 = 0$  and  $k \cdot e = 0$ , the condition (12.132) requires the coefficient of  $\delta_{\mu\nu} k^2 - k_\mu k_\nu$  to vanish when  $k^2 = 0$ . Since the first term in (12.134) satisfies this condition, it implies that the second

term vanishes, i.e.,

$$Z_3 = 1 + \frac{8ie^2}{(2\pi)^4} \int_0^1 x(1-x)dx \int \frac{d^4p}{(p^2 + m^2 - i\epsilon)^2}. \tag{12.135}$$

Therefore, carrying out the Feynman integral,

$$\Pi_{\text{ren}}^*(k)_{\mu\nu} = \frac{e^2}{2\pi^2} (k^2\delta_{\mu\nu} - k_\mu k_\nu) \int_0^1 x(1-x) \ln \frac{m^2}{m^2 + x(1-x)k^2 - i\epsilon} dx. \tag{12.136}$$

In particular, when  $|k^2| \ll m^2$ ,

$$\Pi_{\text{ren}}^*(k)_{\mu\nu} \approx \frac{e^2}{60\pi^2} (\delta_{\mu\nu}k^2 - k_\mu k_\nu) \left( -\frac{k^2}{m^2} \right). \tag{12.137}$$

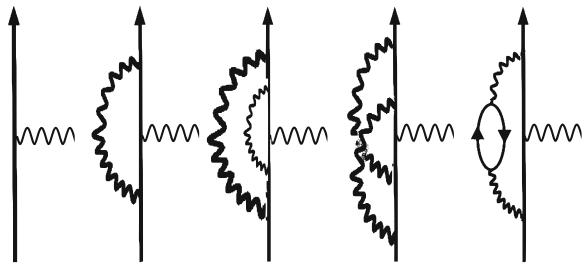
We see that, from the renormalization condition (12.132),  $Z_3$  is uniquely determined by (12.135), and it turns out that  $D'_F$  is divergenceless.

### 12.6 Renormalization of Vertex Functions

We have seen that the propagator becomes finite in the renormalized interaction picture. However, there is one thing that does not become finite without multiplicative renormalization, namely the vertex function. We now consider its renormalization. Up to now, we have investigated the proper self-energy diagrams, but these are all related to two-point functions or propagators. We now consider the corrections to the vertex function  $\gamma_\mu$  shown in Fig. 12.6.

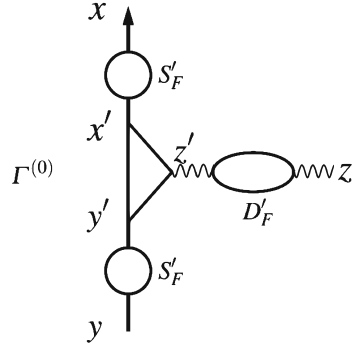
These are diagrams in which the propagator, including the self-energy diagram, is removed from the Feynman diagrams of the three-point function. They give the correction to the vertex operator. With this correction, the vertex function  $\gamma_\mu$  is replaced by the vertex function  $\Gamma_\mu^{(0)}$ . This is the same as replacing  $S_F$  and  $D_F$  by  $S'_F$  and  $D'_{F_{\mu\nu}}$ .

Fig. 12.6 Corrections to the vertex function





**Fig. 12.7** Feynman diagram for the complete three-point function



Thus, the Feynman diagram corresponding to the complete three-point function (Fig. 12.7) can be written like the one above, i.e., the general three-point function can be expressed as a product of three propagators and one vertex function. Corresponding to this diagram, we have the expression

$$\begin{aligned}
 \langle \mathbf{0} | T^*[\psi(x), \bar{\psi}(y), A_\nu(z)] | \mathbf{0} \rangle &= \frac{\langle \mathbf{0} | T^*[\psi(x), \bar{\psi}(y), A_\nu(z), U(\infty, -\infty)] | \mathbf{0} \rangle}{\langle \mathbf{0} | U(\infty, -\infty) | \mathbf{0} \rangle} \\
 &= -eZ_2Z_3^{1/2} \int d^4x' d^4y' d^4z' S'_F(x-x') \Gamma_\mu^{(0)}(x', y'; z') \\
 &\quad \times S'_F(y'-y) D'_{F\mu\nu}(z'-z) . \tag{12.138}
 \end{aligned}$$

Here, the Heisenberg operators and the interaction picture are renormalized. In the renormalized interaction picture, a propagator is automatically renormalized, but the vertex function is not. To the lowest order,

$$\Gamma_\mu^{(0)}(x, y, z) = \gamma_\mu \delta^4(x-z) \delta^4(y-z) . \tag{12.139}$$

Since in general  $\Gamma_\mu^{(0)}$  becomes a function of  $(x-z)$  and  $(y-z)$ , we define its Fourier representation by

$$\Gamma_\mu^{(0)}(x, y, z) = \frac{1}{(2\pi)^8} \int d^4p d^4q e^{ip \cdot (x-z) + iq \cdot (z-y)} \Gamma_\mu^{(0)}(p, q) . \tag{12.140}$$

Therefore, according to (12.139), to lowest order,

$$\Gamma_\mu^{(0)}(p, q) = \gamma_\mu . \tag{12.141}$$

We renormalize the vertex function in which higher-order corrections are included by the equations

$$\Gamma_\mu(p, q) = Z_1 \Gamma_\mu^{(0)}(p, q) , \quad (12.142)$$

$$\bar{u}(p) \Gamma_\mu(p, p) u(p) = \bar{u}(p) \gamma_\mu u(p) . \quad (12.143)$$

However, in (12.143), we have assumed that  $p$  is on the electron mass shell. Replacing  $\Gamma_\mu^{(0)}$  on the right-hand side of (12.138) by  $\Gamma$ , it turns out that the coefficient in front of the integral on the right-hand side of that equation is given by

$$e_{\text{obs}} = e Z_1^{-1} Z_2 Z_3^{1/2} . \quad (12.144)$$

In fact, this combination corresponds to the electric charge observed in experiments. This is manifested by several properties called *low-energy theorems*. Comparing (12.144) with (12.76), the permittivity of the vacuum is

$$\sqrt{\epsilon} = Z_1 Z_2^{-1} Z_3^{-1/2} . \quad (12.145)$$

Although the definition of the renormalized interaction picture is unchanged, it is more useful to rewrite  $e$  as  $e_{\text{obs}}$ . Then we write

$$\begin{aligned} e Z_2 Z_3^{1/2} &= e Z_1^{-1} Z_2 Z_3^{1/2} - (1 - Z_1) e Z_1^{-1} Z_2 Z_3^{1/2} \\ &= e_{\text{obs}} - (1 - Z_1) e_{\text{obs}} \\ &\equiv e_{\text{obs}} - \delta e . \end{aligned} \quad (12.146)$$

Thus, in the interaction part of the Lagrangian density, we rewrite as follows:

$$i e Z_2 Z_3^{1/2} \bar{\psi} \gamma_\mu \psi A_\mu = i (e_{\text{obs}} - \delta e) \bar{\psi} \gamma_\mu \psi A_\mu . \quad (12.147)$$

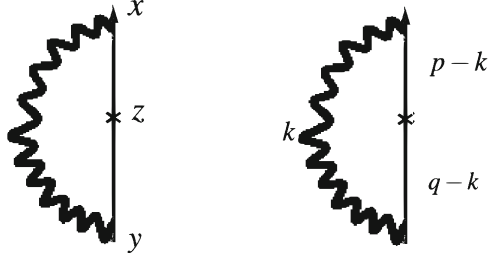
So from now on, we use powers of  $e_{\text{obs}}$  rather than powers of  $e$  in the perturbation theory. We have

$$\delta e = (1 - Z_1) e_{\text{obs}} = \mathcal{O}(e_{\text{obs}}^3) . \quad (12.148)$$

The relation between  $\Gamma_\mu$  and  $\Gamma_\mu^{(0)}$  is

$$(e_{\text{obs}} - \delta e) \Gamma_\mu^{(0)}(x, y, z) = e_{\text{obs}} \Gamma_\mu(x, y, z) . \quad (12.149)$$

**Fig. 12.8** Feynman diagram corresponding to the second term on the right-hand side of (12.150) in the  $x$ -space (left) and the  $p$ -space (right)



Computing the left-hand side up to order  $e_{\text{obs}}^3$ ,

$$e_{\text{obs}}\Gamma_{\mu}(x, y, z) = (e_{\text{obs}} - \delta e)\gamma_{\mu}\delta^4(x - z)\delta^4(y - z) + e_{\text{obs}}^2\gamma_{\lambda}S_{\text{F}}(x - z)\gamma_{\mu}S_{\text{F}}(z - y)\gamma_{\lambda}D_{\text{F}}(x - y) . \tag{12.150}$$

The Feynman diagram corresponding to the second term on the right-hand side is shown in the  $x$ -space and the  $p$ -space in Fig. 12.8.

Taking the Fourier transform of (12.150), we obtain

$$e_{\text{obs}}\Gamma_{\mu}(p, q) = (e_{\text{obs}} - \delta e)\gamma_{\mu} + \frac{(-i)^3}{(2\pi)^4}e_{\text{obs}}^3 \int d^4k \gamma_{\lambda} \frac{1}{i(p - k) \cdot \gamma + m - i\epsilon} \gamma_{\mu} \frac{1}{i(q - k) \cdot \gamma + m - i\epsilon} \gamma_{\lambda} \frac{1}{k^2 - i\epsilon} . \tag{12.151}$$

To determine  $\delta e$ , we use (12.143). Rationalizing the denominator of (12.151), the numerator becomes

$$N = \gamma_{\lambda}[i(p - k) \cdot \gamma + m]\gamma_{\mu}[i(q - k) \cdot \gamma + m]\gamma_{\lambda} . \tag{12.152}$$

Summing over  $\lambda$ , we can use the following formulas to calculate products of the  $\gamma$ -matrices:

- (1)  $\gamma_{\lambda}\gamma_a\gamma_{\lambda} = -2\gamma_a$  ,
- (2)  $\gamma_{\lambda}\gamma_a\gamma_b\gamma_{\lambda} = 4\delta_{ab}$  ,
- (3)  $\gamma_{\lambda}\gamma_a\gamma_b\gamma_c\gamma_{\lambda} = -2\gamma_c\gamma_b\gamma_a$  .

The numerator assumes the form

$$N = -2i(q - k) \cdot \gamma \gamma_{\mu} i(p - k) \cdot \gamma - 4im[(p - k)_{\mu} + (q - k)_{\mu}] - 2m^2\gamma_{\mu} . \tag{12.153}$$

Equation (12.151) then becomes

$$e_{\text{obs}}\Gamma_{\mu}(p, q) = (e_{\text{obs}} - \delta e)\gamma_{\mu} \quad (12.154)$$

$$+ \frac{(-ie_{\text{obs}})^3}{(2\pi)^4} \int d^4k \frac{N}{[(p-k)^2 + m^2 - i\epsilon][(p-k)^2 + m^2 - i\epsilon](k^2 - i\epsilon)} .$$

In order to carry out the  $k$ -integral, we use the following formula and change of variables:

$$\frac{1}{abc} = 2 \int_0^1 dx_1 \int_0^1 dx_2 \int_0^1 dx_3 \delta(1 - \Sigma x_i) \frac{1}{(x_1a + x_2b + x_3c)^3} , \quad (12.155)$$

$$P = \frac{1}{2}(p + q) , \quad \Delta = p - q . \quad (12.156)$$

Up to a factor of 2, the denominator of the integral in (12.154) can be written

$$\int_0^1 dx_1 \int_0^1 dx_2 \int_0^1 dx_3 \delta(1 - \sum x_i) f(x_1, x_2) = \int_D dx_1 dx_2 f(x_1, x_2) , \quad (12.157)$$

where

$$f(x_1, x_2) = \left[ \left( P^2 + \frac{\Delta^2}{4} + m^2 \right) (x_1 + x_2) + (P - k) \cdot \Delta (x_1 - x_2) \right. \quad (12.158)$$

$$\left. - 2k \cdot P (x_1 + x_2) + k^2 - i\epsilon \right]^{-3} .$$

The domain of integration  $D$  is shown in Fig. 12.9 (left). Changing the variables in (12.157) according to

$$x_1 + x_2 = u , \quad x_1 - x_2 = 2v , \quad (12.159)$$

the domain of integration for  $u$  and  $v$  is given by  $D'$ , depicted in Fig. 12.9 (right). We now change the integration variable from  $k$  to  $k'$ :

$$k' = k - uP - v\Delta . \quad (12.160)$$

The part corresponding to (12.155) becomes

$$2 \int_{D'} du dv \left[ k'^2 + \left( P^2 + \frac{\Delta^2}{4} + m^2 \right) u + 2P \cdot \Delta v - (uP + v\Delta)^2 - i\epsilon \right]^{-3} . \quad (12.161)$$

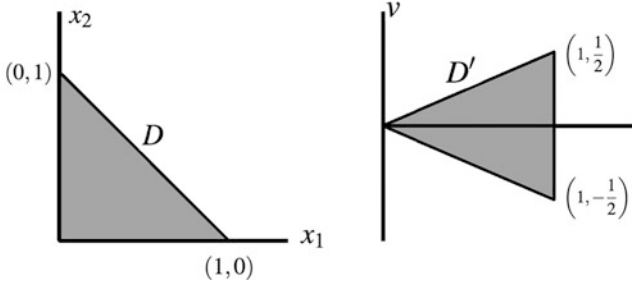


Fig. 12.9 Domains of integration for evaluating (12.154)

We now express the numerator in terms of  $k'$ . We can drop terms linear in  $k'$  because they give zero when we carry out the  $k'$ -integral. Since the integral is still complicated, we assume that  $p$  and  $q$  are on the mass shell, sandwich it between  $\bar{u}(p)$  and  $u(q)$ , and use the relations

$$\begin{aligned} \bar{u}(p)(i p \cdot \gamma + m) &= \bar{u}(p) \left( i P \cdot \gamma + \frac{1}{2} i \Delta \cdot \gamma + m \right) = 0, \\ (i q \cdot \gamma + m) u(p) &= \left( i P \cdot \gamma - \frac{1}{2} i \Delta \cdot \gamma + m \right) u(q) = 0. \end{aligned}$$

Moreover, when  $\bar{u}(p) A u(q) = \bar{u}(p) B u(q)$ , we write  $A \sim B$ . From the mass-shell condition,

$$p^2 + m^2 = q^2 + m^2 = P^2 + \frac{\Delta^2}{4} + m^2 = 0, \quad P \cdot \Delta = 0. \quad (12.162)$$

Therefore, the term in square brackets in (12.161) simplifies as follows:

$$k'^2 + m^2 u^2 + \Delta^2 \left( \frac{u^2}{4} - v^2 \right) - i \epsilon. \quad (12.163)$$

Furthermore, assuming that  $\Delta$  is small, we only keep terms linear in  $\Delta$  in the denominator and the numerator, and drop those in  $\Delta^2$ . Using the formula

$$2i P_\mu \sim -\sigma_{\mu\nu} \Delta_\nu - 2m \gamma_\mu, \quad (12.164)$$

we obtain

$$e_{\text{obs}} \Gamma_\mu(p, q) \sim (e_{\text{obs}} - \delta e) \gamma_\mu - 4 \frac{i e_{\text{obs}}^3}{(2\pi)^4} \int_{D'} du dv \int d^4 k' \frac{N'}{(k'^2 + m^2 u^2 - i \epsilon)^3} + \mathcal{O}(\Delta^2), \quad (12.165)$$

where

$$N' = -u(1-u)m\sigma_{\mu\nu}\Delta_\nu + m^2\gamma_\mu[(1-u)^2 - 4(1-u) + 1] + \frac{1}{2}\gamma_\mu k'^2. \quad (12.166)$$

In the limit  $\Delta \rightarrow 0$ , the right-hand side should be equal to  $e_{\text{obs}}\gamma_\mu$ , so

$$\delta e = -\frac{4ie_{\text{obs}}^3}{(2\pi)^4} \int_{D'} du dv \int d^4k' \frac{m^2(u^2 + 2u - 2) + k'^2/2}{(k'^2 + m^2u^2 - i\epsilon)^3}. \quad (12.167)$$

The right-hand side turns out to diverge logarithmically. Inserting (12.167) into (12.165),

$$\begin{aligned} e_{\text{obs}}\Gamma_\mu(p, q) &\sim e_{\text{obs}}\gamma_\mu + 4\frac{ie_{\text{obs}}^3}{(2\pi)^4} \int_{D'} du dv \int d^4k' \frac{u(1-u)m\sigma_{\mu\nu}\Delta_\nu}{(k'^2 + m^2u^2 - i\epsilon)^3} + \mathcal{O}(\Delta^2) \\ &= e_{\text{obs}}\gamma_\mu - \frac{\alpha}{2\pi} \frac{e_{\text{obs}}}{2m} \sigma_{\mu\nu}\Delta_\nu + \mathcal{O}(\Delta^2). \end{aligned} \quad (12.168)$$

Here,  $\alpha$  is, of course, equal to  $e_{\text{obs}}^2/4\pi$ . This equation is the one derived as the third-order perturbation. The effective Hamiltonian which yields the same result up to the first order in  $\Delta$  is

$$\mathcal{H}_{\text{eff}} = -\frac{\alpha}{2\pi} \frac{e_{\text{obs}}}{4m} \bar{\psi} \sigma_{\mu\nu} \psi F_{\mu\nu}. \quad (12.169)$$

This term, which is gauge invariant, is called the *Pauli term*. When there is only a magnetic field, applying the non-relativistic approximation, we have

$$\mathcal{H}_{\text{eff}} = -\frac{\alpha}{2\pi} \frac{e_{\text{obs}}}{2m} \boldsymbol{\sigma} \cdot \mathbf{H}. \quad (12.170)$$

This tells us that the electron acquires a supplementary magnetic moment in addition to  $e_{\text{obs}}/2m$  in the Dirac theory. This increase is called the *anomalous magnetic moment*. As a result, the magnetic moment of the electron up to this order is given by

$$\frac{e_{\text{obs}}}{2m} \left( 1 + \frac{\alpha}{2\pi} \right). \quad (12.171)$$

This result was obtained by Schwinger and by Tomonaga et al. [116, 117], and it matches experimental values well. It is considered to be a great achievement of renormalization theory, which thus succeeded in explaining what we call the *Lamb shift* in the hydrogen atom, and has provided a foundation for the development of field theory.

So far, we have introduced  $Z_1$ ,  $Z_2$ , and  $Z_3$  as multiplicative renormalization constants. From (12.148) and (12.167),  $Z_1$  is given by

$$Z_1 = 1 + \frac{4ie_{\text{obs}}^2}{(2\pi)^4} \int_{D'} du dv \int d^4k' \frac{m^2(u^2 + 2u - 2) + k'^2/2}{(k'^2 + m^2u^2 - i\epsilon)^3}. \quad (12.172)$$

Although we have not provided an explicit computation of  $Z_2$ , it is given by (12.107), and computing to this order, it coincides with  $Z_1$ . We will discuss this equality in the next section.

## 12.7 Ward–Takahashi Identity

We now investigate in detail the relation (12.115) between the bare electric charge  $e$  and the observed electric charge  $e_{\text{obs}}$ . So far, we have only considered the electron. Let us consider the case where there is a wide variety of charged particles  $a, b, \dots$ . Then,

$$(e_a)_{\text{obs}} = Z_{1a}^{-1} Z_{2a} Z_3^{1/2} e_a, \quad (e_b)_{\text{obs}} = Z_{1b}^{-1} Z_{2b} Z_3^{1/2} e_b, \dots \quad (12.173)$$

We consider the reaction

$$a + b \longrightarrow c + d. \quad (12.174)$$

In this case, it is the bare electric charge that is conserved by Noether's theorem:

$$e_a + e_b = e_c + e_d. \quad (12.175)$$

However, we know experimentally that charge conservation holds for the renormalized electric charges:

$$(e_a)_{\text{obs}} + (e_b)_{\text{obs}} = (e_c)_{\text{obs}} + (e_d)_{\text{obs}}. \quad (12.176)$$

In order for these two conservation laws to hold simultaneously,  $Z_1^{-1} Z_2$  cannot depend on the type of charged particle, i.e.,

$$Z_{1a}^{-1} Z_{2a} = Z_{1b}^{-1} Z_{2b} = Z_{1c}^{-1} Z_{2c} = Z_{1d}^{-1} Z_{2d}. \quad (12.177)$$

Ward discovered that these equalities can be replaced by the following, which imply them [118]:

$$Z_{1a} = Z_{2a}, \dots \quad (12.178)$$

The equation  $Z_1 = Z_2$  is referred to as the *Ward identity*. We shall now give its proof.

We take  $\psi^{(0)}$  and  $\bar{\psi}^{(0)}$  to be non-renormalized operators of the electric field. Setting

$$\mathbf{J}_\mu = i\bar{\psi}^{(0)}\gamma_\mu\psi^{(0)} = iZ_2\bar{\psi}\gamma_\mu\psi, \quad (12.179)$$

and using the fact that non-renormalized Heisenberg operators satisfy the canonical commutation relations for  $x_0 = y_0$ , we have

$$[\mathbf{J}_0(x), \psi^{(0)}(y)] = -\psi^{(0)}(y)\delta^3(x-y), \quad [\mathbf{J}_0(x), \bar{\psi}^{(0)}(y)] = \bar{\psi}^{(0)}(y)\delta^3(x-y). \quad (12.180)$$

Thus, if  $\mathbf{J}_\mu$  is included in the T-product, then using  $\partial_\mu \mathbf{J}_\mu = 0$ , we obtain

$$\begin{aligned} \partial_\mu T[\mathbf{J}_\mu(z), \psi^{(0)}(x), \bar{\psi}^{(0)}(y)] &= T[[\mathbf{J}_0(z), \psi^{(0)}(x)]\delta(z_0 - x_0), \bar{\psi}^{(0)}(y)] \\ &\quad + T[\psi^{(0)}(x), [\mathbf{J}_0(z), \bar{\psi}^{(0)}(y)]\delta(z_0 - y_0)] \\ &= [\delta^4(z-y) - \delta^4(z-x)]T[\psi^{(0)}(x), \bar{\psi}^{(0)}(y)]. \end{aligned} \quad (12.181)$$

As we have seen before, at the point where the order of two time variables in the T-product are switched, a delta function in time shows up. Later, we will use a generalization of (12.181).

The renormalized interaction picture is defined by (12.96) and (12.97). On the other hand, expressing (12.181) in terms of renormalized operators,

$$\begin{aligned} \partial_\mu T[\mathbf{J}_\mu(z), \psi(x), \bar{\psi}(y)] &= [\delta^4(z-y) - \delta^4(z-x)]T[\psi(x), \bar{\psi}(y)] \quad (12.182) \\ &\quad - \left[ \psi(z)\frac{\delta}{\delta\psi(z)} - \bar{\psi}(z)\frac{\delta}{\delta\bar{\psi}(z)} \right] T[\psi(x), \bar{\psi}(y)]. \end{aligned}$$

We will use this formula later by generalizing it to a certain extent.

In order to apply the reduction formula for the electromagnetic field, we introduce the differential operator

$$D_{\mu\nu}(\partial) = \delta_{\mu\nu}\square - \partial_\mu\partial_\nu + \frac{1}{\alpha}\partial_\mu\partial_\nu. \quad (12.183)$$

Applying the reduction formula to Green's function in the renormalized interaction picture, we compute the quantity

$$\begin{aligned} D_{\mu\nu}(\partial)T^*[A_\nu(x)\dots, U(\infty, -\infty)] &= T^*\left[i\frac{\delta}{\delta A_\mu(x)}\dots, U(\infty, -\infty)\right] \quad (12.184) \\ &\quad - T^*[\dots[\mathcal{L}_{\text{int}}(x)]_{A_\mu}, U(\infty, -\infty)], \end{aligned}$$



where an explicit expression for the Euler derivative is

$$[\mathcal{L}_{\text{int}}(x)]_{A_\mu} = eZ_3^{1/2} J_\mu - (1 - Z_3)(\delta_{\mu\sigma} \square - \partial_\mu \partial_\sigma) A_\sigma(x). \quad (12.185)$$

Inserting this into (12.184) and taking the derivative in the second term of (12.185) outside  $T^*$ , we have

$$\begin{aligned} D_{\mu\nu}(\partial)T^*[A_\nu(x) \dots, U(\infty, -\infty)] &= T^*\left[i\frac{\delta}{\delta A_\mu(x)} \dots, U(\infty, -\infty)\right] \\ &\quad - eZ_3^{1/2}T^*[J_\mu(x), \dots, U(\infty, -\infty)] \\ &\quad + (1 - Z_3)(\delta_{\mu\sigma} \Delta - \partial_\mu \partial_\sigma)T^*[A_\sigma(x), \dots, U(\infty, -\infty)]. \end{aligned} \quad (12.186)$$

Here we use the relation

$$eZ_3^{1/2} = e(Z_1^{-1}Z_2Z_3^{1/2})Z_1Z_2^{-1} = e_{\text{obs}}Z_1Z_2^{-1}. \quad (12.187)$$

Next, we differentiate (12.186) with respect to  $x_\mu$  and use (12.182). Since the last term vanishes, we obtain

$$\begin{aligned} \partial_\mu D_{\mu\nu}(\partial)T^*[A_\nu(x), \dots, U(\infty, -\infty)] &= \partial_\mu T^*\left[i\frac{\delta}{\delta A_\mu(x)} \dots, U(\infty, -\infty)\right] \\ &\quad + e_{\text{obs}}Z_1Z_2^{-1}T^*\left[\left(\psi(x)\frac{\delta}{\delta\psi(x)} - \bar{\psi}(x)\frac{\delta}{\delta\bar{\psi}(x)}\right) \dots, U(\infty, -\infty)\right]. \end{aligned} \quad (12.188)$$

Taking the vacuum expectation value of the above equation and using the Gell-Mann–Low formula, the relation between the Green's functions involving the renormalized Heisenberg operators is

$$\begin{aligned} \partial_\mu D_{\mu\nu}(\partial)\langle\mathbf{0}|T^*[A_\nu(x) \dots]|\mathbf{0}\rangle &= \partial_\mu\langle\mathbf{0}|i\frac{\delta}{\delta A_\mu(x)}T^*[\dots]|\mathbf{0}\rangle \\ &\quad + e_{\text{obs}}Z_1Z_2^{-1}\langle\mathbf{0}|T^*\left[\left(\psi(x)\frac{\delta}{\delta\psi(x)} - \bar{\psi}(x)\frac{\delta}{\delta\bar{\psi}(x)}\right) \dots\right]|\mathbf{0}\rangle. \end{aligned} \quad (12.189)$$

In the equation above, dots stand for a suitable product of Heisenberg operators.

One feature of this equation is that only renormalized operators show up. Since  $Z_1Z_2^{-1} = 1$ , as will be shown later, we can now obtain the *Ward–Takahashi identity* [119]:

$$\begin{aligned} \partial_\mu D_{\mu\nu}(\partial)\langle\mathbf{0}|T^*[A_\nu(x) \dots]|\mathbf{0}\rangle &= \partial_\mu\langle\mathbf{0}|i\frac{\delta}{\delta A_\mu(x)}T^*[\dots]|\mathbf{0}\rangle \\ &\quad + e_{\text{obs}}\langle\mathbf{0}|T^*\left[\left(\psi(x)\frac{\delta}{\delta\psi(x)} - \bar{\psi}(x)\frac{\delta}{\delta\bar{\psi}(x)}\right) \dots\right]|\mathbf{0}\rangle. \end{aligned} \quad (12.190)$$

To understand the relevance of (12.189), we first insert  $A_\sigma(y)$  in place of the dots to obtain

$$\partial_\mu D_{\mu\nu}(\partial_x) \langle \mathbf{0} | T^* [A_\nu(x), A_\sigma(y)] | \mathbf{0} \rangle = i \partial_\sigma \delta^4(x-y) . \quad (12.191)$$

Expressing this in momentum space and taking into account the equation

$$\partial_\mu D_{\mu\nu}(\partial) = \frac{1}{\alpha} \square \partial_\nu , \quad (12.192)$$

we find

$$\frac{1}{\alpha} k^2 k_\nu D'_{F\nu\sigma}(k) = k_\sigma . \quad (12.193)$$

We now write Dyson's equation for  $D'_F$  in an arbitrary gauge, viz.,

$$D'_{F\nu\sigma}(k) = D_{F\nu\sigma}(k) - D_{F\nu\lambda}(k) \Pi_{\lambda\mu}^*(k) D'_{F\mu\sigma}(k) , \quad (12.194)$$

where

$$D_{F\mu\sigma}(k) = \left( \delta_{\nu\sigma} - \frac{k_\nu k_\sigma}{k^2 - i\epsilon} \right) \frac{1}{k^2 - i\epsilon} + \alpha \frac{k_\nu k_\sigma}{(k^2 - i\epsilon)^2} \quad (12.195)$$

and

$$\Pi_{\lambda\mu}^*(k) = \Pi_{\text{ren}}^*(k)_{\lambda\mu} = (k_\lambda k_\mu - \delta_{\lambda\mu} k^2) \Pi^*(-k^2) . \quad (12.196)$$

From this, we obtain the solution of (12.194) in the form

$$D'_{F\nu\sigma}(k) = \left( \delta_{\nu\sigma} - \frac{k_\nu k_\sigma}{k^2 - i\epsilon} \right) \frac{1}{k^2 - i\epsilon} \frac{1}{1 - \Pi^*(-k^2)} + \alpha \frac{k_\nu k_\sigma}{(k^2 - i\epsilon)^2} . \quad (12.197)$$

This certainly satisfies (12.193).

Another example is

$$\begin{aligned} \partial_\mu D_{\mu\nu}(\partial_z) \langle \mathbf{0} | T^* [\psi(x), \bar{\psi}(y), A_\nu(z)] | \mathbf{0} \rangle \\ = -e_{\text{obs}} Z_1 Z_2^{-1} [\delta^4(y-z) - \delta^4(x-z)] \langle \mathbf{0} | T^* [\psi(x), \bar{\psi}(y)] | \mathbf{0} \rangle . \end{aligned} \quad (12.198)$$

We combine (12.191) and this equation with the following equation:

$$\begin{aligned} \langle \mathbf{0} | T^* [\psi(x), \bar{\psi}(y), A_\nu(z)] | \mathbf{0} \rangle \\ = -e_{\text{obs}} \int d^4 x' d^4 y' d^4 z' S'_F(x-x') \Gamma_\mu(x', y'; z') S'_F(y'-y) D'_{F\nu\sigma}(z'-z) . \end{aligned} \quad (12.199)$$

Inserting the right-hand side of (12.199) into the left-hand side of (12.198) and using (12.191) to cancel  $D'_F$ , we obtain the following equation in momentum space:

$$-i(p-q)_\nu S'_F(p) \Gamma_\nu(p, q) S'_F(q) = Z_1 Z_2^{-1} [S'_F(p) - S'_F(q)] . \quad (12.200)$$

Alternatively, differentiating both sides with respect to  $p_\mu$  and setting  $q = p$ ,

$$\Gamma_\mu(p, p) = -i Z_1 Z_2^{-1} \frac{\partial}{\partial p_\mu} S'_F(p)^{-1} . \quad (12.201)$$

We sandwiching this between  $\bar{u}(p)$  and  $u(p)$  and use the relations

$$\begin{aligned} S'_F(p)^{-1} &= ip \cdot \gamma + m + \mathcal{O}((ip \cdot \gamma + m)^2) , \\ \bar{u}(p)(ip \cdot \gamma + m) &= (ip \cdot \gamma + m)u(p) = 0 . \end{aligned}$$

This leads finally to

$$\bar{u}(p) \gamma_\mu u(p) = Z_1 Z_2^{-1} \bar{u}(p) \gamma_\mu u(p) , \quad (12.202)$$

which yields Ward's identity

$$Z_1 = Z_2 . \quad (12.203)$$

Thus, in the Ward–Takahashi identity (12.189), we can replace  $Z_1 Z_2^{-1}$  by 1 to obtain the Ward–Takahashi identity given by (12.190). In general, an equation obtained by subtracting a divergence of the Green's function is called a Ward–Takahashi identity.

Equations (12.201) and (12.203) were found by Ward [118], while (12.200) was derived by Takahashi [119]. Equation (12.190) for the Green's function was proved by the author in the Fermi gauge [120], but it was eventually shown that it can be extended to an arbitrary gauge.

## 12.8 Integral Representation of the Propagator

The propagator with higher-order corrections can be expressed by a certain kind of integral representation as the superposition of free-field propagators with different masses. This was established by Umezawa et al. in the early 1950s [121–123].

### 12.8.1 Integral Representation

We consider a neutral scalar field and make the following general assumptions:

1. There exists a four-vector operator  $P_\mu$ , which stands for the energy–momentum, satisfying

$$[P_\mu, P_\nu] = 0, \quad (12.204)$$

$$[\varphi(x), P_\mu] = \frac{1}{i} \partial_\mu \varphi(x). \quad (12.205)$$

2. There exists a set  $\{\Phi_{k,\alpha}\}$  of eigenstates of  $P_\mu$  which form a complete system. Here,  $k_\mu$  is an eigenvalue of  $P_\mu$  and  $\alpha$  is another quantity specifying the states:

$$P_\mu \Phi_{k,\alpha} = k_\mu \Phi_{k,\alpha}. \quad (12.206)$$

Of course, as  $\{\Phi\}$ , we can choose either  $\{\Phi^{\text{in}}\}$  or  $\{\Phi^{\text{out}}\}$ . In what follows, we assume the existence of the vacuum and write it as  $|\mathbf{0}\rangle$ . Affixing a prime to the invariant function in the presence of interactions,

$$\begin{aligned} \langle \mathbf{0} | \varphi(x) \varphi(y) | \mathbf{0} \rangle &= i \Delta^{(+)\prime}(x-y), \\ \langle \mathbf{0} | [\varphi(x), \varphi(y)] | \mathbf{0} \rangle &= i \Delta^\prime(x-y), \\ \langle \mathbf{0} | T[\varphi(x) \varphi(y)] | \mathbf{0} \rangle &= \Delta^\prime_F(x-y). \end{aligned} \quad (12.207)$$

Then, to carry the Fourier expansion, we introduce the matrix element

$$\langle \mathbf{0} | \varphi(x) | \mathbf{k}, \alpha \rangle = a_{k,\alpha} e^{ik \cdot x}. \quad (12.208)$$

Hence,

$$\begin{aligned} \langle \mathbf{0} | \varphi(x) \varphi(y) | \mathbf{0} \rangle &= \sum_{k,\alpha} \langle \mathbf{0} | \varphi(x) | \mathbf{k}, \alpha \rangle \langle \mathbf{k}, \alpha | \varphi(y) | \mathbf{0} \rangle \\ &= \sum_{k,\alpha} |a_{k,\alpha}|^2 e^{ik \cdot (x-y)}. \end{aligned} \quad (12.209)$$

We now introduce the Lorentz invariant function

$$\rho(-p^2) = (2\pi)^3 \sum_{k,\alpha} |a_{k,\alpha}|^2 \delta^4(p-k) \geq 0. \quad (12.210)$$

Inserting this into (12.209), for the time-like four-momentum  $p$ , we have

$$\begin{aligned}
 i\Delta^{(+)\prime}(x-y) &= \frac{1}{(2\pi)^3} \int d^4 p \theta(p_0) \rho(-p^2) e^{ip \cdot (x-y)} \quad (-p^2 \geq 0) \\
 &= \int_0^\infty d\kappa^2 \rho(\kappa^2) \frac{1}{(2\pi)^3} \int d^4 p \theta(p_0) \delta(p^2 + \kappa^2) e^{ip \cdot (x-y)} \\
 &= \int_0^\infty d\kappa^2 \rho(\kappa^2) i\Delta^{(+)}(x-y; \kappa^2) .
 \end{aligned} \tag{12.211}$$

Generalizing this to an arbitrary function, we obtain

$$\Delta^{(+)\prime}(x) = \int_0^\infty d\kappa^2 \rho(\kappa^2) \Delta^{(+)}(x; \kappa^2) , \tag{12.212}$$

so in the presence of interactions, the invariant two-point function can be expressed by the integral representation as a superposition of invariant two-point functions for free fields with different masses.

The spectral function  $\rho(\kappa^2)$  for a free field with the mass  $m$  is

$$\rho(\kappa^2) = \delta(\kappa^2 - m^2) . \tag{12.213}$$

If interactions are introduced, we can decompose into contributions from the one-particle intermediate states and the multi-particle intermediate states:

$$\rho(\kappa^2) = c\delta(\kappa^2 - m^2) + \sigma(\kappa^2) . \tag{12.214}$$

We note, however, that  $\sigma$  vanishes below the lowest invariant mass  $2m$  in the two-particle system, i.e.,

$$\sigma(\kappa^2) = \theta(\kappa^2 - 4m^2) \sigma(\kappa^2) . \tag{12.215}$$

If we apply the renormalization condition, namely that the normalization of the wave function becomes the same as that of the free field in the one-particle state, then  $c$  in (12.214) becomes unity, i.e.,

$$\rho(\kappa^2) = \delta(\kappa^2 - m^2) + \sigma(\kappa^2) . \tag{12.216}$$

We now introduce the Fourier representation of the propagator:

$$\Delta'_F(x) = \frac{-i}{(2\pi)^4} \int d^4 k e^{ik \cdot x} \Delta'_F(-k^2) . \tag{12.217}$$

Hence,

$$\Delta'_F(-k^2) = \frac{1}{k^2 + m^2 - i\epsilon} + \int_{4m^2}^{\infty} dk^2 \frac{\sigma(\kappa^2)}{k^2 + \kappa^2 - i\epsilon} . \quad (12.218)$$

We introduce the renormalization constant  $Z_\varphi$ , noting that, for the electron field  $Z_\psi = Z_2$  and for the electromagnetic field  $Z_A = Z_3$  :

$$\varphi^{(0)}(x) = Z_\varphi^{1/2} \varphi(x) , \quad (12.219)$$

where the superscript (0) denotes unrenormalized quantities. Therefore,

$$\langle \mathbf{0} | [\varphi^{(0)}(x), \varphi^{(0)}(y)] | \mathbf{0} \rangle = Z_\varphi \int dk^2 \rho(\kappa^2) i\Delta(x - y; \kappa^2) . \quad (12.220)$$

Differentiating both sides of this equation with respect to  $x_0$  and then setting  $x_0 = y_0$ , the left-hand side becomes  $-i\delta^3(x - y)$  from the canonical commutation relation. In addition, on the right-hand side, using (4.13),

$$\left. \frac{\partial}{\partial x_0} \Delta(x - y; \kappa^2) \right|_{x_0=y_0} = -\delta^3(x - y) , \quad (12.221)$$

whence

$$-i\delta^3(x - y) = -i\delta^3(x - y) Z_\varphi \int dk^2 \rho(\kappa^2) . \quad (12.222)$$

The integral representation of the renormalization constant can now be obtained immediately as

$$Z_\varphi^{-1} = \int dk^2 \rho(\kappa^2) = 1 + \int dk^2 \sigma(\kappa^2) \geq 1 . \quad (12.223)$$

This inequality originates from the positive-definite metric assumed for the scalar field. Thus,

$$1 \geq Z_\varphi \geq 0 . \quad (12.224)$$

### 12.8.2 Self-Energy

As shown above, the renormalization constant is written in terms of the integral representation via the spectral function. But what about the self-energy? To discuss

this problem, we consider the following simple model:

$$\mathcal{L} = -(\partial_\lambda \Phi^\dagger \partial_\lambda \Phi + M_0^2 \Phi^\dagger \Phi) - \frac{1}{2}[(\partial_\lambda \varphi)^2 + m_0^2 \varphi^2] - g_0 \Phi^\dagger \Phi \varphi . \quad (12.225)$$

Here  $\varphi$  stands for a neutral scalar field,  $\Phi$  and  $\Phi^\dagger$  for charged scalars, and  $m_0$  and  $M_0$  for the bare masses associated with each field, which are related to the observed masses  $m$  and  $M$  via

$$M^2 = M_0^2 + \delta M^2 , \quad m^2 = m_0^2 + \delta m^2 . \quad (12.226)$$

We introduce multiplicative renormalizations via the relations

$$\Phi^{(0)} = Z_2^{1/2} \Phi , \quad \varphi^{(0)} = Z_3^{1/2} \varphi , \quad g_0 = Z_1 Z_2^{-1} Z_3^{-1/2} g . \quad (12.227)$$

The renormalized field equations are

$$(\square - m^2)\varphi = Z_1 Z_3^{-1} g \Phi^\dagger \Phi - \delta m^2 \varphi , \quad (12.228)$$

$$(\square - M^2)\Phi = Z_1 Z_2^{-1} g \Phi \varphi - \delta M^2 \Phi . \quad (12.229)$$

Using (12.228), we can derive an integral representation of  $\delta m^2$ . Likewise, we can use (12.229) to derive an integral representation of  $\delta M^2$ . Expressing (12.228) in terms of  $m_0$ ,

$$(\square_x - m_0^2)\langle \mathbf{0} | [\varphi(x), \varphi(y)] | \mathbf{0} \rangle = Z_1 Z_3^{-1} g \langle \mathbf{0} | [\Phi^\dagger(x)\Phi(x), \varphi(y)] | \mathbf{0} \rangle . \quad (12.230)$$

Differentiating this equation with respect to  $y_0$  and setting  $y_0 = x_0$ , the right-hand side becomes

$$\langle \mathbf{0} | [\Phi^\dagger(x)\Phi(x), \dot{\varphi}(y)] | \mathbf{0} \rangle = 0 \quad (x_0 = y_0) . \quad (12.231)$$

It thus turns out that the left-hand side also vanishes. The left-hand side should still vanish if we carry out the operation discussed above in the integral representation:

$$(\square^2 - m_0^2)\langle \mathbf{0} | [\varphi(x), \varphi(y)] | \mathbf{0} \rangle = i \int d\kappa^2 \rho_\varphi(\kappa^2) (\kappa^2 - m_0^2) \Delta(x - y; \kappa^2) .$$

Differentiating this with respect to  $y_0$  and setting  $y_0 = x_0 = 0$ , Eq. (12.221) implies

$$i\delta^3(x - y) \int d\kappa^2 (\kappa^2 - m_0^2) \rho_\varphi(\kappa^2) = 0 ,$$

and hence,

$$\int d\kappa^2 (\kappa^2 - m^2) \rho_\varphi(\kappa^2) = -\delta m^2 \int d\kappa^2 \rho_\varphi(\kappa^2) = -Z_3^{-1} \delta m^2 . \quad (12.232)$$

Thus,

$$\delta m^2 = -Z_3 \int d\kappa^2 (\kappa^2 - m^2) \rho_\varphi(\kappa^2) = -Z_3 \int d\kappa^2 (\kappa^2 - m^2) \sigma(\kappa^2) < 0 . \quad (12.233)$$

This equation implies that  $\delta m^2$  is always negative. This is, of course, a predictable result. If we assume that, in (12.226),  $m^2$  is positive and finite and  $m_0^2$  is positive and infinite, then  $\delta m^2$  can only be negative and infinite.

### 12.8.3 Integral Representation of the Electromagnetic Field Propagator

We now apply the previous discussion for the scalar field to the electromagnetic field. We first introduce a propagator for the electromagnetic field:

$$\langle \mathbf{0} | T^* [A_\mu(x) A_\nu(y)] | \mathbf{0} \rangle = \frac{-i}{(2\pi)^4} \int d^4k e^{ik \cdot (x-y)} D'_{F_{\mu\nu}}(k) . \quad (12.234)$$

For the free electromagnetic field,  $D_{F_{\mu\nu}}$  is given by (12.195). Furthermore, it is clear from (12.197) that the effects of the interaction appear only in the transverse part of the wave. Recalling that

$$D_{\mu\nu}(\partial) A_\nu = -j_\mu , \quad (12.235)$$

$$j_\mu = -ie_{\text{obs}} Z_2 \bar{\psi} \gamma_\mu \psi - (1 - Z_3) (\delta_{\mu\sigma} \square - \partial_\mu \partial_\sigma) A_\sigma , \quad (12.236)$$

we introduce the current  $j_\mu$  and define the function

$$\langle \mathbf{0} | T^* [j_\mu(x) j_\nu(y)] | \mathbf{0} \rangle = \frac{-i}{(2\pi)^4} \int d^4k e^{ik \cdot (x-y)} \Pi_{\mu\nu}(k) . \quad (12.237)$$

Therefore,

$$k^4 \text{Im} D'_{F_{\mu\nu}}(k) = \text{Im} \Pi_{\mu\nu}(k) . \quad (12.238)$$



We now seek an integral representation of  $\Pi_{\mu\nu}$ . From the Lorentz covariance,

$$\begin{aligned} \langle \mathbf{0} | \mathbf{j}_\mu(x) \mathbf{j}_\nu(y) | \mathbf{0} \rangle &= i \delta_{\mu\nu} \int d\kappa^2 \sigma_1(\kappa^2) \Delta^{(+)}(x-y; \kappa^2) \\ &\quad + i \partial_\mu \partial_\nu \int d\kappa^2 \sigma_2(\kappa^2) \Delta^{(+)}(x-y; \kappa^2). \end{aligned} \quad (12.239)$$

Note also that, since  $\partial_\mu \mathbf{j}_\mu = 0$ ,

$$\sigma_1(\kappa^2) + \kappa^2 \sigma_2(\kappa^2) = 0. \quad (12.240)$$

We thus have

$$\sigma_1(\kappa^2) = \kappa^2 \sigma(\kappa^2), \quad \sigma_2(\kappa^2) = -\sigma(\kappa^2), \quad (12.241)$$

$$\langle \mathbf{0} | \mathbf{j}_\mu(x) \mathbf{j}_\nu(y) | \mathbf{0} \rangle = i \int d\kappa^2 \sigma(\kappa^2) (\delta_{\mu\nu} \square - \partial_\mu \partial_\nu) \Delta^{(+)}(x-y; \kappa^2). \quad (12.242)$$

Taking  $\mu = \nu = 4$ ,

$$-\langle \mathbf{0} | \mathbf{j}_0(x) \mathbf{j}_0(y) | \mathbf{0} \rangle = i \int d\kappa^2 \sigma(\kappa^2) \nabla^2 \Delta^{(+)}(x-y; \kappa^2). \quad (12.243)$$

Because  $\mathbf{j}_\mu$  is gauge invariant, the indefinite metric cannot appear on the left-hand side. Hence,

$$(2\pi)^3 \sum_n |\langle \mathbf{n} | \mathbf{j}_0(0) | \mathbf{0} \rangle|^2 \delta^4(p_n - k) = \mathbf{k}^2 \sigma(-k^2) \geq 0, \quad (12.244)$$

i.e.,  $\sigma(\kappa^2)$  is positive-definite as in the case of the scalar field:

$$\sigma(\kappa^2) \geq 0. \quad (12.245)$$

Using this result, for the integral representation of  $\Pi_{\mu\nu}$ , we have

$$\Pi_{\mu\nu}(k) = (k_\mu k_\nu - \delta_{\mu\nu} k^2) \Pi(-k^2), \quad (12.246)$$

$$\Pi(-k^2) = \int d\kappa^2 \frac{\sigma(\kappa^2)}{k^2 + \kappa^2 - i\epsilon} + \text{const.} \quad (12.247)$$

Since the renormalization condition requires  $\Pi(0)$  to vanish,

$$\Pi(-k^2) = \Pi(-k^2) - \Pi(0) = -k^2 \int \frac{d\kappa^2}{\kappa^2} \frac{\sigma(\kappa^2)}{k^2 + \kappa^2 - i\epsilon}. \quad (12.248)$$

At the lowest order in the perturbation theory,

$$\sigma(\kappa^2) = \frac{e_{\text{obs}}^2}{12\pi^2} \left(1 + \frac{2m^2}{\kappa^2}\right) \sqrt{1 - \frac{4m^2}{\kappa^2}} \theta(\kappa^2 - 4m^2). \quad (12.249)$$

The integral representation of the electromagnetic field propagator is

$$D'_{F_{\mu\nu}}(k) = \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2 - i\epsilon}\right) \left[ \frac{1}{k^2 - i\epsilon} + \int \frac{d\kappa^2}{\kappa^2} \frac{\sigma(\kappa^2)}{k^2 + \kappa^2 - i\epsilon} \right] + \alpha \frac{k_\mu k_\nu}{(k^2 - i\epsilon)^2}. \quad (12.250)$$

### 12.8.4 Goto–Imamura–Schwinger Term

The integral representation of the vacuum expectation value of the commutator of  $\mathbf{j}_\mu$  and  $\mathbf{j}_\nu$  is

$$\langle \mathbf{0} | [j_\mu(x), j_\nu(y)] | \mathbf{0} \rangle = i \int d\kappa^2 \sigma(\kappa^2) (\delta_{\mu\nu} \square - \partial_\mu \partial_\nu) \Delta(x - y; \kappa^2). \quad (12.251)$$

To obtain this result, we have only used the gauge invariance of  $\partial_\mu \mathbf{j}_\mu = 0$  and  $\mathbf{j}_0$ .

Assuming that  $\mu = 0$  and  $\nu = k$  ( $1, 2, 3$ ) and setting  $x_0 = y_0$ ,

$$\begin{aligned} i\langle \mathbf{0} | [j_0(x), j_k(y)] | \mathbf{0} \rangle &= - \int d\kappa^2 \sigma(\kappa^2) [\partial_0 \partial_k \Delta(x - y; \kappa^2)]_{y_0=x_0} \\ &= \frac{\partial}{\partial x_k} \delta^3(x - y) \int d\kappa^2 \sigma(\kappa^2). \end{aligned} \quad (12.252)$$

Since  $\sigma(\kappa^2) \geq 0$ , the right-hand side does not vanish. Since  $j_\mu = -ie\bar{\psi}\gamma_\mu\psi$  at the lowest order, this contradicts the result obtained from the canonical commutation relation, i.e., using the commutation relation,

$$[j_0(x), j_k(y)] = 0 \quad (x_0 = y_0). \quad (12.253)$$

The term which survives on the right-hand side of (12.252) is called the *Goto–Imamura–Schwinger term* [124, 125]. We carry out the following computation in the interaction picture using  $\partial_\mu j_\mu = 0$ :

$$\partial_\mu T[j_\mu(x), j_\nu(y)] = \delta(x_0 - y_0)[j_0(x), j_\nu(y)] \neq 0. \quad (12.254)$$

The reason why this vanished in (12.121) is that we used the  $T^*$ -product instead of the  $T$ -product, i.e.,

$$\langle 0 | T^* [j_\mu(x), j_\nu(y)] | 0 \rangle = \int d\kappa^2 \sigma(\kappa^2) (\delta_{\mu\nu} \square - \partial_\mu \partial_\nu) \Delta_F(x - y; \kappa^2). \quad (12.255)$$

In this case, it turns out that (12.121) obviously holds. As shown above, the representation which includes a product of field operators at the same point often has a singularity. A different result is often obtained from the one derived by a simple computation.

As a similar example, if we consider a neutral vector field and assume that the mass is zero, then since  $\partial_\mu \varphi_\mu = 0$ ,

$$\begin{aligned} [\varphi_\mu(x), \varphi_\nu(y)] &= i \left( \delta_{\mu\nu} - \frac{1}{m^2} \frac{\partial^2}{\partial x_\mu \partial x_\nu} \right) \Delta(x - y; m^2) \\ &= i \int d\kappa^2 \frac{\delta(\kappa^2 - m^2)}{m^2} (\delta_{\mu\nu} \square - \partial_\mu \partial_\nu) \Delta(x - y; \kappa^2). \end{aligned} \quad (12.256)$$

Therefore,

$$\sigma(\kappa^2) = \frac{1}{m^2} \delta(\kappa^2 - m^2), \quad (12.257)$$

$$i \langle 0 | [\varphi_0(x), \varphi_k(y)] | 0 \rangle = \frac{\partial}{\partial x_k} \delta^3(x - y) \frac{1}{m^2}. \quad (12.258)$$

The derivation here is exactly the same as for the Goto–Imamura–Schwinger term.

In this chapter, we have discussed the concept of renormalization by asking how we can separate the Lagrangian into the free part and the interaction part. We have shown that, at the next to lowest order, the divergence disappears, or rather that the renormalization constant can be chosen so as to make the divergence vanish. In the renormalization theory, the following stance is taken: the observed parameters can be replaced by finite experimental values, while the bare parameters which are not observed directly are allowed to be divergent. But if we now compute higher order corrections, can we be sure there will still be no divergence? Although this issue has been discussed by many people, we shall not pursue it here. We will discuss this aspect from a slightly different point of view in the last chapter.

# Chapter 13

## Classification of Hadrons and Models



Up to and including Chap. 12, we have described the fundamental framework of quantum field theory, focusing mainly on issues related to QED. However, as mentioned in Chap. 1, an important part of elementary particle theory concerns the strong interactions. Hence, we need to consider the classification, structure, and modelling of hadrons, and in this chapter, we shall turn to a more phenomenological discussion about such issues, including the group-theoretical classification of hadrons, the quark model, the parton model, and so on.

### 13.1 Unitary Groups

We start with a simple review of group theory. In a set  $G$  consisting of elements  $a, b, \dots$ , we define a product denoted  $ab$ : if  $a$  and  $b$  are elements of  $G$ , then the product  $ab$  is also an element of  $G$ . If this product satisfies the following four axioms, then we call  $G$  a *group*:

1. If  $a, b \in G$ , then their product  $ab$  is uniquely defined and  $ab \in G$ .
2. There is a unique identity element  $e$  in  $G$ , such that, for an arbitrary element  $a$  in  $G$ ,

$$ae = ea = a .$$

3. For an arbitrary element  $a$  in  $G$ , there is a unique inverse element, denoted  $a^{-1}$ , such that

$$aa^{-1} = a^{-1}a = e .$$

4. The product is associative, i.e., for arbitrary elements  $a, b, c$ ,

$$a(bc) = (ab)c .$$

### 13.1.1 Representations of a Group

If there exists a matrix  $D(a)$  for each element  $a$  in  $G$  such that

$$D(a)D(b) = D(ab) , \quad (13.1)$$

$$D(a^{-1}) = D(a)^{-1} , \quad (13.2)$$

$$D(e) = E \quad (\text{unit matrix}) , \quad (13.3)$$

then we call this set of matrices a *representation* of the group  $G$ . The vector space on which the matrix operators act is called a *representation space* and the dimension of this vector space is called the *dimension of the representation*.

The transformation of a vector in the  $n$ -dimensional representation is described by

$$\xi_a \rightarrow \xi'_a = D_a^b \xi_b , \quad a, b = 1, 2, \dots, n . \quad (13.4)$$

In addition, the contravariant representation  $\bar{\xi}^a$  associated with  $\xi_a$  is specified by the transformation

$$\bar{\xi}^a \rightarrow \bar{\xi}^{a'} = \bar{\xi}^b (D^{-1})_b^a . \quad (13.5)$$

The quantity

$$\bar{\xi}^a \xi_a \quad (\text{summing over } a) \quad (13.6)$$

is then an invariant.

Two representations  $D_1$  and  $D_2$  are said to be equivalent if there exists a matrix  $U$  such that, for every element  $a$  in  $G$ ,

$$D_2(a) = U D_1(a) U^{-1} . \quad (13.7)$$

If, for an arbitrary element  $a$ , the matrix  $D(a)$  can be diagonalized to the form

$$D(a) = \begin{pmatrix} D_1(a) & 0 & 0 \\ 0 & D_2(a) & 0 \\ 0 & 0 & \ddots \end{pmatrix} , \quad (13.8)$$

the representation is said to be *reducible* and we write

$$D = D_1 \oplus D_2 \oplus \cdots \oplus D_k . \quad (13.9)$$

A representation which does not permit such a decomposition is said to be *irreducible*.

### 13.1.2 Direct Product Representation

Suppose there are two representations  $D^\alpha$  and  $D^\beta$ , with dimensions  $\alpha$  and  $\beta$ , respectively. If we now use the bases  $\xi^\alpha(i)$  and  $\xi^\beta(j)$  of the two representation spaces to define

$$\xi^{\alpha\beta}(ij) = \xi^\alpha(i)\xi^\beta(j) , \quad (13.10)$$

we can construct an  $\alpha\beta$ -dimensional representation. This is called a *direct product representation*, and we write it as

$$D^{\alpha\beta} = D^\alpha \otimes D^\beta . \quad (13.11)$$

In general, the direct product representation is reducible.

### 13.1.3 Lie Groups

We call a transformation group parameterized by  $n$  continuous variables a *Lie group*, and  $n$  is called the *order* of the group.

### 13.1.4 Orthogonal Group $O(n)$

For  $n$  real variables  $x_1, \dots, x_n$ , the set of linear transformations  $O(n)$  of those variables which leave the quadratic form

$$\sum_{i=1}^n x_i^2 = x_1^2 + \cdots + x_n^2 \quad (13.12)$$

invariant is called the *orthogonal group*.

### 13.1.5 Unitary Group $U(n)$

Similarly, the set of linear transformations  $U(n)$  of the complex variables  $\xi_1, \dots, \xi_n$  which leave the quantity

$$\sum_{i=1}^n |\xi_i|^2 = |\xi_1|^2 + \dots + |\xi_n|^2 \quad (13.13)$$

invariant is called the *unitary group*. This is the group of  $n$ -dimensional unitary matrices. Assuming that  $U$  is one of these matrices,

$$U^\dagger U = U U^\dagger = E, \quad U^{-1} = U^\dagger. \quad (13.14)$$

The order  $N$  of this group is  $n^2$ . Considering the transformation generated by a composition of infinitesimal transformations in this group, as we did for the proper Lorentz transformation, a general element of  $U(n)$  can be written as

$$U = \exp(i\epsilon^a F_a), \quad a = 1, 2, \dots, N, \quad (13.15)$$

where we have summed over  $a$ , and  $F_a$  is a Hermitian matrix, i.e.,

$$F_a^\dagger = F_a. \quad (13.16)$$

Moreover, an infinitesimal transformation has the form

$$U = 1 + i\epsilon^a F_a, \quad (13.17)$$

where  $\epsilon^a$  is an infinitesimal parameter. The group  $SU(n)$  comprises all unitary matrices with unit determinant, i.e., matrices  $U$  such that

$$\det U = 1. \quad (13.18)$$

This is called the *special unitary group*. For the matrix written in the form (13.15), the condition (13.18) can be expressed by

$$\text{Tr } F_a = 0. \quad (13.19)$$

There are in fact  $N = n^2 - 1$  independent matrices  $F_a$ . The  $F_a$  are called *generators*. The set  $F$  of these generators satisfies the commutation relations

$$[F_i, F_j] = i f_{ijk} F_k. \quad (13.20)$$

This makes it a Lie algebra. The  $f_{ijk}$  are called *structure constants*. The commutators in (13.20) satisfy the following two conditions:

$$\text{Antisymmetry: } [F_i, F_j] = -[F_j, F_i] , \quad (13.21)$$

$$\text{Jacobi identity: } [F_i, [F_j, F_k]] + [F_j, [F_k, F_i]] + [F_k, [F_i, F_j]] = 0 . \quad (13.22)$$

In addition, the maximum number of the  $F_1, F_2, \dots, F_N$  which can be simultaneously diagonalized, in the quantum mechanical sense, is called the *rank* of the algebra. For instance, in  $SU(2)$ , choosing the three Pauli matrices  $\sigma_i/2$  ( $i = 1, 2, 3$ ) to be the  $F$ 's, we have

$$[J_i, J_j] = i\epsilon_{ijk} J_k , \quad (13.23)$$

where  $J_i = \sigma_i/2$ . Since only  $J_3$  is in general diagonal, the rank of the algebra is unity.

We consider a representation of  $U(n)$  as in (13.4),

$$\xi'_a = U_a{}^b \xi_b . \quad (13.24)$$

Taking the complex conjugate, since  $U(n)$  is unitary,

$$\xi_a^{*'} = \xi_b^* (U_a{}^b)^* = \xi_b^* (U^{-1})_b{}^a .$$

Thus, comparing with (13.5), we can choose the contravariant vector to be

$$\xi_a^* = \bar{\xi}^a . \quad (13.25)$$

### 13.1.6 Special Unitary Group $SU(2)$

A two-dimensional representation of  $U(2)$  is given by

$$U = e^{i\epsilon_0} e^{i\hat{\epsilon}\hat{\sigma}/2} . \quad (13.26)$$

The first factor is a phase and the second factor rotates in the isospin space. Moving from  $U(2)$  to  $SU(2)$ , the general form is

$$U = \exp(i\hat{\epsilon}\hat{\sigma}/2) . \quad (13.27)$$

This is a representation of the isospin group for the nucleon. In this case, if we take the three components of the quantized operators which express the isospin to be  $I_1$ ,



$I_2$ , and  $I_3$ , then they satisfy

$$[I_i, I_j] = i\epsilon_{ijk} I_k . \quad (13.28)$$

The two-dimensional representation of  $I_i$  is  $\sigma_i/2$ . Since this group is isomorphic to the rotation group in three-dimensional space, the direct product representation has the decomposition

$$D_I \otimes D_{I'} = D_{I+I'} \oplus D_{I+I'-1} \oplus \cdots \oplus D_{|I-I'|} . \quad (13.29)$$

Although  $SU(2)$  shows up in many topics in physics and is of fundamental importance, we will now discuss  $SU(3)$ , which is an extension of  $SU(2)$ .

## 13.2 The Group $SU(3)$

The group  $SU(2)$  was introduced in relation to the isospin, and all its representations were obtained from the fundamental representation using p and n (the proton and neutron, respectively) as its basis. If we consider only  $I_3$  as a hadron quantum number, the hadrons can be classified using  $SU(2)$  with rank 1, but if we include strangeness, as discussed in Chap. 1, a group of rank 2 is required. And hence,  $SU(3)$  comes on the scene.

The group  $SU(3)$  has  $3^2 - 1 = 8$  generators, and we write them as  $F_1, F_2, \dots, F_8$ . Since a traceless three-dimensional diagonal matrix has only two independent components, the group has rank two. And so, as in the case of p and n, we consider a three-component basis:

$$\text{Sakata's model} \quad \text{p, n, } \Lambda , \quad (13.30)$$

$$\text{Gell-Mann's quark model} \quad \text{u, d, s} . \quad (13.31)$$

Although we discussed both Sakata's model [64] and the quark model [74, 75] in Chap. 1, in the following we will only describe the quark model. Just like p and n, u and d correspond to the up-state and the down-state of the isospin, respectively, while s corresponds to the quark with strangeness. The electric charge  $Q$  of the hadron is connected with the third component of its isospin by the relations

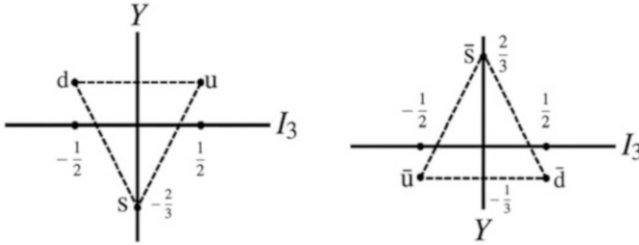
$$Q = e \left( I_3 + \frac{Y}{2} \right) , \quad (13.32)$$

$$Y = B + S . \quad (13.33)$$

Here  $Y$  is the *hypercharge*, an integer specific to each isospin multiplet.  $B$  and  $S$  are the *baryon number* and the *strangeness*, respectively. The baryon number is +1 for

**Table 13.1** Quantum numbers of the three different quarks

	$B$	$I_3$	$S$	$Q/e$
u	1/3	1/2	0	2/3
d	1/3	-1/2	0	-1/3
s	1/3	0	-1	-1/3



**Fig. 13.1** Weight diagram for quarks (*left*) and anti-quarks (*right*)

the nucleon and  $-1$  for the anti-nucleon. It is also  $+1$  for  $\Lambda^0, \Sigma^+, \Sigma^0, \Sigma^-, \Xi^0, \Xi^-$ , and so on, which decay into the nucleon by emitting a  $\pi$ -meson or a photon.

Although  $I_3$  and  $S$  are conserved for the strong interaction and the electromagnetic interaction, they are not conserved for the weak interaction and can vary by 1. If we denote quarks such as u, d, and s generically by q, we can consider that, in the quark model, mesons correspond to bound states of the form  $q\bar{q}$ , and baryons to bound states of the form  $qqq$ . In this model, the electric charges and the baryon numbers of the quarks are not integers but fractions, being integer multiples of  $1/3$ . Their values are shown in Table 13.1.

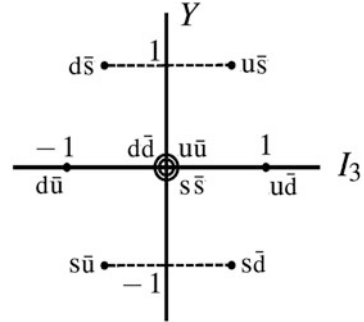
Now the quark, anti-quark, and their bound states form multiplets. Each member of a multiplet is distinguished by the values of  $I_3$  and  $Y$ . In order to make the structure of the multiplets clear, we define a so-called *weight diagram* in the  $I_3$ - $Y$  plane (see Fig. 13.1). The isospin distinguishes members in the horizontal direction, such as u and d. We simply call this  $I$ -spin. On the other hand, we sometimes introduce the  $U$ -spin distinguishing d and s and the  $V$ -spin distinguishing s and u. We will return to this when discussing the mass formula.

The weight diagram of the representation created by the direct product of two representations can be made by superimposing the center of one weight diagram on each particle point in the other weight diagram. For instance, from the direct product of q and  $\bar{q}$ , we obtain the weight diagram shown in Fig. 13.2. The center of this diagram is occupied by three particles.

The different states  $q_i \bar{q}^j$  can be expressed in a matrix form:

$$D = (q_i \bar{q}^j) = \begin{pmatrix} u\bar{u} & u\bar{d} & u\bar{s} \\ d\bar{u} & d\bar{d} & d\bar{s} \\ s\bar{u} & s\bar{d} & s\bar{s} \end{pmatrix}. \tag{13.34}$$

**Fig. 13.2** Weight diagram for the quark–anti-quark system



We decompose this direct product as follows:

$$D_i^j = \frac{1}{3} \delta_i^j D_k^k + \left( D_i^j - \frac{1}{3} \delta_i^j D_k^k \right). \tag{13.35}$$

The first term is a scalar singlet and the second term is a traceless tensor octet. The coefficient of the first term is

$$\frac{1}{3} (u\bar{u} + d\bar{d} + s\bar{s}) = \frac{1}{\sqrt{3}} \left( \frac{u\bar{u} + d\bar{d} + s\bar{s}}{\sqrt{3}} \right). \tag{13.36}$$

This belongs to the one-dimensional representation of  $SU(3)$ . Since this meson is  $0^-$ , it corresponds to the particle called  $\eta'(958)$ , where 958 indicates the mass in units of MeV. We thus replace the matrix elements of  $D$  by the  $0^-$  mesons. For example,

$$(13.36) \longrightarrow \frac{1}{\sqrt{3}} \eta'.$$

We now consider the octet, in which the off-diagonal terms are replaced directly by suitable mesons without the need for any adjustments. On the other hand, the diagonal terms are first decomposed in terms of the isospin. For instance, the second term in (13.35), corresponding to  $i = j = 1$ , is

$$\begin{aligned} \frac{1}{3} (2u\bar{u} - d\bar{d} - s\bar{s}) &= \frac{1}{6} (u\bar{u} + d\bar{d} - 2s\bar{s}) + \frac{1}{2} (u\bar{u} - d\bar{d}) \\ &\rightarrow \frac{1}{\sqrt{6}} \eta^0 + \frac{1}{\sqrt{2}} \pi^0. \end{aligned} \tag{13.37}$$

The off-diagonal terms are replaced as follows:

$$u\bar{s} \rightarrow K^+, \quad d\bar{s} \rightarrow K^0, \quad s\bar{u} \rightarrow K^-, \quad s\bar{d} \rightarrow \bar{K}^0. \tag{13.38}$$

These replacements are useful when discussing only group-theoretic properties, without any relation to the details of the interactions. Thus, the part of  $D$  corresponding to the  $0^-$  mesons is

$$D(0^-) = \frac{1}{\sqrt{3}}\eta' + \begin{pmatrix} \frac{\eta^0}{\sqrt{6}} + \frac{\pi^0}{\sqrt{2}} & \pi^+ & K^+ \\ \pi^- & \frac{\eta^0}{\sqrt{6}} - \frac{\pi^0}{\sqrt{2}} & K^0 \\ K^- & \bar{K}^0 & -\frac{2}{\sqrt{6}}\eta^0 \end{pmatrix}. \quad (13.39)$$

The decomposition of the direct product into irreducible representations is

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}, \quad (13.40)$$

where  $\mathbf{3}$  and  $\bar{\mathbf{3}}$  are the three-dimensional representations to which the quark and the anti-quark belong, respectively. A similar decomposition can be carried out for the  $1^-$  mesons. Thus, the part of  $D$  corresponding to the  $1^-$  mesons is

$$D(1^-) = \frac{1}{\sqrt{3}}\omega + \begin{pmatrix} \frac{\phi^0}{\sqrt{6}} + \frac{\rho^0}{\sqrt{2}} & \rho^+ & K^{*+} \\ \rho^- & \frac{\phi^0}{\sqrt{6}} - \frac{\rho^0}{\sqrt{2}} & K^{*0} \\ K^{*-} & \bar{K}^{*0} & -\frac{2}{\sqrt{6}}\phi^0 \end{pmatrix}. \quad (13.41)$$

### 13.2.1 Generators of $SU(3)$

The generators of  $SU(2)$  in the two-dimensional spinor representation are

$$I_i \rightarrow \frac{1}{2}\sigma_i. \quad (13.42)$$

What about  $SU(3)$ ? Considering the  $2 \times 2$  part of  $D(0^-)$ , the isospin-1 part can be written as

$$\begin{pmatrix} \frac{\pi^0}{\sqrt{2}} & \pi^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \pi^3 & \pi^1 - i\pi^2 \\ \pi^1 + i\pi^2 & -\pi^3 \end{pmatrix} = \frac{1}{\sqrt{2}} \hat{\tau} \hat{\pi}. \quad (13.43)$$

For a Hermitian field, we write the octet part in the form

$$\frac{1}{\sqrt{2}}(\lambda_i)_a{}^b M^i = \frac{1}{\sqrt{2}}\hat{\lambda}\hat{M}, \quad (13.44)$$

where  $M$  stands for the eight-component Hermitian meson field.

We use the following numbering for, say,  $0^-$  :

$$\begin{pmatrix} \frac{\pi^0}{\sqrt{2}} + \frac{\eta^0}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} + \frac{\eta^0}{\sqrt{6}} & K^0 \\ K^- & \bar{K}^0 & -\frac{2\eta^0}{\sqrt{6}} \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} M^3 + \frac{M^8}{\sqrt{3}} & M^1 - iM^2 & M^4 - iM^5 \\ M^1 + iM^2 & -M^3 + \frac{M^8}{\sqrt{3}} & M^6 - iM^7 \\ M^4 + iM^5 & M^6 + iM^7 & -\frac{2}{\sqrt{3}}M^8 \end{pmatrix}. \quad (13.45)$$

Thus, corresponding to (13.42) in the three-dimensional representation, the generators  $\lambda_i$  generalize  $\sigma_i$  and replace the generators  $F_i$  for  $SU(3)$ :

$$F_i \longrightarrow \frac{\lambda_i}{2}. \quad (13.46)$$

Following Gell-Mann, these eight matrices  $\lambda_i$  can be chosen as follows:

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned} \quad (13.47)$$

Using this choice for the  $\lambda_i$ , the structure constants can be determined from

$$\left[ \frac{\lambda_i}{2}, \frac{\lambda_j}{2} \right] = if_{ijk} \frac{\lambda_k}{2}. \quad (13.48)$$

The  $\lambda_i$  are all Hermitian matrices and normalized so that

$$\text{Tr}(\lambda_i \lambda_j) = 2\delta_{ij}. \quad (13.49)$$

This choice is important. The symmetries of the structure constants can be derived from this equation. Furthermore, using (13.49), the anti-commutation relations can be written in the form

$$\left\{ \frac{\lambda_i}{2}, \frac{\lambda_j}{2} \right\} = \frac{1}{3} \delta_{ij} + d_{ijk} \frac{\lambda_k}{2}. \quad (13.50)$$

This introduces another set of structure constants.

The two sets of structure constants  $f_{ijk}$  and  $d_{ijk}$  above are completely anti-symmetric and completely symmetric in the indices  $i$ ,  $j$ , and  $k$ , respectively. To see this, we use (13.49):

$$4i f_{ijk} = \text{Tr}([\lambda_i, \lambda_j] \lambda_k) = \text{Tr}(\lambda_j [\lambda_k, \lambda_i]) = \text{Tr}(\lambda_i [\lambda_j, \lambda_k]). \quad (13.51)$$

This shows that  $f_{ijk}$  is completely anti-symmetric. Similarly, the symmetry of  $d_{ijk}$  is obvious from the equation

$$4d_{ijk} = \text{Tr}(\{\lambda_i, \lambda_j\} \lambda_k) = \text{Tr}(\lambda_j \{\lambda_k, \lambda_i\}) = \text{Tr}(\lambda_i \{\lambda_j, \lambda_k\}). \quad (13.52)$$

### 13.2.2 $I$ -, $U$ -, and $V$ -Spin

The  $SU(3)$  algebra has several kinds of  $SU(2)$  algebras as subalgebras, and many useful consequences can be derived from this. Firstly, examining the content of (13.32), we obtain

$$F_3 = I_3, \quad F_8 = \frac{\sqrt{3}}{2} Y. \quad (13.53)$$

Thus, for the quarks,

$$Q = e \left( F_3 + \frac{F_8}{\sqrt{3}} \right). \quad (13.54)$$

This relation holds for all hadrons composed of quarks and anti-quarks. We define the  $I$ -,  $U$ -, and  $V$ -spin as follows:

$$\begin{aligned} I_{\pm} &= F_1 \pm iF_2, & U_{\pm} &= F_6 \pm iF_7, & V_{\pm} &= F_4 + iF_5, \\ I_3 &= F_3, & U_3 &= -\frac{1}{2}F_3 + \frac{\sqrt{3}}{2}F_8, & V_3 &= -\frac{1}{2}F_3 - \frac{\sqrt{3}}{2}F_8. \end{aligned} \quad (13.55)$$

Hence,  $I, U,$  and  $V$  generate three  $SU(2)$  subalgebras. Moreover, among these three algebras, the following condition is satisfied:

$$I_3 + U_3 + V_3 = 0. \tag{13.56}$$

It is straightforward to show that the commutation relations among these are

$$\begin{aligned} [I_3, I_{\pm}] &= \pm I_{\pm}, & [I_3, U_{\pm}] &= \mp \frac{1}{2} U_{\pm}, & [I_3, V_{\pm}] &= \mp \frac{1}{2} V_{\pm}, \\ [U_3, I_{\pm}] &= \mp \frac{1}{2} I_{\pm}, & [U_3, U_{\pm}] &= \pm U_{\pm}, & [U_3, V_{\pm}] &= \mp \frac{1}{2} V_{\pm}, \\ [[V_3, I_{\pm}] &= \mp \frac{1}{2} I_{\pm}, & [V_3, U_{\pm}] &= \mp \frac{1}{2} U_{\pm}, & [V_3, V_{\pm}] &= \pm V_{\pm}, \\ [Y, I_{\pm}] &= 0, & [Y, U_{\pm}] &= \pm U_{\pm}, & [Y, V_{\pm}] &= \pm V_{\pm}, \\ [I_+, I_-] &= 2I_3, & [U_+, U_-] &= 2U_3, & [V_+, V_-] &= 2V_3, \\ [I_{\pm}, U_{\pm}] &= \pm V_{\pm}, & [U_{\pm}, V_{\pm}] &= \pm I_{\mp}, & [V_{\pm}, I_{\pm}] &= \pm U_{\mp}. \end{aligned} \tag{13.57}$$

### 13.2.3 Three-Body Quark Systems

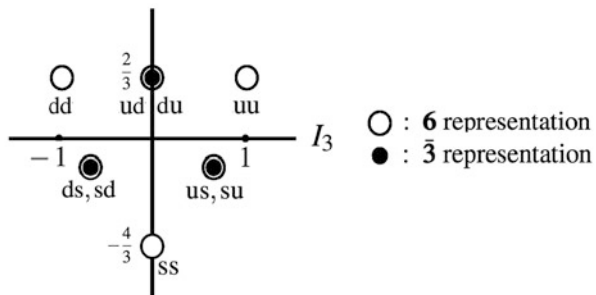
As we have seen before, a meson is a quark–anti-quark bound state  $q\bar{q}$ , whereas a baryon (or heavy particle) is a three-quark bound state  $qqq$ . We thus begin with the weight diagram for a two-quark state  $qq$ .

We can carry out the following direct-product decomposition corresponding to (13.40):

$$\mathbf{3} \otimes \mathbf{3} = \bar{\mathbf{3}} \oplus \mathbf{6}. \tag{13.58}$$

Since three of the six points appearing in Fig. 13.3 are double points, we have to make a suitable linear combination of the relevant states. To do this, we have to

**Fig. 13.3** Weight diagram for the two-quark system



make symmetric and anti-symmetric combinations:

$$q_i q_j = T_{ij} = \frac{1}{2}(T_{ij} - T_{ji}) + \frac{1}{2}(T_{ij} + T_{ji}) = \frac{1}{2}T_{[ij]} + \frac{1}{2}T_{(ij)} . \quad (13.59)$$

The anti-symmetric one is equal to the contravariant vector

$$\epsilon^{ijk} T_{[ij]} = T^k . \quad (13.60)$$

This is a three-dimensional representation and transforms like  $\bar{q}^k$ , so it is the representation  $\bar{\mathbf{3}}$ . The symmetric part, on the other hand, has six components and corresponds to the representation  $\mathbf{6}$ . This confirms (13.58).

We now look at the transformation of  $\epsilon$  under  $SU(3)$ :

$$\epsilon'_{ijk} = U_i^l U_j^m U_k^n \epsilon_{lmn} = \epsilon_{ijk} \det U .$$

Thus, since  $\det U = 1$ ,

$$\epsilon'_{ijk} = \epsilon_{ijk} . \quad (13.61)$$

Similarly,

$$\epsilon'^{ijk} = \epsilon^{ijk} . \quad (13.62)$$

We can now list the states belonging to the  $\bar{\mathbf{3}}$  representation and the  $\mathbf{6}$  representation:

$$\{\bar{\mathbf{3}}\} \quad \left\{ \begin{array}{l} I = 0 \quad \frac{1}{\sqrt{2}}(\text{ud} - \text{du}) , \\ I = \frac{1}{2} \quad \frac{1}{\sqrt{2}}(\text{us} - \text{su}) , \quad \frac{1}{\sqrt{2}}(\text{ds} - \text{sd}) , \end{array} \right. \quad (13.63)$$

$$\{\mathbf{6}\} \quad \left\{ \begin{array}{l} I = 0 \quad \text{ss} \\ I = \frac{1}{2} \quad \frac{1}{\sqrt{2}}(\text{us} + \text{su}) , \quad \frac{1}{\sqrt{2}}(\text{ds} + \text{sd}) , \\ I = 1 \quad \text{uu} , \quad \frac{1}{\sqrt{2}}(\text{ud} + \text{du}) , \quad \text{dd} . \end{array} \right. \quad (13.64)$$

We now consider a general tensor:

$$T_{(q)}^{(p)} = T_{ij\dots k}^{\alpha\beta\dots\gamma} , \quad \text{such as } \bar{q}^\alpha \bar{q}^\beta \dots \bar{q}^\gamma q_i q_j \dots q_k , \quad (13.65)$$



and an irreducible representation, i.e., one where the tensor cannot be reduced to a lower-dimensional tensor by contractions with the tensors

$$\delta_i^j, \quad \epsilon^{ijk}, \quad \epsilon_{ijk}. \quad (13.66)$$

Reducible tensors can be reduced to lower-dimensional tensors by contractions, as in the following examples:

$$\delta_\alpha^i T_{ij\dots l}^{\alpha\beta\dots\delta} = B_{j\dots l}^{\beta\dots\delta}, \quad \epsilon_{\mu\alpha\beta} T_{ij\dots l}^{\alpha\beta\dots\delta} = C_{\mu ij\dots l}^{\gamma\dots\delta}, \quad \epsilon^{mij} T_{ij\dots l}^{\alpha\beta\dots\delta} = D_{k\dots l}^{m\alpha\beta\dots\delta}.$$

If  $T$  is irreducible, it turns out that the right-hand sides of the equations above vanish. This is the condition that  $T$  be symmetric in  $\alpha, \beta, \dots$  and  $i, j, \dots$ , and the trace vanishes, i.e.,

$$T_{ij\dots l}^{i\beta\dots\delta} = 0. \quad (13.67)$$

If  $T_{(q)}^{(p)}$  satisfies the condition above, there is one irreducible representation  $D(p, q)$ . Here,  $p$  and  $q$  are the numbers of superscripts and subscripts, respectively. In this case, the dimension  $N$  of  $D(p, q)$  is given by

$$N = (1 + p)(1 + q) \left[ 1 + \frac{1}{2}(p + q) \right]. \quad (13.68)$$

**Proof** First we ask how many tensors there are which are symmetric in  $p$  indices. We consider as indices the case where the numbers 1, 2, and 3 are  $x, y$ , and  $z$ . We must therefore count how many choices of  $x, y$ , and  $z$  there are. This is equivalent to counting the number of solutions to  $x + y + z = p$ .  $x$  can take any of the values  $x = 0, 1, \dots, p$ . Once  $x$  is fixed, we need to find the number of solutions to

$$y + z = p - x. \quad (13.69)$$

This yields  $p - x + 1$ . The total will thus be

$$\sum_{x=0}^p (p - x + 1) = \frac{1}{2}(p + 1)(p + 2). \quad (13.70)$$

Now, for the  $p$  superscripts and the  $q$  subscripts, the number of symmetric tensors is

$$\frac{1}{4}(p + 1)(p + 2)(q + 1)(q + 2). \quad (13.71)$$

We now take into account the vanishing trace condition. To do this, we reduce the number of symmetric tensors by changing  $p$  into  $p - 1$  and  $q$  into  $q - 1$  and

subtracting the result:

$$\begin{aligned} N &= \frac{1}{4}(p+1)(p+2)(q+1)(q+2) - \frac{1}{4}p(p+1)q(q+1) \\ &= (1+p)(1+q) \left[ 1 + \frac{1}{2}(p+q) \right], \end{aligned}$$

which proves (13.68).

We can now specify the following examples of  $D(p, q)$ :

$$\begin{aligned} D(0, 0) &= \{\mathbf{1}\}, & D(1, 1) &= \{\mathbf{8}\}, & D(0, 3) &= \{\mathbf{10}\}, \\ D(0, 1) &= \{\mathbf{3}\}, & D(0, 2) &= \{\mathbf{6}\}, & D(3, 0) &= \{\bar{\mathbf{10}}\}, \\ D(1, 0) &= \{\bar{\mathbf{3}}\}, & D(2, 0) &= \{\bar{\mathbf{6}}\}, & \dots & \end{aligned} \quad (13.72)$$

And so we come back to the first issue, i.e., investigating the representation of the three-quark system:

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = (\mathbf{3} \oplus \mathbf{6}) \otimes \mathbf{3} = (\mathbf{1} \oplus \mathbf{8}) \oplus (\mathbf{8} \oplus \mathbf{10}). \quad (13.73)$$

Nucleons such as  $p$  and  $n$  belong to the eight-dimensional representation here. An excited state of the nucleon known as  $\Delta$ , which arises in  $\pi$ -meson–nucleon scattering, belongs to the ten-dimensional representation.

In the same vein as (13.39) and (13.41), we now express the eight-dimensional representation which includes the nucleon in the form of a matrix:

$$\mathcal{B} = \begin{pmatrix} \frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda^0}{\sqrt{6}} & \Sigma^+ & p \\ \Sigma^- & -\frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda^0}{\sqrt{6}} & n \\ \Xi^- & \Xi^0 & -\frac{2}{\sqrt{6}}\Lambda^0 \end{pmatrix}. \quad (13.74)$$

We can now derive the anti-baryon octet from the baryon octet using the charge conjugation operator  $\mathcal{C}$ , whence

$$\mathcal{C}^{-1} \mathcal{B}^T \mathcal{C} = C \bar{\mathcal{B}}, \quad (13.75)$$

where  $T$  stands for the transposition of the  $3 \times 3$  matrix. This is a generalization of (3.198). Finally, imitating (13.74), we denote the matrix representations of  $D(0^-)$  and  $D(1^-)$  by  $\mathcal{P}$  and  $\mathcal{V}$ , respectively.

In the treatment above, the idea of using the quark as the fundamental representation is a mathematical tool for deriving other representations. Physically, even if there are no quarks, we can derive all observable representations by

considering eight-dimensional baryon representations and eight-dimensional meson representations. We call this the *eightfold way* [52, 67, 68], in contrast with Sakata's model, which starts with the three-dimensional representation of the baryon,

### 13.2.4 Mass Formulas

$SU(3)$  is not a symmetry group in the strict sense. In the real world, this symmetry is broken. However, by treating this symmetry-breaking perturbatively, meaningful consequences can be obtained, both qualitatively and semi-quantitatively. In the following, we investigate the mass difference between hadrons and also the magnetic moment, using a perturbative computation in the context of the quark model. Later in this chapter, we will show the corresponding computation based on the full quark model.

The most important assumption introduced here is that, when we split the Hamiltonian into the invariant part under  $SU(3)$  and the other part, viz.,

$$H = H_{\text{inv}} + H_{\text{SB}} , \quad (13.76)$$

the symmetry-breaking part  $H_{\text{SB}}$  has the same transformation property as  $F_8$  and we can treat this term as a perturbation. Now,

$$\frac{\sqrt{3}}{2} F_8 = U_3 + \frac{1}{2} I_3 . \quad (13.77)$$

We introduce the following  $U$ -spin triplet:

$$\mathbf{n} , \quad -\frac{1}{2} \Sigma^0 + \frac{\sqrt{3}}{2} \Lambda^0 , \quad \Xi^0 . \quad (13.78)$$

The matrix elements of  $U_3$  and  $I_3$  in the eight-dimensional representation are

$$\begin{aligned} \langle \mathbf{n} | U_3 | \mathbf{n} \rangle &= 1 , \\ \left\langle -\frac{1}{2} \Sigma^0 + \frac{\sqrt{3}}{2} \Lambda^0 \middle| U_3 \middle| -\frac{1}{2} \Sigma^0 + \frac{\sqrt{3}}{2} \Lambda^0 \right\rangle &= 0 , \\ \langle \Xi^0 | U_3 | \Xi^0 \rangle &= -1 , \end{aligned} \quad (13.79)$$

$$\begin{aligned} \langle \mathbf{n} | I_3 | \mathbf{n} \rangle &= -\frac{1}{2} , \\ \left\langle -\frac{1}{2} \Sigma^0 + \frac{\sqrt{3}}{2} \Lambda^0 \middle| I_3 \middle| -\frac{1}{2} \Sigma^0 + \frac{\sqrt{3}}{2} \Lambda^0 \right\rangle &= 0 , \\ \langle \Xi^0 | I_3 | \Xi^0 \rangle &= \frac{1}{2} . \end{aligned} \quad (13.80)$$

If  $H_{SB}$  transforms in the same way as (13.77),

$$\begin{aligned} \langle n | H_{SB} | n \rangle + \langle \Xi^0 | H_{SB} | \Xi^0 \rangle &= 2 \left\langle -\frac{1}{2} \Sigma^0 + \frac{\sqrt{3}}{2} \Lambda^0 \middle| H_{SB} \middle| -\frac{1}{2} \Sigma^0 + \frac{\sqrt{3}}{2} \Lambda^0 \right\rangle \\ &= \frac{1}{2} \langle \Sigma^0 | H_{SB} | \Sigma^0 \rangle + \frac{3}{2} \langle \Lambda^0 | H_{SB} | \Lambda^0 \rangle. \end{aligned} \quad (13.81)$$

Thus, adding the common mass of the invariant part to each diagonal element, which is taken as a shift in the mass, we deduce the relation

$$m(n) + m(\Xi) = \frac{1}{2} m(\Sigma) + \frac{3}{2} m(\Lambda). \quad (13.82)$$

Additionally, for the 10-dimensional representation, considering the spin quartet,

$$\Delta^-, \quad \Sigma^{*-}, \quad \Xi^{*-}, \quad \Omega^-, \quad (13.83)$$

and carrying out a similar computation,

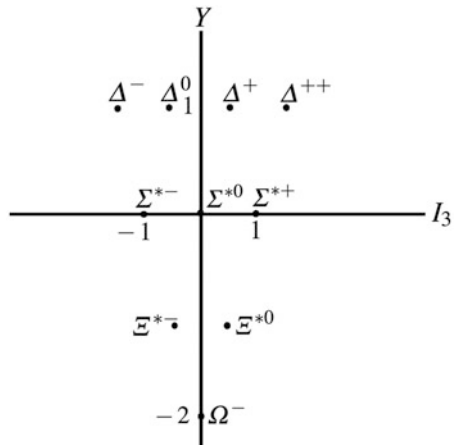
$$m(\Delta) - m(\Sigma^*) = m(\Sigma^*) - m(\Xi^*) = m(\Sigma^*) - m(\Omega). \quad (13.84)$$

These are called the *Gell-Mann mass formulas* [67]. The weight diagram of the ten-dimensional representation is shown in Fig. 13.4. It was using these formulas that  $\Omega^-$  was found in [71].

Okubo also showed that, using the isospin  $I$  and hypercharge  $Y$ , the mass of a particle in one of the  $SU(3)$  multiplets can be expressed in the form [70]

$$m = a + bY + c \left[ I(I + 1) - \frac{Y^2}{4} \right]. \quad (13.85)$$

**Fig. 13.4** Weight diagram for the decuplet



Although this equation holds for both bosons and fermions, a better agreement with experiment is obtained for bosons by writing the equation for  $m^2$ .

### 13.2.5 Baryon Magnetic Moments

In the example above, we have assumed that the term  $H_{SB}$  which breaks the symmetry transforms in the same way as  $Y$  (which is proportional to  $F_8$ ). It may be assumed that the electromagnetic interaction, which was not considered there, has the same transformation property as  $Q$ . Consequently, the magnetic moment can be considered to have the same transformation property. Note that  $Q$  commutes with the  $U$ -spin:

$$[U_{\pm}, Q] = 0, \quad [U_3, Q] = 0. \quad (13.86)$$

Moreover, since the electromagnetic interaction is a sum of the isospin singlet and triplet, we have

$$\mu(\Sigma^+) + \mu(\Sigma^-) = 2\mu(\Sigma^0), \quad (13.87)$$

where  $\mu(\Sigma)$  stands for the magnetic moment of  $\Sigma$ , and  $\Sigma^+$ ,  $\Sigma^0$ , and  $\Sigma^-$  belong to the isospin triplet.

Since the magnetic moments of particles belonging to the same multiplet of the  $U$ -spin are all equal, we can classify the baryon octet by the  $U$ -spin:

$$\begin{array}{ll} U - \text{spin singlet} & U_0^0 = \frac{1}{2}(\sqrt{3}\Sigma^0 + \Lambda^0), \\ U - \text{spin doublet} & \Sigma^+, \quad p, \quad \text{and} \quad \Xi^-, \quad \Sigma^-, \\ U - \text{spin triplet} & \Xi^0, \quad U_1^0 = \frac{1}{2}(\Sigma^0 - \sqrt{3}\Lambda^0), \quad n. \end{array}$$

From these, we immediately obtain the following equalities:

$$\mu(\Sigma^+) = \mu(p), \quad \mu(\Xi^-) = \mu(\Sigma^-), \quad \mu(\Xi^0) = \mu(U_1^0) = \mu(n). \quad (13.88)$$

Furthermore, the sum of the eigenvalues of  $Q$  over all members of the octet vanishes:

$$\sum_{\{8\}} \mu = 0. \quad (13.89)$$

Since (13.87), (13.88), and (13.89) amount to six conditions, we can express the eight magnetic moments in terms of just two magnetic moments, say  $\mu(p)$  and  $\mu(n)$ . It is more useful to use  $\Sigma^0$  and  $\Lambda^0$  as particles, rather than  $U_0^0$  and  $U_1^0$ , so we consider the latter as eigenstates of  $\mu$  and use the condition

$$\langle U_0^0 | \mu | U_1^0 \rangle = 0. \quad (13.90)$$

This yields the relations

$$2\mu(\Sigma^0 \Lambda^0) \equiv 2\langle \Sigma^0 | \mu | \Lambda^0 \rangle = \sqrt{3}[\mu(\Sigma^0) - \mu(\Lambda^0)], \quad (13.91)$$

$$\mu(U_1^0) = \frac{3}{2}\mu(\Lambda^0) - \frac{1}{2}\mu(\Sigma^0). \quad (13.92)$$

Here,  $\mu(\Sigma^0 \Lambda^0)$  is called a *transition magnetic moment*. It contributes to the decay  $\Sigma^0 \rightarrow \Lambda^0 + \gamma$ . We can now obtain the magnetic moments of all members of the octet:

$$\begin{aligned} \mu(\Sigma^+) &= \mu(p), & \mu(\Sigma^0) &= -\frac{1}{2}\mu(n), & \mu(\Sigma^-) &= -\mu(p) - \mu(n), \\ \mu(\Lambda^0) &= \frac{1}{2}\mu(n), & \mu(\Xi^0) &= \mu(n), & \mu(\Xi^-) &= -\mu(p) - \mu(n). \end{aligned} \quad (13.93)$$

Taking  $e/2m_p$  as the basic unit, known as the *nuclear magneton*, and using

$$\mu(p) = 2.79, \quad \mu(n) = -1.91, \quad (13.94)$$

we can compute the magnetic moments of all members of the baryon octet and compare them with experimental data. The results are displayed in Table 13.2.

Although the agreement of theoretical values and experimental values is qualitatively good, it is not so good quantitatively. Later, we will return to a more quantitative discussion using the quark model. In any case, the discussions so far are group theoretical, and not based directly on the quark model.

**Table 13.2** Choosing  $C_j$  and  $O_j$

	$\mu(\Lambda^0)$	$\mu(\Sigma^+)$	$\mu(\Sigma^-)$	$\mu(\Xi^-)$	$\mu(\Xi^0)$
Theoretical value	-0.96	2.79	-0.88	-0.88	-1.91
Experimental value	$-0.613 \pm 0.004$	$2.38 \pm 0.02$	$-1.10 \pm 0.05$	$-1.85 \pm 0.75$	$-1.250 \pm 0.014$

### 13.2.6 $SU(3)$ -Invariant Interactions

So far, we have treated the octets of baryons, pseudo-scalar mesons, and vector mesons. These are expressed by  $3 \times 3$  matrices like  $\mathcal{B}$ ,  $\mathcal{P}$ , and  $\mathcal{V}$ . We now consider the transformation under  $SU(3)$  (13.24). For this transformation, the matrix  $D$  given by (13.34) transforms by

$$D \rightarrow UDU^{-1} . \quad (13.95)$$

Thus,  $\mathcal{B}$ ,  $\mathcal{P}$ , and  $\mathcal{V}$ , and so on, all transform in the same way as (13.95). This means that any trace of the product of these matrices will be invariant under the  $SU(3)$  transformations. Moreover, under charge conjugation,

$$\mathcal{C}^{-1} \mathcal{P}^T \mathcal{C} = \mathcal{P} , \quad \mathcal{C}^{-1} \mathcal{V}^T \mathcal{C} = -\mathcal{V} . \quad (13.96)$$

These correspond to (13.75) in the case of baryons.

An  $SU(3)$ -invariant Yukawa-type interaction is

$$\mathcal{H}_{\text{int}} = \sqrt{2}D \text{Tr}(\bar{\mathcal{B}}\{\mathcal{M}, \mathcal{B}\}) + \sqrt{2}F \text{Tr}(\bar{\mathcal{B}}[\mathcal{M}, \mathcal{B}]) , \quad (13.97)$$

for  $\mathcal{M} = \mathcal{P}$  or  $\mathcal{V}$ . The reason why we use the notation  $D$  and  $F$  is that  $d$  and  $f$  appear as the Clebsch–Gordan coefficients connecting different kinds of members when we compute the trace. In the case of the pseudo-scalar meson, the Yukawa-type coupling constant  $G$  appearing in (6.96) is given by

$$G = D + F . \quad (13.98)$$

Note also that, if there are interactions corresponding to

$$\mathcal{V} \rightarrow \mathcal{P} + \mathcal{P} , \quad \mathcal{V} \rightarrow \mathcal{V} + \mathcal{P} , \quad (13.99)$$

it is easy to see that only type  $F$  or type  $D$  can occur, respectively.

### 13.2.7 Casimir Operator

Up to now, we have introduced matrices corresponding to the octets  $\mathcal{B}$ ,  $\mathcal{P}$ , and  $\mathcal{V}$ . Considering a member corresponding to their matrix elements, say the one such that  $M$  in (13.45) is replaced by  $F$ , we write it as

$$\mathcal{F} = \sum_{i=1}^8 \frac{\lambda_i}{\sqrt{2}} F_i . \quad (13.100)$$

Although the  $F_i$  transform under  $SU(3)$ , their transformation rules are once again given by (13.95). Hence,

$$\text{Tr } \mathcal{F}^2, \quad \text{Tr } \mathcal{F}^3, \quad (13.101)$$

are invariant under the  $SU(3)$  transformations, and therefore,

$$[F_i, \text{Tr } \mathcal{F}^n] = 0. \quad (13.102)$$

An operator which is a polynomial in the generators  $F_i$  and commutes with all the  $F_i$  as just described is called a *Casimir operator*. It is clear from the definition that the Casimir operator has the same eigenvalue for all members of the same irreducible representation. The operator

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2, \quad (13.103)$$

for the rotation group in three-dimensional space is an example of a Casimir operator, and its eigenvalue  $J(J+1)$  specifies the irreducible representation.

In the case of  $SU(3)$ , there are only two independent quantities in  $\text{Tr } \mathcal{F}^n$ . If we take

$$C_n = \text{Tr } \mathcal{F}^n, \quad (13.104)$$

then it is straightforward to show that the following recursion equation holds:

$$C_{n+3} = \frac{1}{2}C_2C_{n+1} + \frac{1}{3}C_3C_n. \quad (13.105)$$

Using  $C_1 = 0$ , all the  $C_n$  can be expressed in terms of  $C_2$  and  $C_3$ . Thus, we can choose  $C_2$  and  $C_3$  as independent Casimir operators. For instance,  $C_4$  is expressed in terms of  $C_2$  by

$$\text{Tr } \mathcal{F}^4 = \frac{1}{2}(\text{Tr } \mathcal{F}^2)^2, \quad C_4 = \frac{1}{2}C_2^2. \quad (13.106)$$

We can express  $C_2$  and  $C_3$  in terms of the  $F_i$ :

$$C_2 = \text{Tr } \mathcal{F}^2 = \sum_i F_i^2, \quad C_3 = \text{Tr } \mathcal{F}^3 = \frac{1}{\sqrt{2}} \sum_{i,j,k} d_{ijk} F_i F_j F_k. \quad (13.107)$$

We write the eigenvalue of  $C_2$  for a given irreducible representation  $R$  as  $C_2(R)$ . In particular, we denote the eigenvalue for the adjoint representation by  $C_2(G)$ . Now we know that, for the  $SU(3)$  group, the irreducible representations can be expressed



as  $D(p, q)$ , and it can be shown that, for  $R = D(p, q)$ ,

$$C_2(R) = \frac{1}{3}(p^2 + pq + q^2) + p + q . \quad (13.108)$$

This result will be used when discussing the gauge theory later on.

In this section we have used quarks as a mathematical tool and discussed how to apply the  $SU(3)$  group to classify hadrons. Later, we will explain why quarks should be considered as real physical entities.

### 13.3 Universality of $\pi$ -Meson Decay Interactions

So far, we have discussed the classification of hadrons. We shall now consider leptons. As mentioned in Chap. 1, leptons are fermions which are not affected by the strong interaction. These include those carrying electric charges, such as the electron and the muon, as well as the neutral neutrino. The latter has several forms, e.g., the electron neutrino  $\nu_e$ , which forms a doublet with the electron, and the muon neutrino  $\nu_\mu$ , which forms a doublet with the muon. We call the pairs  $(e, \nu_e)$  and  $(\mu, \nu_\mu)$  the *first generation* and the *second generation*, respectively. It seems that the lepton number in each generation is conserved. M.L. Perl found a third charged lepton in [126]. This is called the *tauon* or *tau-lepton* and denoted by  $\tau$ . Of course, there is also a tau neutrino which is paired with the tauon. We thus know of three generations of leptons:

$$(e, \nu_e) , \quad (\mu, \nu_\mu) , \quad (\tau, \nu_\tau) . \quad (13.109)$$

The masses of the charged leptons are

$$\begin{aligned} m_e &= 0.5110034 \pm 0.0000014 \text{ MeV} , & m_\mu &= 105.65943 \pm 0.00018 \text{ MeV} , \\ m_\tau &= 1776.86 \pm 0.12 \text{ MeV} . \end{aligned} \quad (13.110)$$

Since the mass of the  $\pi$ -meson or pion is

$$m_\pi = 139.5637 \pm 0.0007 \text{ MeV} , \quad (13.111)$$

the pion can decay into lepton pairs of the first and second generations:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu , \quad \pi^+ \rightarrow e^+ + \nu_e . \quad (13.112)$$

However, it is the tau-lepton that can decay into the pion:

$$\tau^+ \rightarrow \pi^+ + \bar{\nu}_\tau . \quad (13.113)$$

We assume that the interaction between the lepton  $l$  ( $l = e, \mu, \tau$ ), the  $l$ -neutrino  $\nu_l$ , and  $\pi$  has the form

$$\mathcal{H}_{\text{decay}} = \frac{ig}{m_\pi} \bar{\psi}_{\nu_l} \gamma_\lambda (1 + \gamma_5) \psi_l \partial_\lambda \varphi + \text{h.c.} \quad (13.114)$$

In Sect. 9.2, assuming that the coupling constant  $g$  has the same value for the first and second generations, and computing the branching ratio of (13.112), we obtained (9.61), i.e.,

$$\frac{\Gamma(\pi^+ \rightarrow e^+ + \nu_e)}{\Gamma(\pi^+ \rightarrow \mu^+ + \nu_\mu)} = 1.29 \times 10^{-4}. \quad (13.115)$$

This agrees well with the experimental value of  $(1.266 \pm 0.023) \times 10^{-4}$ . Let us assume that it is the same for the third generation. The decay width for the reaction (13.113) is

$$\Gamma(\tau^+ \rightarrow \pi^+ + \bar{\nu}_\tau) = \frac{g^2}{8\pi} m_\tau \left( \frac{m_\tau}{m_\pi} \right)^2 \left[ 1 - \left( \frac{m_\pi}{m_\tau} \right)^2 \right]. \quad (13.116)$$

In this case, since we have not summed over the spins for  $\tau^+$ , but averaged over them because the reaction differs from (13.112) (in the sense that the charged lepton  $\tau$  is in the initial, and not the final state),  $g^2/4\pi$  has been altered to  $g^2/8\pi$ . Then, taking the ratio with (9.59),  $g^2$  drops out. If we compute the ratio, the value expected theoretically is

$$\frac{\Gamma(\tau^+ \rightarrow \pi^+ + \bar{\nu}_\tau)}{\Gamma(\pi^+ \rightarrow \mu^+ + \nu_\mu)} \approx 1.0 \times 10^4. \quad (13.117)$$

In order to investigate this ratio experimentally, we write

$$\begin{aligned} \frac{\Gamma(\tau^+ \rightarrow \pi^+ + \bar{\nu}_\tau)}{\Gamma(\pi^+ \rightarrow \mu^+ + \nu_\mu)} &= \frac{\Gamma(\tau^+ \rightarrow \text{all})}{\Gamma(\pi^+ \rightarrow \text{all})} r(\tau^+ \rightarrow \pi^+ + \nu_\tau) \\ &= \frac{\tau(\pi^+)}{\tau(\tau^+)} r(\tau^+ \rightarrow \pi^+ + \nu_\tau), \end{aligned} \quad (13.118)$$

where  $r(\tau^+ \rightarrow \pi^+ + \nu_\tau)$  stands for the branching ratio for  $\tau^+$  to decay as in (13.113), and  $\tau(\pi^+)$  and  $\tau(\tau^+)$  are the mean lifetimes of  $\pi^+$  and  $\tau^+$ , respectively:

$$\begin{aligned} \tau(\pi^+) &= (2.6030 \pm 0.0023) \times 10^{-8} \text{ s}, \\ \tau(\mu^+) &= (2.19714 \pm 0.00007) \times 10^{-6} \text{ s}, \\ \tau(\tau^+) &= (3.4 \pm 0.5) \times 10^{-13} \text{ s}. \end{aligned} \quad (13.119)$$

Inserting these values into (13.118) and equating it with (13.117), we obtain

$$r(\tau^+ \rightarrow \pi^+ + \nu_\tau) = (13 \pm 2)\% . \quad (13.120)$$

In the case of  $\pi^+$ , we have assumed that they almost all decay into  $\mu^+ + \nu_\mu$ . Equation (13.120) does not contradict the experimental value  $(10.7 \pm 1.6)\%$ . This confirms that the coupling constant  $g$  does not depend on the generation in the lepton-decay interactions of  $\pi$ .

## 13.4 Beta-Decay

The interaction term (13.114) given in the previous section is in fact phenomenological. We may consider that this interaction derives from a more fundamental one. Before the birth of gauge theory, the fundamental weak interaction theory was Fermi's, and the oldest and most typical example of such an interaction was the beta-decay of an atomic nucleus.

A theory of beta-decay, viz.,

$$n \rightarrow p + e^- + \bar{\nu}_e , \quad (13.121)$$

was proposed by Fermi in 1934 [13]. The interaction term was considered to be

$$\mathcal{H}_\beta = \sum_j C_j \bar{\psi}_p O_j \psi_n \cdot \bar{\psi}_e O_j \psi_\nu + \sum_j C'_j \bar{\psi}_p O_j \psi_n \cdot \bar{\psi}_e O_j \gamma_5 \psi_\nu + \text{h.c.} , \quad (13.122)$$

where the  $O_j$  are the Dirac matrices corresponding to the scalar, the vector, the tensor, the axial vector, and the pseudo-scalar (SVTAP), introduced in Sect. 3.9. From the point of view of the Pauli–Gürsey group introduced in Sect. 7.7, this is not general enough. A term with  $\psi_\nu$  replaced by  $C\bar{\psi}_\nu$  should be added to (13.122). However, we consider here only the interaction with the restricted form above. Therefore, we are actually considering only a subgroup of the Pauli–Gürsey group:

$$\mathcal{G}_1 \times \mathcal{G}_2 \in \mathcal{G}_F . \quad (13.123)$$

Phase transformations under  $\mathcal{G}_F$  are

$$\psi \rightarrow e^{i\alpha + i\beta\gamma_5} \psi , \quad \bar{\psi} \rightarrow \bar{\psi} e^{-i\alpha + i\beta\gamma_5} . \quad (13.124)$$

These transformations lead to the following transformations of the coupling constants:

$$C_j \rightarrow e^{i\alpha} (C_j \cos \beta + iC'_j \sin \beta) , \quad C'_j \rightarrow e^{i\alpha} (C'_j \cos \beta + iC_j \sin \beta) . \quad (13.125)$$

An invariant combination of the coupling constants under these transformations is

$$C_i^* C_j + C_i'^* C_j' . \tag{13.126}$$

Experimentally, only this combination is measured. It was Lee and Yang who introduced a term including  $C_j'$  into (13.122) and considered the possibility of non-conservation of parity [45]. That was in 1956. They pointed out that it was only by making the replacement

$$C_i^* C_j \rightarrow C_i^* C_j + C_i'^* C_j' \tag{13.127}$$

in several kinds of calculation in the previous theory without  $C_j'$ , that agreement could be obtained with previously obtained experimental data for beta-decay.

Furthermore, restricting to the transformation (13.124), fixed points exist for  $C_i$  and  $C_j'$  under the transformation (13.125):

$$\text{when } \alpha = -\beta , \quad C_j = C_j' , \quad \text{and when } \alpha = \beta , \quad C_j = -C_j' . \tag{13.128}$$

Fixed points imply that the interaction (13.122) is invariant under this restricted transformation, and as discussed in Sect. 7.7, this leads to invariance under the restricted transformation (7.146). As shown in Sect. 3.7, experiment implies that

$$C_j = C_j' . \tag{13.129}$$

Thus, (13.122) can be written in the form

$$\mathcal{H}_\beta = \sum_j C_j \bar{\psi}_p O_j \psi_n \cdot \bar{\psi}_e O_j (1 + \gamma_5) \psi_\nu + \text{h.c.} \tag{13.130}$$

As shown above, the neutrino is left-handed. Lee and Yang chose  $C_j$  and  $O_j$  as in Table 13.3.

We shall now determine what kinds of coupling constants exist. In beta-decay, there is one for the emission of  $e^-$  and one for the emission of  $e^+$ . We can then obtain the following equation for the dependence of the helicity  $\boldsymbol{\sigma} \cdot \mathbf{p}/p$  of the emitted  $e^\mp$  on the coupling constants:

$$h(e^\pm) = \mp \frac{|(1)|^2(|C_S|^2 - |C_V|^2) + |\langle \boldsymbol{\sigma} \rangle|^2(|C_T|^2 - |C_A|^2) v}{|(1)|^2(|C_S|^2 + |C_V|^2) + |\langle \boldsymbol{\sigma} \rangle|^2(|C_T|^2 + |C_A|^2) c} , \tag{13.131}$$

**Table 13.3** Choice of  $C_j$  and  $O_j$  made by Lee and Yang

Type	Scalar	Vector	Tensor	Axial vector	Pseudo-scalar
$C_j$	$C_S$	$C_V$	$C_T$	$C_A$	$C_P$
$O_j$	1	$\gamma_\mu$	$\frac{1}{\sqrt{2}}\sigma_{\mu\nu}$	$-i\gamma_\mu\gamma_5$	$\gamma_5$

where  $\langle \mathbf{1} \rangle$  and  $\langle \boldsymbol{\sigma} \rangle$  are the nuclear matrix elements of Fermi type and Gamow–Teller type. That is, assuming that  $\boldsymbol{\psi}$  is a non-relativistic Heisenberg operator,

$$\langle \mathbf{1} \rangle = \langle \mathbf{B} | \boldsymbol{\psi}_p^\dagger(0) \boldsymbol{\psi}_n(0) | \mathbf{A} \rangle, \quad \langle \boldsymbol{\sigma} \rangle = \langle \mathbf{B} | \boldsymbol{\psi}_p^\dagger(0) \boldsymbol{\sigma} \boldsymbol{\psi}_n(0) | \mathbf{A} \rangle. \quad (13.132)$$

In deriving (13.131), we have omitted the so-called *Fierz interference term* between the VA-type and the ST-type, and we have also omitted the Coulomb correction. We now prove this formula, assuming for simplicity that there is only a scalar-type interaction.

We start with the state  $|i\rangle$  and end up with the state  $S|i\rangle$ . Therefore,

$$S|i\rangle = \sum_n |n\rangle S_{ni}. \quad (13.133)$$

Depending on experimental conditions, we may or may not have to take a sum over all states  $|n\rangle$ . We write observable states as  $(n)$ . The expectation value of the physical quantity  $A$  in the final state is given by

$$\langle A \rangle = \frac{\sum_{(m)} \sum_{(n)} \langle i | T^\dagger | m \rangle \langle m | A | n \rangle \langle n | T | i \rangle}{\sum_n \langle i | T^\dagger | n \rangle \langle n | T | i \rangle}, \quad (13.134)$$

where  $S$  is replaced by  $T$  and common factors in the numerator and denominator have been eliminated. We now choose the eigenstates of  $A$  as the partially complete system  $(n)$ :

$$A|n\rangle = a_n|n\rangle. \quad (13.135)$$

If we choose  $A = \mathbf{p} \cdot \boldsymbol{\sigma}$  for the electron, then  $|n\rangle$  can be an eigenstate of the Dirac Hamiltonian  $H(p)$ . This is because

$$[\boldsymbol{\sigma} \cdot \mathbf{p}, H(p)] = 0. \quad (13.136)$$

In order to compute  $\langle A \rangle$ , we only have to make the following replacement for the Casimir operator of the electron in a molecule:

$$-i\mathbf{p} \cdot \boldsymbol{\gamma} + m \rightarrow (-i\mathbf{p} \cdot \boldsymbol{\gamma} + m)(\boldsymbol{\sigma} \cdot \mathbf{p}) = (\boldsymbol{\sigma} \cdot \mathbf{p})(-i\mathbf{p} \cdot \boldsymbol{\gamma} + m). \quad (13.137)$$

We thus start with

$$\mathcal{T}_{fi} \propto \bar{u}_e(p)(1 + \gamma_5)v_v(q)\langle \mathbf{1} \rangle. \quad (13.138)$$

Therefore,

$$\langle \boldsymbol{\sigma} \cdot \mathbf{p} \rangle = \frac{\sum \text{Tr}[(1 + \gamma_5)(-iq \cdot \boldsymbol{\gamma})(-i\mathbf{p} \cdot \boldsymbol{\gamma} + m)(\boldsymbol{\sigma} \cdot \mathbf{p})] \cdot |\langle \mathbf{1} \rangle|^2}{\sum \text{Tr}[(1 + \gamma_5)(-iq \cdot \boldsymbol{\gamma})(-i\mathbf{p} \cdot \boldsymbol{\gamma} + m)] \cdot |\langle \mathbf{1} \rangle|^2}, \quad (13.139)$$

where the four-momentum  $p$  of the electron has been fixed. The sum over the final states can for all practical purposes be replaced by the following phase space integral:

$$\sum \longrightarrow \int \frac{d^3q}{2q_0} \int \frac{d^3P_B}{2P_{B0}} \delta^4(P_f - P_i) \dots \quad (13.140)$$

Since an atomic nucleus  $B$  is heavy,  $P_{B0}$  can be treated as constant. Then,  $d^3P_B$  cancels  $\delta^3(P_f - P_i)$  for the momentum. Regarding  $\mathbf{q}$ , its direction becomes free only if energy is conserved, as in the case of scattering in a potential:

$$\sum \longrightarrow \int \frac{d^3q}{2q_0} \delta(E - p_0 - q_0) \dots \quad (13.141)$$

The energy can be approximated by the mass difference with the atomic nucleus, i.e.,  $E = P_{A0} - P_{B0} \approx M_A - M_B$ .

Regarding the traces,

$$\text{Tr}(\text{numerator}) = 4(q_0 p^2 - p_0 \mathbf{p} \cdot \mathbf{q}), \quad \text{Tr}(\text{denominator}) = 4(q_0 p_0 - \mathbf{p} \cdot \mathbf{q}). \quad (13.142)$$

Hence, carrying out the  $q$ -integral in polar coordinates, we have

$$\begin{aligned} \int \frac{d^3q}{2q_0} \delta(E - p_0 - q_0) (q_0 p^2 - p_0 \mathbf{p} \cdot \mathbf{q}) &= 2\pi \int q dq \delta(E - p_0 - q_0) q_0 p^2, \\ \int \frac{d^3q}{2q_0} \delta(E - p_0 - q_0) (q_0 p_0 - \mathbf{p} \cdot \mathbf{q}) &= 2\pi \int q dq \delta(E - p_0 - q_0) p_0 q_0. \end{aligned}$$

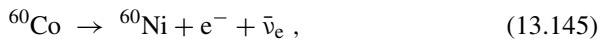
Taking into account the fact that  $q_0 = q$  and carrying out the  $q$ -integral, it turns out that  $q = q_0$  is replaced by  $E - p_0$ . Therefore, the ratio of the two terms yields

$$\langle \boldsymbol{\sigma} \cdot \mathbf{p} \rangle = \frac{q_0 p^2}{p_0 q_0} = p \frac{p}{p_0} = p \frac{v}{c}. \quad (13.143)$$

Thus, the helicity is

$$\langle h \rangle = \langle \boldsymbol{\sigma} \cdot \mathbf{p} / p \rangle = \frac{v}{c}. \quad (13.144)$$

A generalization of this derivation yields (13.131). The helicity of  $e^-$  was measured for



by Fraunfelder in [127, 128]. The result was

$$h(e^-) = -\left(\frac{v}{c}\right) \pm 10\% . \quad (13.146)$$

Comparing this with (13.131), we see that beta-decay is a VA combination. Thus, the interaction term is

$$\mathcal{H}_\beta = (C_V \bar{\psi}_p \gamma_\lambda \psi_n - C_A \bar{\psi}_p \gamma_\lambda \gamma_5 \psi_n) \bar{\psi}_e \gamma_\lambda (1 + \gamma_5) \psi_\nu + \text{h.c.} \quad (13.147)$$

Taking

$$C_V = \frac{G}{\sqrt{2}} , \quad C_A = \frac{G}{\sqrt{2}} x , \quad (13.148)$$

$G$  itself is determined from

$$^{14}\text{O} \rightarrow ^{14}\text{N}^* + e^+ + \nu_e , \quad (13.149)$$

corresponding to the transition  $0^+ \rightarrow 0^+$ , which is not affected by  $C_A$ . This is because  $(\mathbf{1})$  is obtained group-theoretically in this case, using the fact that  $^{14}\text{C}$ ,  $^{14}\text{N}^*$ , and  $^{14}\text{O}$  form an isospin triplet. We will mention an issue regarding this intermediate process in the next section.

In fact,  $x$  can be determined from the beta-decay of a polarized meson, for example:

$$G = \frac{1.01}{m_p^2} \times 10^{-5} , \quad x = -1.254 \pm 0.006 . \quad (13.150)$$

Since  $x$  is negative, we call this the V – A theory. Explicitly, the interaction term is

$$\mathcal{H}_\beta = \frac{G}{\sqrt{2}} \bar{\psi}_p \gamma_\lambda (\mathbf{1} + 1.25 \gamma_5) \psi_n \cdot \bar{\psi}_e \gamma_\lambda (\mathbf{1} + \gamma_5) \psi_\nu + \text{h.c.} \quad (13.151)$$

The factor of  $1/\sqrt{2}$  originates from the replacement in (13.127). For this reason, when using the old coupling constant  $G$ , we need to write it in this form.

## 13.5 Universality of the Fermi Interaction

We have already mentioned the universality of the lepton decay of the  $\pi$ -meson with interaction (13.114), since it does not depend on the lepton generation. The interaction term specified in (13.114) is phenomenological. The Fermi interaction discussed above can be considered to cover more phenomena and is therefore more

fundamental. We will discuss this idea step by step. So we begin this section by interpreting the universality of (13.114) from the point of view of the universality of the Fermi interaction. Equation (13.151) specifies the interaction between nucleons and leptons in the first generation, and we assume that the same form holds universally for leptons in all generations:

$$\mathcal{H}_F = \frac{G}{\sqrt{2}} \bar{\psi}_p \gamma_\lambda (\mathbf{1} + 1.25 \gamma_5) \psi_n \cdot \bar{\psi}_l \gamma_\lambda (\mathbf{1} + \gamma_5) \psi_{\nu_l} + \text{h.c.} \quad (13.152)$$

If we assume that  $\pi^+ \rightarrow \mu^+ + \nu_\mu$  occurs through this interaction, then

$$\begin{aligned} & \langle \mu^+, \nu_\mu | \mathcal{H}_F(x) | \pi^+ \rangle \\ &= 1.25 \frac{G}{\sqrt{2}} \langle \mathbf{0} | \bar{\psi}_n(x) \gamma_\lambda \gamma_5 \psi_p(x) | \pi^+ \rangle \langle \mu^+, \nu_\mu | \bar{\psi}_{\nu_\mu}(x) \gamma_\lambda (\mathbf{1} + \gamma_5) \psi_\mu(x) | \mathbf{0} \rangle. \end{aligned} \quad (13.153)$$

From its transformation property, we may write the first factor as

$$1.25 \frac{G}{\sqrt{2}} \langle \mathbf{0} | \bar{\psi}_n(x) \gamma_\lambda \gamma_5 \psi_p(x) | \pi^+ \rangle = -\frac{g}{m_\pi} p_\lambda \frac{e^{ip \cdot x}}{\sqrt{2 p_0 V}}, \quad (13.154)$$

where  $p$  is the four-momentum of  $\pi^+$ . Therefore, phenomenologically, as far as this process is concerned, we can make the following replacement:

$$1.25 \frac{G}{\sqrt{2}} \bar{\psi}_n \gamma_\lambda \gamma_5 \psi_p \rightarrow \frac{ig}{m_\pi} \partial_\lambda \varphi. \quad (13.155)$$

The fact that  $g$  does not depend on the generation in (13.114) can be interpreted as a consequence of the universality wherein  $G$  does not depend on the generation in (13.152). We can thus consider (13.114) as a phenomenological expression of (13.152) as far as this process is concerned. Let us go a little further with this universality.

As far as the lepton part goes, the universal form

$$\bar{\psi}_l \gamma_\lambda (\mathbf{1} + \gamma_5) \psi_{\nu_l} \quad (13.156)$$

appears in the Fermi interaction, but the nucleon part still has a complicated form. It may thus be better to consider the nucleon part as still being a phenomenological expression. Let us treat this part slightly more abstractly. From (7.30), we write down the isospin current density for the nucleon–meson system as

$$\hat{j}_\lambda = i \bar{\psi} \gamma_\lambda \hat{\tau} \psi - \hat{\varphi} \times \partial_\lambda \hat{\varphi}, \quad (13.157)$$

where we have considered nucleons and  $\pi$ -mesons as elementary particles here. As long as the theory preserves charge independence, the isospin current density



is always expressed in terms of the fundamental field operators as the Noether current density. However, the main properties of isospin do not depend on what field we choose as the fundamental field, being determined by group-theoretical considerations. This is similar in the scattering theory: the scattering amplitude for stable particles obeys the scattering formula regardless of whether the particles are elementary or composite.

The vector part of the nucleon term in the Fermi interaction is

$$i\bar{\psi}_p\gamma_\lambda\psi_n = i\left(\bar{\psi}\gamma_\lambda\frac{\tau}{2}\psi + i\bar{\psi}\gamma_\lambda\frac{\tau_2}{2}\psi\right). \quad (13.158)$$

Note that this is the nucleon part of  $j_{1\lambda} + ij_{2\lambda}$  in the isospin current density (13.157). To give the phenomenological vector part of the nucleon greater generality, we replace it with the isospin current density, which can be written without specifying the fundamental fields. Hence, we rewrite the vector part in (13.152) as

$$\mathcal{H}_V = -i\frac{G}{\sqrt{2}}(j_{1\lambda} + ij_{2\lambda})\bar{\psi}_l\gamma_\lambda(\mathbf{1} + \gamma_5)\psi_{l'} + \text{h.c.} \quad (13.159)$$

The isospin current density satisfies the conservation law, so Gell-Mann called this replacement the CVC hypothesis [49], where CVC is an abbreviation for the conserved vector current. Since this vector current can be used for all hadrons with fixed isospins, it becomes a truly universal expression. For example, consider the beta-decay of  $\pi^+$ :

$$\pi^+ \rightarrow \pi^0 + e^+ + \nu_e. \quad (13.160)$$

If the momentum is conserved, the matrix elements of the isospin are

$$\langle \mathbf{p} | I_1 + iI_2 | \mathbf{n} \rangle = 1, \quad \langle \pi^+ | I_1 + iI_2 | \pi^0 \rangle = -\sqrt{2}. \quad (13.161)$$

Regarding their densities when the momentum transfer is almost zero,

$$\langle \mathbf{p} | j_{10} + ij_{20} | \mathbf{n} \rangle = \frac{1}{V} \langle \mathbf{p} | I_1 + iI_2 | \mathbf{n} \rangle = \frac{1}{V}, \quad (13.162)$$

$$\langle \pi^+ | j_{10} + ij_{20} | \pi^0 \rangle = \frac{1}{V} \langle \pi^+ | I_1 + iI_2 | \pi^0 \rangle = -\frac{\sqrt{2}}{V}. \quad (13.163)$$

This result does not depend on whether nucleons and  $\pi$ -mesons are elementary or composite particles. Generalizing the above equation,

$$\langle \pi^+ | j_{1\lambda} + ij_{2\lambda} | \pi^0 \rangle = -\frac{\sqrt{2}}{V} \frac{q_\lambda}{q_0}, \quad (13.164)$$

where we assume that  $\pi^+$  and  $\pi^0$  have almost the same momentum  $q_\lambda$ , and neglect the difference in their momenta. From this, we can immediately compute the decay width of the beta-decay of  $\pi^+$ :

$$\frac{\Gamma(\pi^+ \rightarrow \pi^0 + e^+ + \nu_e)}{\Gamma(\pi^+ \rightarrow \mu^+ + \nu_\mu)} = 1.0 \times 10^{-8} . \quad (13.165)$$

This agrees extremely well with the experimental value of  $(1.033 \pm 0.034) \times 10^{-8}$ . Assuming that the CVC hypothesis is correct, it thus turns out that the interaction (13.159) is extremely general and can be used for all hadrons.

So far in the discussion, we have only appealed to the properties of the isospin, and in particular the commutation relation

$$[I_i, I_j] = i\epsilon_{ijk} I_k . \quad (13.166)$$

This commutation relation therefore characterizes the universality. It is thus tempting to consider the same thing for the axial-vector part of the interaction. We thus make the substitution

$$1.25 \frac{G}{\sqrt{2}} i \bar{\psi}_p \gamma_\lambda \gamma_5 \psi_n \rightarrow \frac{G}{\sqrt{2}} (j_{1\lambda}^5 + i j_{2\lambda}^5) . \quad (13.167)$$

The factor of 1.25 can be considered as phenomenological, i.e., it is given by

$$\langle \mathbf{p} | j_{1\lambda}^5 + i j_{2\lambda}^5 | \mathbf{n} \rangle = 1.25 \frac{i}{V} \bar{u}(\mathbf{p}) \gamma_\lambda \gamma_5 u(\mathbf{n}) . \quad (13.168)$$

As will soon be shown, this is a calculable factor. We can now define the pseudo-scalar isospin by

$$\hat{I}^5 = \int d^3x \hat{j}_0^5 . \quad (13.169)$$

Since  $\hat{I}^5$  is itself an isospin vector, the commutation relation with  $\hat{I}$  is given by

$$[I_i, \hat{I}_j^5] = i\epsilon_{ijk} \hat{I}_k^5 . \quad (13.170)$$

The remaining problems here are the normalization of  $\hat{I}^5$  and, in association with that, the closure of the commutation relations between  $\hat{I}$  and  $\hat{I}^5$ . Gell-Mann assumed the following equation:

$$[I_i^5, I_j^5] = i\epsilon_{ijk} I_k . \quad (13.171)$$

The algebra closed in the way just described is called a *current algebra* [52, 53].

To test this algebra, starting with (13.171), we have to see whether we can prove that  $-C_A/C_V = 1.25$ . To do this, we use the partially conserved axial-vector current (PCAC) hypothesis [129], according to which  $\partial_\lambda \hat{j}_\lambda^5$  can express the  $\pi$ -meson field and the dispersion formula. The result is the Adler–Weisberger formula [54–56]:

$$\left(\frac{C_V}{C_A}\right)^2 = 1 + \frac{2}{\pi} \left(\frac{m_p}{G}\right)^2 \int_{m_\pi}^\infty \frac{k \, d\omega}{\omega^2} [\sigma_-(\omega) - \sigma_+(\omega)], \quad (13.172)$$

where  $\omega$  and  $k$  are the incident energy and momentum of  $\pi$  in the laboratory frame, and  $\sigma_-(\omega)$  and  $\sigma_+(\omega)$  are the total cross-sections for the following processes:

$$\sigma_-(\omega) : \pi^- + p \rightarrow \text{all}, \quad \sigma_+(\omega) : \pi^+ + p \rightarrow \text{all}, \quad (13.173)$$

$G$  is the coupling constant appearing in the phenomenological interaction between nucleons and  $\pi$ -mesons given in (6.96). Its approximate value is  $G^2/4\pi \approx 15$ . Inserting an experimental value into the right-hand side of (13.172), the result almost coincides with  $C_A/C_V = -1.25$ , and this therefore supports the current algebra hypothesis. For a proof of (13.172), the reader is referred to my book *Fields and Particles* [2].

So far, regarding the Fermi interaction, the following universality has been established. First, for the lepton part, the interactions are common to all generations. Next, for the hadron part, it can be expressed by the current density satisfying the current algebra hypothesis. Moreover, it is easy to check that exactly the same commutation relations hold for the lepton part as for the hadron part, i.e., taking

$$i\bar{\psi}_l \gamma_\lambda (1 + \gamma_5) \psi_{\nu_l} \rightarrow j'_{1\lambda} - i j'_{2\lambda}, \quad (13.174)$$

and writing spatial integrals of  $j'_{10}$  and  $j'_{20}$  as  $I'_1$  and  $I'_2$ , respectively, we find that  $I'_1$  and  $I'_2$  satisfy the same commutation relations as  $I_1 + I_1^5$  and  $I_2 + I_2^5$ . Thus, algebraically, the lepton part and the hadron part have the universal properties.

In order to investigate the Fermi interaction among leptons, such as in the reaction

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu, \quad (13.175)$$

we assume an interaction with the same form:

$$\mathcal{H}_{\mu e} = \frac{G_F}{\sqrt{2}} [\bar{\psi}_{\nu_e} \gamma_\lambda (1 + \gamma_5) \psi_e] [\bar{\psi}_\mu \gamma_\lambda (1 + \gamma_5) \psi_{\nu_\mu}] + \text{h.c.} \quad (13.176)$$

Since it was explained in detail how to compute this decay width in Sect. 9.5, we only show the result:

$$\Gamma(\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu) = \frac{1}{192\pi^2} G_F^2 m_\mu^5. \quad (13.177)$$

This coupling constant does not have a correction from the strong interaction, so it can be determined from the lifetime and the mass of the muon:

$$G_F = \frac{1.03}{m_p^2} \times 10^{-5} . \quad (13.178)$$

Similarly, in the beta-decay of the neutron,

$$\Gamma(n \rightarrow p + e^- + \bar{\nu}_e) = \frac{0.94}{60\pi^3} (|C_V|^2 + 3|C_A|^2) \Delta^5 , \quad (13.179)$$

where  $\Delta = m_\mu - m_p$ . Comparing (13.178) with (13.150), there is a slight difference between  $G$  and  $G_F$ . Cabibbo provided an interpretation of this difference. The current algebra is based on the isospin group  $SU(2)$ , and the weak interaction wherein a hadron decays into a lepton pair has a small component which changes the strangeness. This extends the current algebra to  $SU(3)$ . Cabibbo's interpretation was that the hadron part of the Fermi interaction is the isospin density but slightly altered by an  $SU(3)$  rotation [57]. The angle of rotation is known as the *Cabibbo angle*. Writing it as  $\theta_c$ ,

$$G = G_F \cos \theta_c . \quad (13.180)$$

Cabibbo's theory will be discussed in detail in the next section. Here we discuss the quark model. Expressing the commutation relation in terms of field operators at equal times  $x_0 = y_0$ , we can use the density to write

$$[\mathbf{J}_{i0}^5(x), \mathbf{J}_{j0}^5(y)] = i\epsilon_{ijk} \mathbf{J}_{k0}(x) \delta^3(x - y) + \text{GIS term} , \quad (13.181)$$

where the GIS term is the term which vanishes, like the Goto–Imamura–Schwinger term, when we carry out the spatial integral. Assuming that the fields of the  $\pi$ -meson and the nucleon are fundamental fields, the first term on the right-hand side,  $\mathbf{J}_{k\lambda}(x)$ , or  $\hat{\mathbf{j}}_\lambda(x)$ , is given by (13.157). It remains to express  $\hat{\mathbf{j}}_\lambda^5$ . We shall not give an explicit expression here, as  $\hat{\mathbf{j}}_\lambda^5$  satisfying (13.181) is highly non-linear and hence complicated. The dynamics in a system with the symmetry corresponding to such a non-linear current expression is called *chiral dynamics*. This gives a phenomenological representation of the current algebra, but it would be too complicated to consider the fundamental theory. It would be utterly impossible to implement the renormalization. We thus turn to the expression in terms of quarks.

## 13.6 Quark Model in Weak Interactions

In Sect. 13.2, we introduced three types of quarks, u, d, and s. These quarks were invented as a mathematical tool and as carriers of the fundamental representation, rather than as real particles. However, once we recognize that the current which

satisfies the current algebra exists in the weak interaction as a physical entity, the way we express the current density in terms of field operators becomes a real physical problem, i.e., a problem of what the fundamental fields are in real life.

Denoting the quark field with three components u, d, and s as  $q$ , and using the Gell-Mann matrices  $\lambda_i$  which mix these three components, we introduce the current density

$$\mathcal{F}_{j\mu} = i\bar{q}\gamma_\mu \frac{\lambda_j}{2} q, \quad \mathcal{F}_{j\mu}^5 = i\bar{q}\gamma_\mu\gamma_5 \frac{\lambda_j}{2} q. \quad (13.182)$$

We also define the quantities

$$F_j = \int d^3x \mathcal{F}_{j0}, \quad F_j^5 = \int d^3x \mathcal{F}_{j0}^5. \quad (13.183)$$

These are not necessarily conserved. Their equal-time commutation relations are

$$[F_i, F_j] = if_{ijk} F_k, \quad [F_i, F_j^5] = if_{ijk} F_k^5, \quad [F_i^5, F_j^5] = if_{ijk} F_k. \quad (13.184)$$

Thus, it turns out that, using quarks, it is a straightforward matter to express the current algebra.

Here we have constructed the  $SU(3)$  current algebra, which has  $SU(2)$  as a subalgebra. To understand the algebraic structure of (13.184) in detail, we build the combination

$$F_i^{(\pm)} = \frac{1}{2}(F_i \pm F_i^5), \quad (13.185)$$

which yields the commutation relations

$$[F_i^{(+)}, F_j^{(-)}] = 0, \quad [F_i^{(+)}, F_j^{(+)}] = if_{ijk} F_k^{(+)}, \quad [F_i^{(-)}, F_j^{(-)}] = if_{ijk} F_k^{(-)}. \quad (13.186)$$

This algebraic structure is thus a direct product of the algebras  $SU(3)_L$  and  $SU(3)_R$ , generated by  $F^{(+)}$  corresponding to the left-handed current and  $F^{(-)}$  corresponding to the right-handed current, respectively:

$$SU(3)_L \times SU(3)_R. \quad (13.187)$$

As described above, the quark model is the simplest way to satisfy the current algebra and the best expression from the point of view of high-energy behaviour and renormalizability. Here, we no longer consider the quark model as a simple mathematical tool, but take the quarks to be real physical entities. Cabibbo defined

the current carrying electric charge which appeared in the Fermi interaction in [57] by

$$J_\mu = [\mathcal{F}_{1\mu}^{(\mu)} - i\mathcal{F}_{2\mu}^{(+)}] \cos \theta_C + [\mathcal{F}_{4\mu}^{(+)} - i\mathcal{F}_{5\mu}^{(+)}] \sin \theta_C + \sum_l i\bar{\psi}_l \gamma_\mu (1 + \gamma_5) \psi_{\nu_l}, \quad (13.188)$$

and wrote the universal Fermi interaction as

$$\mathcal{H}_F = \frac{G_F}{\sqrt{2}} J_\mu^+ J_\mu. \quad (13.189)$$

From these, for instance, the beta-decay interaction, which does not change the strangeness, is

$$\mathcal{H}_\beta = \frac{G_F}{\sqrt{2}} \cos \theta_C [\mathcal{F}_{1\mu}^{(+)} + i\mathcal{F}_{2\mu}^{(+)}] i\bar{\psi}_e \gamma_\mu (\mathbf{1} + \gamma_5) \psi_{\nu_e} + \text{h.c.} \quad (13.190)$$

Taking into account (13.180), this coincides with (13.152). In addition, it turns out that the 0-components of  $J_\mu$  defined by (13.188) and those of its Hermitian conjugate  $J_\mu^\dagger$  satisfy the algebra  $SU(2)_L$ .

The interaction (13.189) consists of the charged current, and the neutral current is missing. However, this will be added later, in the gauge theory. The quantum number discriminating between the three types of quarks u, d, and s is called *flavour*. Isospin and strangeness are examples of flavours. In fact, there is further discrimination between quarks, known as *colour*. This will be discussed in the next section.

To explain contributions changing the strangeness in the weak interaction, Cabibbo suggested that  $d'$  and  $s'$  might differ from d and s as mass eigenstates, rewriting the hadron part of the Fermi interaction as

$$\bar{\psi}_p \gamma_\lambda (\mathbf{1} + 1.25\gamma_5) \psi_n \rightarrow \bar{u} \gamma_\lambda (\mathbf{1} + \gamma_5) d'. \quad (13.191)$$

Since both d and s have electric charge  $-e/3$ , he proposed the transformation

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos \theta_C & \sin \theta_C \\ -\sin \theta_C & \cos \theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \quad (13.192)$$

This is the origin of the form (13.188). In 1970, Glashow et al. introduced a fourth quark called charm, denoted by c. Assuming that it carried an electric charge of  $+2e/3$ , this completed a second generation of quarks [130]:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}. \quad (13.193)$$

Then, in order to explain the violation of CP invariance, Kobayashi and Maskawa predicted a third generation of quarks, replacing (13.193) by [131]

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix}. \quad (13.194)$$

The quarks  $t$  and  $b$  are called the *top quark* and the *bottom quark*, respectively. Extending (13.192), they introduced the mixing matrix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} c_1 & s_1 c_2 & s_1 s_2 \\ -s_1 c_3 & c_1 c_2 c_3 - s_2 s_3 e^{-i\delta} & c_1 s_2 c_3 + c_2 s_3 e^{-i\delta} \\ -s_1 s_3 & c_1 c_2 s_3 + s_2 c_3 e^{-i\delta} & c_1 s_2 s_3 - c_2 c_3 e^{-i\delta} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (13.195)$$

This is called the *Kobayashi–Maskawa matrix*. The factor of  $e^{-i\delta}$  leads to time reversal or CP violation, and accordingly, the third-generation quarks are needed here. Moreover,

$$s_i = \sin \theta_i, \quad c_i = \cos \theta_i, \quad (13.196)$$

where  $\theta_i$  ( $i = 1, 2, 3$ ) are three angles, extending the Cabibbo angle.

We have introduced the quark model by considering the weak interactions, but what about the strong interactions?

### 13.7 Quark Model in Strong Interactions

In Sect. 13.2, we discussed the classification of hadrons, the mass formula, the magnetic moment, and so on, based on the  $SU(3)$  algebra. In this section, we use the quark model to show that these results can be further improved.

The most important thing here is that quarks have, besides flavour, another degree of freedom called *colour*. Let us see why a new degree of freedom is needed.

The quark model is supported by the following group theoretical considerations. Since each meson is a bound state  $q\bar{q}$ , for flavour  $SU(3)$ ,

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}, \quad (13.197)$$

which means that a singlet and an octet are possible. There is experimental support for this, namely, both the singlet and the octet exist, and there is nothing else. Each baryon has the form  $qqq$  and we have

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{10}. \quad (13.198)$$

Experimentally, it is precisely these multiplets that have been detected. The quarks forming the basic triplet have not been detected. The only thing that has been detected experimentally are states belonging to  $\mathbf{8}$ ,  $\mathbf{8} \otimes \mathbf{8}$ , and so on. This is what the eightfold way means.

Now let us consider a member  $\Delta^{++}$  of the decuplet. This particle is a bound state of three u quarks, i.e., uuu, and its spin is  $3/2$ . Assuming that the orbital angular momenta are each in the S-state, it can be considered that the three  $1/2$ -spins are parallel and the composite spin is  $3/2$ . If so, both the spatial wave function and the spin wave function are symmetric, which leads to a contradiction, since the wave function represents three identical fermions. Thus, there must be another, unknown degree of freedom. The wave function here should be completely anti-symmetric with respect to this degree of freedom, which we shall call colour for the time being. We assume the existence of three kinds of colour states 1, 2, and 3. We note, however, that this colour degree of freedom has never been observed at the level of hadrons. Assuming that hadrons do have an extra degree of freedom, this yields several contradictions, unless we consider that hadrons are the singlets for this new degree of freedom.

In this case, apart from the spatial part, the wave function of  $\Delta^{++}$  may be written as

$$\epsilon_{\alpha\beta\gamma} u_{\uparrow}^{\alpha} u_{\uparrow}^{\beta} u_{\uparrow}^{\gamma}, \quad (13.199)$$

where  $u$  stands for the flavour,  $\uparrow$  indicates spin up, and  $\alpha$ ,  $\beta$ , and  $\gamma$  indicate the colour degree of freedom. When we carry out an  $SU(3)$  transformation on the three kinds of colour state, (13.199) is invariant, so this is a colour singlet. All hadrons will thus be invariant under this colour  $SU(3)$  transformation. This means that the quark model must be a theory which is invariant under the  $SU(3)$  group.

### 13.7.1 Mass Formula

We will show that we can improve the mass formula in Sect. 13.2 by using the quark model. To do so, we first express the hadrons as quark bound states:

$$\begin{aligned} \pi^+ &\sim u\bar{d}, & \pi^- &\sim d\bar{u}, & K^+ &\sim u\bar{s}, & K^0 &\sim d\bar{s}, \\ p &\sim uud, & n &\sim udd, & \Lambda &\sim uds, & \Sigma^+ &\sim uus, \\ \Delta^{++} &\sim uuu, & \Omega^- &\sim sss, & & \text{etc.} \end{aligned}$$

The quarks have the following charges:

$$Q_u = \frac{2}{3}e, \quad Q_d = Q_s = -\frac{1}{3}e. \quad (13.200)$$



In some cases, we have to consider mixed states in order to distinguish the flavour singlet from the flavour octet:

$$\eta_1 \sim \frac{1}{\sqrt{3}}(u\bar{u} + d\bar{d} + s\bar{s}), \quad \eta_8 \sim \frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} - 2s\bar{s}). \quad (13.201)$$

A rough formula for the mass can be obtained by taking a sum over the quark masses. The following formulas can be derived from both the quark model and the previous  $SU(3)$  method. Particle names stand for masses here:

$$\begin{aligned} \frac{K}{496 \text{ MeV}} &= \frac{(\pi + 3\eta)/4}{446 \text{ MeV}}, & \frac{\rho}{776 \text{ MeV}} &= \frac{\omega}{783 \text{ MeV}}, & \frac{\phi - K^*}{127 \text{ MeV}} &= \frac{K^* - \rho}{116 \text{ MeV}}, \\ \frac{\Sigma^* - \Delta}{152 \text{ MeV}} &= \frac{\Xi^* - \Sigma^*}{150 \text{ MeV}} = \frac{\Omega^- - \Xi^*}{139 \text{ MeV}}, & \frac{(N + \Xi)/2}{1129 \text{ MeV}} &= \frac{(\Sigma + 3\Lambda)/4}{1135 \text{ MeV}}. \end{aligned} \quad (13.202)$$

In the quark model, we can include further corrections from the interaction. However, to do this, we will have to borrow a result from the gauge theory discussed later. The short-range potential among  $q\bar{q}$  or  $qq$  is given as follows, assuming that the composite system belongs to the  $\mathbf{1}$ -colour representation or the  $\bar{\mathbf{3}}$ -colour representation:

$$V_{q\bar{q}}(\mathbf{1}) = -\frac{4}{3}\alpha_s/r, \quad V_{qq}(\bar{\mathbf{3}}) = -\frac{2}{3}\alpha_s/r, \quad (13.203)$$

where  $\alpha_s$  is a coupling constant characterizing the interaction among particles carrying colour. Its expectation value  $\Delta E$  is given up to a multiplicative constant by

$$\Delta E \propto \frac{\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2}{m_1 m_2}. \quad (13.204)$$

We can thus add a correction term to the equation wherein the mass  $M$  of the composite system is considered to be a sum of masses  $m_i$  of the components:

$$M(q\bar{q}) = m_1 + m_2 + a(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2)/m_1 m_2, \quad (13.205)$$

$$M(qqq) = \sum_{i=1}^3 + a' \sum_{i < j} (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2)/m_i m_j. \quad (13.206)$$

These can be computed if the spin wave functions of the hadrons are known. The results are shown in Tables 13.4 and 13.5. Here, both the meson case and the baryon case involve three parameters, but given the simplicity of the formula, it cannot be helped that the quarks have to be attributed slightly different masses to fit the mesons and to fit the baryons. Even so, the qualitative agreement is good.

**Table 13.4** Meson masses in MeV. Choice of parameters:  $m_u = m_d = 310$  MeV,  $m_s = 483$  MeV,  $a/m_u^2 = 160$  MeV

Meson	$\pi(138)$	$K(496)$	$\eta(549)$	$\rho(776)$	$\omega(783)$	$K^*(892)$	$\phi(1020)$
Mass	140	485	559	780	780	896	1032

**Table 13.5** Baryon masses in MeV. Choice of parameters:  $m_u = m_d = 363$  MeV,  $m_s = 538$  MeV,  $a/m_u^2 = 50$  MeV

Baryon	$N(939)$	$\Lambda(1116)$	$\Sigma(1193)$	$\Xi(1318)$	$\Delta(1232)$	$\Sigma^*(1384)$	$\Xi^*(1533)$	$\Omega(1672)$
Mass	939	1114	1179	1327	1239	1381	1529	1682

### 13.7.2 Magnetic Moments

Just as we have done for mass, we can use the spin wave function to express the magnetic moments of baryons by

$$\mu_i = \frac{Q_i}{2m_i}, \quad i = u, d, s. \quad (13.207)$$

For instance, the wave function of the proton is

$$|p \uparrow\rangle = \sqrt{\frac{2}{3}}|uu, S_z = 1\rangle \times |d \downarrow\rangle - \sqrt{\frac{1}{3}}|uu, S_z = 0\rangle \times |d \uparrow\rangle, \quad (13.208)$$

where the symbols have the obvious meaning. The magnetic moment can then be computed from

$$\mu(p) = \frac{2}{3}(2\mu_u) + \left(-\frac{2}{3} + \frac{1}{3}\right)\mu_d = \frac{4}{3}\mu_u - \frac{1}{3}\mu_d. \quad (13.209)$$

We can express the magnetic moments in units of the nuclear magneton, as in Sect. 13.2. In baryons, the values of the  $\mu_i$  in (13.207) are

$$\mu_u = 1.863 \text{ n.m.}, \quad \mu_d = -0.931 \text{ n.m.}, \quad \mu_s = -0.583 \text{ n.m.} \quad (13.210)$$

The magnetic moments of the baryons are shown in Table 13.6. In the table, the magnetic moments of the quarks in (13.210) are expressed directly by the quark symbols u, d, and s. Apart from  $\Xi^-$ , where there is a large error, the agreement with experiment is much improved over Table 13.2.

So far we have discussed three types of quark. It took much longer for other kinds of quark to be “detected”. This is because c and b are much more massive than u, d, and s. As will be discussed later, an isolated quark cannot be observed, so what is observed is always a hadron. In 1974, Ting and Richter independently found a vector particle with a mass of about 3100 MeV, which they called  $J$  and

**Table 13.6** Baryon magnetic moments

Baryon	Expression for magnetic moment	Theoretical value		Experimental value
p	$\frac{4}{3}u - \frac{1}{3}d$	2	0.73	2.793
n	$\frac{4}{3}d - \frac{1}{3}u$	-1	0.86	-1.913
$\Lambda^0$	s	-0	0.58	$-0.613 \pm 0.004$
$\Sigma^0 \rightarrow \Lambda^0$	$\frac{1}{\sqrt{3}}(d - u)$	-1	0.61	$-\left(1.82^{+0.25}_{-0.18}\right)$
$\Sigma^+$	$\frac{4}{3}u - \frac{1}{3}s$	2	0.68	$2.379 \pm 0.020$
$\Sigma^0$	$\frac{4}{3}(u + d) - \frac{1}{3}s$	0	0.82	
$\Sigma^-$	$\frac{4}{3}d - \frac{1}{3}s$	-1	0.05	$-1.10 \pm 0.50$
$\Xi^0$	$\frac{4}{3}s - \frac{1}{3}u$	-1	0.4	$-1.250 \pm 0.014$
$\Xi^-$	$\frac{4}{3}s - \frac{1}{3}d$	-0	0.47	$-1.85 \pm 0.75$

$\psi$ , respectively [132, 133]. This was interpreted as the bound state  $c\bar{c}$ , where  $c$  was identified with the  $c$  quark introduced by Glashow et al. in 1970. Regarding the fourth quark, its existence was considered by Maki and Hara from a different point of view in about 1963 [134, 135]. Then a particle called  $\Upsilon$  with a mass of about 9460 MeV was detected by Lederman and Yamanouchi [136]. It was interpreted as the bound state  $b\bar{b}$ .

Hence, it gradually became clear that there are also three generations of quarks. The next problem was therefore to find the  $t$  quark. The detection of this particle is (now we should use ‘was’ instead of ‘is’) one of the aims of the accelerator TRISTAN, constructed in Tsukuba. Moreover, bound states of these heavy quarks have many excited states, and they have been explained by assuming the potential

$$V(r) = -\frac{4}{3} \frac{\alpha_s}{r} + kr, \quad (13.211)$$

which has proved extremely successful. It is assumed that, the more massive the quark, the better the approximation provided by such a non-relativistic computation.

## 13.8 Parton Model

We have been discussing the quark model, according to which each hadron is a bound state of quarks. A key issue is the way the strong interaction glues quarks together. The *parton model* characterizes this property of the strong interaction. If

we translate this literally, it should be called the *partial-particle model*. In 1969, Feynman assumed that “the nucleon consists of almost free point particles, called partons.” He dubbed this the parton model [137]. Thus, for instance, if the four-momentum of a proton is  $P$  and the four-momenta of the partons making it up are  $p_i$ , then we require

$$p_i \approx x_i P, \quad 1 \geq x_i \geq 0, \quad (13.212)$$

because without such a proportionality relation, almost free partons will fly away on their own and the proton will fall to pieces.

In order to make this idea more explicit, assuming that the parton is in fact a quark, the proton has the composition  $uud$ . If the probabilities for the  $u$  quark and  $d$  quark to exist in the region  $(x, x + \Delta x)$  are  $u(x)\Delta x$  and  $d(x)\Delta x$ , respectively, then we must have

$$\int_0^1 dx u(x) = 2, \quad \int_0^1 dx d(x) = 1. \quad (13.213)$$

We therefore compute the structure function for the inelastic scattering of electrons discussed in Sect. 9.3 using this model. To ensure consistency with (13.212), if the proton mass is denoted by  $M$ , we take the parton mass to be

$$m_i = x_i M. \quad (13.214)$$

We must then compute contributions corresponding to the structure function for each parton. If we write a contribution from one parton as  $w_{\mu\nu}$ , we have

$$2\pi e^2 w_{\mu\nu} = \int d^4 x \frac{p_0}{m} \langle p | j_\mu(x) j_\nu(0) | p \rangle e^{-iQ \cdot x}, \quad (13.215)$$

where we have dropped the subscript  $i$ . Actually, we have to use a matrix element in the Heisenberg picture, but since we use free quarks here, we compute in the interaction picture. Averaging over possible spins of the quark, the equation above becomes

$$\begin{aligned} 2\pi e^2 w_{\mu\nu} &= -e_q^2 \int \frac{d^3 p'}{(2\pi)^3} \frac{p_0}{m} \frac{1}{2} \text{Tr} \left( \gamma_\mu \frac{-i p' \cdot \gamma + m}{2p'_0} \gamma_\nu \frac{-i \cdot \gamma + m}{2p_0} \right) (2\pi)^4 \delta^4(p' - p - Q) \\ &= -2\pi e_q^2 \delta(p'^2 + m^2) \frac{1}{4m} \text{Tr} \left[ \gamma_\mu \left[ -i(P + Q) \cdot \gamma + m \right] \gamma_\nu (-i p \cdot \gamma + m) \right], \end{aligned}$$

where we have assumed that the parton absorbs the momentum  $Q$ , and as a result it has the momentum  $p' = p + Q$ . Therefore,

$$w_{\mu\nu} = - \left( \frac{e_q}{e} \right)^2 \delta[(p + Q)^2 + m^2] \frac{1}{m} (2p_\mu p_\nu + p_\mu Q_\nu + p_\nu Q_\mu - \delta_{\mu\nu} p \cdot Q). \quad (13.216)$$

So, using  $p^2 + m^2 = 0$ ,

$$w_{\mu\nu} = \left(\frac{e_q}{e}\right)^2 \delta(2p \cdot Q + Q^2) \left[ \left( \delta_{\mu\nu} - \frac{Q_\mu Q_\nu}{Q^2} \right) \nu + \left( P_\mu - \frac{P \cdot Q}{Q^2} Q_\mu \right) \left( P_\nu - \frac{P \cdot Q}{Q^2} Q_\nu \right) \frac{2xM}{M^2} \right], \quad (13.217)$$

where, as in Chap. 9,  $\nu$  has been defined by

$$p \cdot Q = xP \cdot Q = -xM\nu. \quad (13.218)$$

Using (9.82), we can now obtain the structure functions of the parton:

$$w_1 = \left(\frac{e_q}{e}\right)^2 \delta(2p \cdot Q + Q^2) \nu = \nu \left(\frac{e_q}{e}\right)^2 \delta(Q^2 - 2M\nu x), \quad (13.219)$$

$$w_2 = \left(\frac{e_q}{e}\right)^2 2xM \delta(2p \cdot Q + Q^2) = 2xM \left(\frac{e_q}{e}\right)^2 \delta(Q^2 - 2M\nu x). \quad (13.220)$$

To derive the structure functions of the proton from these, we have to multiply by  $u(x)$  and  $d(x)$  for  $q = u$  and  $q = d$ , respectively, and integrate with respect to  $x$ . If we take

$$P(x) = \left(\frac{e_u}{e}\right)^2 u(x) + \left(\frac{e_d}{e}\right)^2 d(x) = \frac{4}{9}u(x) + \frac{1}{9}d(x), \quad (13.221)$$

then

$$F_1 = MW_1 = \int dx P(x) \delta\left(\frac{Q^2}{M\nu} - 2x\right) = \frac{1}{2}P\left(\frac{Q^2}{2M\nu}\right), \quad (13.222)$$

$$F_2 = \nu W_2 = \int dx P(x)x \delta\left(\frac{Q^2}{2M\nu} - x\right) = \frac{Q^2}{2M\nu} P\left(\frac{Q^2}{2M\nu}\right). \quad (13.223)$$

This establishes the Bjorken scaling law [105], according to which, in the limits  $Q^2 \rightarrow \infty$  and  $M\nu \rightarrow \infty$ ,  $MW_1$  and  $\nu W_2$  become functions of  $Q^2/2M\nu = x$  alone. Although in the computation above we have not taken these limits explicitly, in the free parton picture we are assuming these limits implicitly because we neglect the binding energy. In this limit, inelastic scattering is called *deep-inelastic scattering*. We thus conclude that the parton model satisfies Bjorken's scaling law for deep inelastic scattering. Moreover, since we have assumed that the spin of the parton is 1/2 in the computation above, it follows that

$$F_2(x) = 2x F_1. \quad (13.224)$$

This relation, which agrees with experiment, is known as the Callan–Gross relation [138].

If the spin of the parton is zero, we may ask how the relation above changes. So, assuming that the parton is a scalar, let us compute  $w_{\mu\nu}$  :

$$\begin{aligned} 2\pi e^2 w_{\mu\nu} &= 2\pi e_p^2 \int \frac{d^3 p'}{2p'_0} \frac{1}{2m} (p + p')_\mu (p + p')_\nu \delta^4(p' - p - Q) \\ &= 2\pi e_p^2 \delta(2p \cdot Q + Q^2) (2p + Q)_\mu (2p + Q)_\nu / m , \end{aligned} \quad (13.225)$$

where  $e_p$  is the electric charge of the parton. In this case,

$$F_1(x) = 0 , \quad F_2(x) = x P(x) . \quad (13.226)$$

In other words,  $F_1$  vanishes. This contradicts the experiment. Hence, it is impossible to construct the proton using only scalar particles, and the parton must have spin 1/2. Note that, in the calculation above, the interference term has been dropped, but this may be justified to some extent because of the large momentum transfer.

To conclude this section, we can assert that “quarks in the hadron are almost free.” In other words, the interaction among quarks becomes weak at short range. If this holds for the strong interaction in general, this is a remarkable feature of the strong interaction which clearly restricts its structure.

If we consider the same thing in momentum space, it turns out that, in the strong interaction, the probability of a process involving a large momentum transfer is very small. As will be discussed later, only a non-Abelian gauge interaction can satisfy this condition.

# Chapter 14

## What Is Gauge Theory?



So far we have introduced a variety of interactions among elementary particles. If we arrange them in order of decreasing strength, we find the strong interaction, the electromagnetic interaction, the weak interaction, and finally the gravitational interaction. Typical in each category are Yukawa's nucleon–meson interaction, the interaction between the electron and the electromagnetic field, the Fermi interaction, and so on. For elementary particles with their very small masses, gravity is not generally considered in elementary particle theory, because it is many orders of magnitude weaker than the others.

These interactions often have various kinds of symmetry. For instance, the nucleon–meson interaction is invariant under rotations in isospin space, and as a consequence, there is a conserved quantity, namely isospin. It is thus natural to ask whether there is a general principle for determining the forms of different interactions and whether there exists a relationship between the form of an interaction and the symmetries of the theory.

The Coulomb potential and the Newtonian potential, already known to classical physics, are both inversely proportional to the distance. Their strengths are proportional to the product of the particles' electric charges and the product of the bodies' masses, respectively. Both the electric charge and the mass are conserved quantities in non-relativistic theories. If we consider that such potentials emerge from fields, then we can say that the origins of these fields are the current densities corresponding to conserved quantities. We call these fields and interactions *gauge fields* and *gauge interactions*. It turns out that in this sort of interaction the gauge field and the conserved quantity, or the symmetry inducing it, are in a one-to-one correspondence, which gives one solution to the question mentioned above.

The best-known example of a gauge field is the electromagnetic field. The origin of the electromagnetic field is a four-dimensional electric current. Its representation can be determined immediately by designating what particles are fundamental. So far we have written the Lagrangian by treating hadrons and leptons as fundamental particles, but it is also easy to write down the Lagrangian by considering quarks

and leptons as fundamental. This involves the same procedure as when we take the electric current in terms of hadrons and rewrite it in terms of quarks. Moreover, when two hadrons are far enough apart, the Coulomb potential depends only on the electric charges of the hadrons and not on whether the hadrons are fundamental particles or composite particles. This is because the electric charge is an additive quantum number.

Yang and Mills proposed in 1954 the existence of a gauge field corresponding to the conserved isospin [139]. Since the isospin symmetry is described by a non-Abelian algebra, this is called a *non-Abelian gauge field*. In 1956, Utiyama introduced the non-Abelian gauge field corresponding to a more general symmetry. He also described the gravitational field as the gauge field corresponding to the conservation of energy–momentum [140]. The assumption that all interactions are of the gauge type, i.e., that there is a one-to-one correspondence between any conserved quantity and a gauge field, is called the *gauge principle*. Essentially, this was proposed by Utiyama. Let us give an example of the kind of result that can be obtained from this principle. Consider the hypothesis of baryon number conservation. If this is correct, there should be a gauge field that corresponds to it. This will be a vector field, and its interaction strength should be proportional to the baryon number. Since the mass is approximately proportional to the baryon number in macroscopic matter, this force looks very much like gravity. However, its effect should have been observed as a deviation from Newton’s law of gravitation. And since the law of gravitation holds to very good accuracy, it is hard to imagine the existence of such a force. Therefore, the basis for assuming the conservation of baryon number is flimsy. This turns out to be connected with the possibility of proton decay.

In the following, we shall present the formalism needed to treat gauge fields.

## 14.1 Gauge Transformations of the Electromagnetic Field

As mentioned above, the electromagnetic field was the first known gauge field. In the mathematical formulation of a gauge field, the gauge transformations have a fundamental importance, not yet discussed above. We thus begin by studying the gauge transformations of the electromagnetic gauge field, and extend the concept to other, non-Abelian gauge fields.

We express a quantum of electric charge of a field  $\psi_a$  as  $e_a = eq_a$ , where  $e$  has units of electric charge and  $q_a$  is a dimensionless number. If  $eQ$  is the electric charge operator, then

$$[\psi_a(x), Q] = q_a\psi_a(x) , \quad (14.1)$$

so  $q_a$  is one of the eigenvalues of  $Q$ . We consider the unitary operator

$$\mathcal{R}(\lambda) = \exp(i\lambda Q) , \quad (14.2)$$



where  $\lambda$  is a real parameter, and we introduce a phase transformation for  $\psi_a(x)$ :

$$\psi_a(x) \longrightarrow \psi'_a(x) = \mathcal{R}^{-1}(\lambda)\psi_a(x)\mathcal{R}(\lambda) = e^{i\lambda q_a}\psi_a(x). \quad (14.3)$$

In a system where the electric charge is conserved, the Lagrangian density is invariant under this transformation. This is called a *global symmetry*. The transformation (14.3) is an example of (14.2), and we write it as

$$U(\lambda) = e^{i\lambda q_a}. \quad (14.4)$$

We now replace the parameter  $\lambda$  by a function of the coordinate, viz.,

$$\lambda \rightarrow \lambda(x). \quad (14.5)$$

When we insert this  $\lambda(x)$ , we call (14.2), (14.3), and (14.4) a *local gauge transformation*, in contrast to the one with constant  $\lambda$ , which is called a *global gauge transformation*. Hence, (14.3) becomes

$$\psi_a(x) \rightarrow U(\lambda(x))\psi_a(x). \quad (14.6)$$

Introducing the gauge field  $A_\mu$ , we define the covariant derivative  $D_\mu$  by

$$D_\mu = \partial_\mu - ie_a A_\mu(x), \quad (14.7)$$

which is intended to act on  $\psi_a(x)$ . To determine how  $A_\mu$  transforms under the gauge transformation, we require  $D_\mu\psi_a$  to transform in the same way as (14.6):

$$D'_\mu\psi'_a(x) = U(\lambda(x))(D_\mu\psi_a(x)). \quad (14.8)$$

Combining (14.6) with (14.8),

$$D'_\mu = U(\lambda(x))D_\mu U(\lambda(x))^{-1}. \quad (14.9)$$

Inserting (14.7) into this equation, the transformation property of  $A_\mu$  is determined to be

$$A_\mu(x) \rightarrow A_\mu(x) + \frac{1}{e}\partial_\mu\lambda(x). \quad (14.10)$$

Note that the transformation above no longer depends on how we choose  $a$ . Moreover, from (14.9),

$$[D'_\mu, D'_\nu] = U(\lambda(x))[D_\mu, D_\nu]U(\lambda(x))^{-1}. \quad (14.11)$$

Since this commutator does not include derivatives, it commutes with  $U$ . It therefore becomes an invariant quantity under the gauge transformation:

$$[D_\mu, D_\nu] = -ie_a(\partial_\mu A_\nu - \partial_\nu A_\mu) = -ie_a F_{\mu\nu} . \quad (14.12)$$

We can also write down the Jacobi identity for  $D$ :

$$[D_\mu, [D_\nu, D_\sigma]] + [D_\nu, [D_\sigma, D_\mu]] + [D_\sigma, [D_\mu, D_\nu]] = 0 . \quad (14.13)$$

Inserting (14.12) into this equation, we find

$$\partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} + \partial_\sigma F_{\mu\nu} = 0 . \quad (14.14)$$

This is called the *Bianchi identity*.

## 14.2 Non-Abelian Gauge Fields

In the previous section, we only considered one conserved quantity, namely  $Q$ . Let us now consider the case where there are  $N$  conserved charges  $Q_1, Q_2, \dots, Q_N$ , which form a Lie algebra, i.e., the commutation relations among them take the form

$$[Q_a, Q_b] = if_{abc} Q_c . \quad (14.15)$$

We consider a field  $\psi(x)$  that carries an  $n$ -dimensional irreducible representation of this algebra, i.e.,

$$[\psi_r(x), Q_a] = (t_a)_{rs} \psi_s(x) , \quad r, s = 1, 2, \dots, n . \quad (14.16)$$

The matrices  $t_a$  form an  $n$ -dimensional representation of the charges  $Q_a$ . Just as representations of  $Q$  for different fields were given by different values of  $q$  in the case of the electromagnetic field, so the  $t_a$  representing the  $Q_a$  will also be different for different fields. By analogy with the last section, we define

$$\mathcal{R}(\lambda) = \exp \left( i \sum_{i=1}^N \lambda^i Q_i \right) , \quad (14.17)$$

and specify a phase transformation for  $\psi(x)$  by

$$\psi(x) \rightarrow \psi'(x) = \mathcal{R}^{-1}(\lambda) \psi(x) \mathcal{R}(\lambda) = U(\lambda) \psi(x) , \quad (14.18)$$

where

$$U(\lambda) = \exp\left(i \sum_{a=1}^N \lambda^a t_a\right). \quad (14.19)$$

We now make this transformation local, viz.,

$$\lambda^a \rightarrow \lambda^a(x),$$

and define the covariant derivative  $D_\mu$  of  $\psi(x)$  by

$$D_\mu = \partial_\mu - ig \sum_{a=1}^N t_a A_\mu^a(x) \equiv \partial_\mu - ig \mathcal{A}_\mu(x). \quad (14.20)$$

Of course,  $\mathcal{A}_\mu$  depends explicitly on the representation  $t$ . As in the last section, the gauge transformation of  $D_\mu$  is

$$D'_\mu = U(\lambda(x)) D_\mu U(\lambda(x))^{-1}. \quad (14.21)$$

This determines the form of the gauge transformation of  $\mathcal{A}_\mu$ :

$$\mathcal{A}'_\mu(x) = U(\lambda(x)) \mathcal{A}_\mu(x) U(\lambda(x))^{-1} - \frac{1}{ig} U(\lambda(x)) \partial_\mu U(\lambda(x))^{-1}. \quad (14.22)$$

If we write down this transformation for the component  $A_\mu^a$ , we see that it does not in fact depend on how we choose the representation  $t$ . For instance, considering an infinitesimal  $\lambda(x)$ ,

$$A_\mu^{a'}(x) = A_\mu^a(x) + f_{abc} A_\mu^b(x) \lambda^c(x) + \frac{1}{g} \partial_\mu \lambda^a(x), \quad (14.23)$$

and the representation  $t$  drops out completely. Additionally, the commutators of the components of the covariant derivative are

$$[D_\mu, D_\nu] = -ig(\partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu - ig[\mathcal{A}_\mu, \mathcal{A}_\nu]) \equiv -ig \mathcal{F}_{\mu\nu}, \quad (14.24)$$

$$\mathcal{F}_{\mu\nu} = \sum_{a=1}^N t_a \left( \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f_{abc} A_\mu^b A_\nu^c \right). \quad (14.25)$$

Once again, we see that the components of  $\mathcal{F}_{\mu\nu}$  do not depend on how we choose  $t$ . The gauge transformation of  $\mathcal{F}_{\mu\nu}$  is given by

$$\mathcal{F}'_{\mu\nu}(x) = U(\lambda(x))\mathcal{F}_{\mu\nu}(x)U(\lambda(x))^{-1} . \quad (14.26)$$

Since the  $t_a$  do not commute in this case,  $\mathcal{F}_{\mu\nu}$  is not gauge invariant. However,

$$\text{Tr}(\mathcal{F}_{\mu\nu}(x)\mathcal{F}_{\mu\nu}(x)) \quad (14.27)$$

is clearly gauge invariant, and can be used as the Lagrangian density of the gauge field.  $A_\mu^a(x)$  is called a *non-Abelian gauge field*. Once again, we can derive the Bianchi identity from (14.13).

### 14.3 Gravitational Field as a Gauge Field

Utiyama was the first to note that the gravitational field is a sort of gauge field [140]. Since the symmetry of the gravitational field under a space-time transformation is different from an internal symmetry that is independent of space-time, many complications arise. So far we have introduced the gauge field from the condition that a given field and its covariant derivative should transform in the same way. However, for a space-time coordinate transformation, it is the rank of the tensor that determines the irreducible representation, and the covariant derivative alters this rank, so the given field and its covariant derivative clearly transform in different ways. For example, for a scalar field  $\phi$ , if we consider  $D_\mu D_\nu \phi$ , the first covariant derivative  $D_\nu$  is the covariant derivative of the scalar field, which is just  $\partial_\nu$ , while the second covariant derivative  $D_\mu$  is the covariant derivative of the vector field  $\phi_\nu \equiv \partial_\nu \phi$ , i.e.,

$$D_\mu \phi_\nu = \partial_\mu \phi_\nu - \Gamma_{\mu\nu}^\lambda \phi_\lambda . \quad (14.28)$$

If we consider  $\Gamma_{\mu\nu}^\lambda$  as a matrix element of the matrix  $\Gamma_\mu$ , viz.,

$$(\Gamma_\mu)_\nu^\lambda = \Gamma_{\mu\nu}^\lambda , \quad (14.29)$$

it turns out that  $\Gamma_\mu$  corresponds to  $\text{ig}\mathcal{A}_\mu$  in the previous section.  $\Gamma_{\mu\nu}^\lambda$  is the Christoffel symbol of the second kind and is defined by

$$\Gamma_{\mu\nu}^\lambda = \frac{1}{2}g^{\lambda\sigma}(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}) . \quad (14.30)$$

For a general coordinate transformation,  $\Gamma$  does not transform as a tensor. An additional term arises, characteristic of the transformation of a gauge field, involving a derivative of the transformation matrix.

We now construct a quantity corresponding to  $\mathcal{F}_{\mu\nu}$ . As mentioned before, the representation of  $D$  depends on what  $D$  acts on. For instance, if we consider a scalar field  $\phi$ , then

$$[D_\mu, D_\nu]\phi = \phi_{; \nu; \mu} - \phi_{; \mu; \nu} , \quad (14.31)$$

where  $;$   $\mu$  indicates the covariant derivative. However, for a vector field  $A_\nu$ ,

$$[D_\sigma, D_\rho]A_\nu = A_{\nu; \rho; \sigma} - A_{\nu; \sigma; \rho} = R_{\nu\rho\sigma}^\lambda A_\lambda . \quad (14.32)$$

Hence, using the matrix  $\Gamma$ , what corresponds to  $\mathcal{F}_{\mu\nu}$  in the previous section is

$$\begin{aligned} R_{\nu\mu\sigma}^\lambda &= (\mathcal{F}_{\rho\sigma})_\nu{}^\lambda = (\partial_\rho \Gamma_\sigma - \partial_\sigma \Gamma_\rho - [\Gamma_\rho, \Gamma_\sigma])_\nu{}^\lambda \\ &= \partial_\rho \Gamma_{\sigma\nu}^\lambda - \partial_\sigma \Gamma_{\rho\nu}^\lambda - \Gamma_{\rho\nu}^\alpha \Gamma_{\sigma\alpha}^\lambda + \Gamma_{\sigma\nu}^\alpha \Gamma_{\rho\alpha}^\lambda . \end{aligned} \quad (14.33)$$

This has essentially the same form as  $\mathcal{F}_{\mu\nu}$  in the previous section. Moreover, the Jacobi identity for the scalar field  $\phi$ , viz.,

$$([D_\nu, [D_\rho, D_\sigma]] + [D_\rho, [D_\sigma, D_\nu]] + [D_\sigma, [D_\nu, D_\rho]])\phi = 0 , \quad (14.34)$$

is equivalent to the following equation (multiplied by  $\partial_\lambda \phi$ ):

$$R_{\nu\rho\sigma}^\lambda + R_{\rho\sigma\nu}^\lambda + R_{\sigma\nu\rho}^\lambda = 0 . \quad (14.35)$$

Similarly, for the vector field  $A_\mu$ , the equation

$$([D_\nu, [D_\rho, D_\sigma]] + [D_\rho, [D_\sigma, D_\nu]] + [D_\sigma, [D_\nu, D_\rho]])A_\mu = 0 \quad (14.36)$$

implies

$$R_{\mu\rho\sigma; \nu}^\lambda + R_{\mu\sigma\nu; \rho}^\lambda + R_{\mu\nu\rho; \sigma}^\lambda = 0 . \quad (14.37)$$

Equations (14.35) and (14.37) are the Bianchi identities in this case.

# Chapter 15

## Spontaneous Symmetry Breaking



In the last chapter we discussed gauge theory. As far as it goes, we remain unable to eliminate other interactions and replace them by gauge interactions. It may happen that the symmetry in the Lagrangian is broken for some reason and is not apparent. This is called a *broken or hidden symmetry*. In this case, a quantum of the gauge field corresponding to the broken or hidden symmetry acquires mass, in contrast to the photon. In this chapter, we will discuss a mechanism whereby the symmetry in the Lagrangian is broken.

### 15.1 Nambu–Goldstone Particles

We assume that a given Lagrangian is invariant under the Lie group  $G$ . We call the ground state of this system the *vacuum* and denote it by  $|\mathbf{0}\rangle$ . In this case, there are two possibilities:

1.  $|\mathbf{0}\rangle$  is invariant under  $G$ , whence it belongs to a one-dimensional representation of  $G$ ;
2.  $|\mathbf{0}\rangle$  is not invariant under  $G$ , and it belongs to a multi-dimensional representation of  $G$ .

The former case is a manifest symmetry and the latter is a hidden symmetry. In the latter case, we say that *the symmetry is spontaneously broken*, and  $|\mathbf{0}\rangle$  is clearly degenerate. There exist many examples like this outside field theory. A well known example is the ferromagnet. In the ground state, the electron spins align in one direction. The direction in which the spins align depends on how the ground state has been prepared, but originally the system is rotationally symmetric, so this ground state is degenerate with respect to spin directions.

**Theorem 15.1** *Nambu–Goldstone theorem.* *When the Lagrangian is invariant under the Lie group  $G$  and its invariance is spontaneously broken, there exist particles whose masses and spins are both zero.*

This was first discovered by Nambu and Jona-Lasinio [141, 142], and then a simpler example was given by Goldstone [143]. The particles in question are called *Nambu–Goldstone particles*.

If  $|\mathbf{0}\rangle$  does not belong to the one-dimensional representation, there exist local scalar operators satisfying

$$\langle \mathbf{0} | \rho_a(x) | \mathbf{0} \rangle \neq 0 . \quad (15.1)$$

Although the  $\rho_a$  themselves belong to the multi-dimensional representation, how they are represented depends on the representation of  $|\mathbf{0}\rangle$ . The expectation value of operators with non-zero spins always disappears because of Lorentz invariance. In the following, we consider the case where we can choose the scalar field operators themselves as the  $\rho_a$ . The original Nambu–Jona-Lasinio model chose a bilinear form of the Fermi field as  $\rho$ . This is because no scalar field was included in their model.

We express an infinitesimal transformation of the group  $G$  in the form

$$\delta\varphi_a = \epsilon_\alpha T_{ab}^\alpha \varphi_b . \quad (15.2)$$

If we rewrite this with  $\epsilon_\alpha \rightarrow \lambda^a$ ,  $T^\alpha \rightarrow it_a$ , it coincides with (14.18). The corresponding Noether current is

$$J_\mu^\alpha = \frac{\partial \mathcal{L}}{\partial \varphi_{a,\mu}} T_{ab}^\alpha \varphi_b . \quad (15.3)$$

It satisfies the local conservation law

$$\partial_\mu J_\mu^\alpha = 0 . \quad (15.4)$$

We thus analyze the global conservation law in the two cases mentioned at the beginning of this section:

1. When a symmetry is manifest, there are global conserved quantities

$$Q^\alpha = \int d^3x J_0^\alpha(x) , \quad (15.5)$$

and the following global conservation law holds:

$$\frac{dQ^\alpha}{dt} = 0 . \quad (15.6)$$

Moreover, since  $|\mathbf{0}\rangle$  belongs to the one-dimensional representation of  $G$ , the generators  $Q^\alpha$  of  $G$  annihilate the vacuum:

$$Q^\alpha |\mathbf{0}\rangle = 0 . \quad (15.7)$$

It is easily shown from the definition of (15.3) that, because the  $Q^\alpha$  are generators of the group  $G$ , we must have

$$i[Q^\alpha, \varphi_a] = T_{ab}^\alpha \varphi_b . \quad (15.8)$$

In this case, it is clear from (15.7) that

$$i\langle \mathbf{0} | [Q^\alpha, \varphi_a] | \mathbf{0} \rangle = T_{ab}^\alpha \langle \mathbf{0} | \varphi_b | \mathbf{0} \rangle = 0 . \quad (15.9)$$

In addition, the generators  $Q^\alpha$  are the same both with and without renormalization, i.e., (15.8) and (15.9) hold for the field operators both with and without renormalization, and they have the same form. However, for the discussion below, it will be more useful to consider the renormalized case.

2. For the broken symmetry, there is a component  $\varphi_b$  satisfying

$$T_{ab}^\alpha \langle \varphi_b \rangle \neq 0 , \quad \langle \varphi_b \rangle = \langle \mathbf{0} | \varphi_b | \mathbf{0} \rangle . \quad (15.10)$$

In this case, we will prove the Nambu–Goldstone theorem. We first study the structure of the two-point function

$$i\langle \mathbf{0} | [J_\mu^\alpha(x), \varphi_\alpha(y)] | \mathbf{0} \rangle \quad (15.11)$$

under the *explicit Lorentz covariance*. First, from the discussion about the integral representation in Sect. 12.8, we can write down the following integral representation as the most general form of (15.11):

$$i\langle \mathbf{0} | [J_\mu^\alpha(x), \varphi_\alpha(y)] | \mathbf{0} \rangle = \partial_\mu \int dm^2 \rho_a^\alpha(m^2) \Delta(x-y; m^2) . \quad (15.12)$$

In addition, the spectral function  $\rho_a^\alpha(m^2)$  is given by

$$p_\mu \theta(p_0) \rho_a^\alpha(-p^2) = (2\pi)^3 i \sum_n \delta^4(p-p_n) \langle \mathbf{0} | J_\mu^\alpha(0) | \mathbf{n} \rangle \langle \mathbf{n} | \varphi_a(0) | \mathbf{0} \rangle . \quad (15.13)$$

Extracting a divergence in (15.12) and using the local conservation law (15.4),

$$\int dm^2 \rho_a^\alpha(m^2) m^2 \Delta(x-y; m^2) = 0 , \quad (15.14)$$

$$m^2 \rho_a^\alpha(m^2) = 0 . \quad (15.15)$$



A general form of  $\rho$  satisfying these equations is

$$\rho_a^\alpha(m^2) = N_a^\alpha \delta(m^2) . \quad (15.16)$$

Thus, denoting the  $\Delta$ -function for  $m^2 = 0$  by  $D$ ,

$$i\langle \mathbf{0} | [J_a^\alpha(x), \varphi_a(y)] | \mathbf{0} \rangle = N_a^\alpha \partial_\mu D(x - y) . \quad (15.17)$$

Now, taking  $\mu = 4$  and using (15.8),

$$N_a^\alpha = T_{ab}^\alpha \langle \varphi_b \rangle \neq 0 . \quad (15.18)$$

Hence,

$$i\langle \mathbf{0} | [J_\mu^\alpha(x), \varphi_a(y)] | \mathbf{0} \rangle = T_{ab}^\alpha \langle \varphi_b \rangle \partial_\mu D(x - y) . \quad (15.19)$$

This equation implies that the scalar field  $\varphi_a$  yields a massless one-particle state. This particle is the Nambu–Goldstone particle. Moreover, in this case the term proportional to  $\partial_\mu \varphi_a$  is included in  $J_\mu^\alpha$ :

$$\int d^3x \partial_\mu \varphi_a \sim V \frac{1}{\sqrt{V}} = \sqrt{V} \rightarrow \infty , \quad V \rightarrow \infty . \quad (15.20)$$

Hence, the integral (15.18) defining  $Q^\alpha$  diverges, so (15.8) can be rewritten as

$$i \int d^3x [J_0^\alpha(x), \varphi_a(y)] = T_{ab}^\alpha \varphi_b(y) , \quad x_0 = y_0 . \quad (15.21)$$

Changing the order of taking the commutator and doing the integration makes no difference. In this case, the global conserved quantities  $Q^\alpha$  are zero. As a consequence, this broken symmetry becomes a hidden symmetry.

No particle whose mass and spin are both zero has yet been observed experimentally. Does this mean that such a broken symmetry cannot be realized in the real world? Or does the Nambu–Goldstone particle disappear as a result of some other mechanism? Solving this issue will be the aim of the present chapter.

## 15.2 Sigma Model

In the last section, we investigated the relationship between the broken symmetry and the Nambu–Goldstone particle. In this section, we will discuss a simple model which realizes this. We consider the system consisting of a pseudo-scalar triplet  $\pi$

and a scalar singlet  $\sigma$ :

$$\pi_1, \pi_2, \pi_3, \sigma. \quad (15.22)$$

To distinguish between the scalar and the pseudo-scalar, we have to introduce fermions and interactions such as

$$i\bar{\psi}\gamma_5\frac{\hat{\tau}}{2}\psi\hat{\phi}, \quad \bar{\psi}\psi\sigma. \quad (15.23)$$

For simplicity, we will only consider bosons. We introduce the notation

$$\varphi_a^2 = \varphi_1^2 + \varphi_2^2 + \varphi_3^2 + \sigma^2 = \hat{\varphi}^2 + \sigma^2, \quad (15.24)$$

where the  $\varphi_i$  are the field operators for the  $\pi_i$  and  $\sigma$  is the field operator for  $\sigma$ .

For the Lagrangian density, we take

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\varphi_a)^2 + \frac{1}{4}m^2\varphi_a^2 - \frac{1}{8}f^2(\varphi_a^2)^2, \quad (15.25)$$

which has  $O(4)$  symmetry. Note that the mass term has the opposite sign to the usual one. Minus the sum of the second and third terms is the potential, viz.,

$$U(\varphi) = -\frac{1}{4}m^2\varphi_a^2 + \frac{1}{8}f^2(\varphi_a^2)^2. \quad (15.26)$$

Then  $\varphi_a = 0$  gives an extremum of  $U(\varphi)$ , but it does not give the minimum. Put another way, the ordinary symmetric vacuum is not stable. Before introducing non-symmetric vacua, let us investigate the algebraic structure of this model. In this system, there are six local conservation laws. We thus introduce the following vector and axial-vector currents:

$$\hat{V}_\mu = -\hat{\boldsymbol{\phi}} \times \partial_\mu \hat{\boldsymbol{\phi}}, \quad \hat{A}_\mu = \hat{\boldsymbol{\phi}} \partial_\mu \sigma - \sigma \partial_\mu \hat{\boldsymbol{\phi}}. \quad (15.27)$$

Both these current densities satisfy the local conservation laws

$$\partial_\mu \hat{V}_\mu = 0, \quad \partial_\mu \hat{A}_\mu = 0. \quad (15.28)$$

If the global conservation laws hold, then taking  $\hat{\boldsymbol{\pi}} = \hat{\boldsymbol{\phi}}$  and  $\pi_0 = \dot{\sigma}$ , the conserved quantities are

$$\hat{I} = \int d^3x (\hat{\boldsymbol{\phi}} \times \hat{\boldsymbol{\pi}}), \quad \hat{I}^5 = \int d^3x (\sigma \hat{\boldsymbol{\pi}} - \hat{\boldsymbol{\phi}} \pi_0). \quad (15.29)$$

Comparing with (13.166), (13.170), and (13.171), the commutation relations among them are

$$[I_i, I_j] = i\epsilon_{ijk} I_k, \quad [I_i, I_j^5] = i\epsilon_{ijk} I_k^5, \quad [I_i^5, I_j^5] = i\epsilon_{ijk} I_k. \quad (15.30)$$

Even if the global conservation laws do not hold, the commutation relations among the densities still do. Neglecting the Goto–Imamura–Schwinger term discussed in Sect. 12.8 and setting  $x_0 = y_0$ ,

$$[V_{i0}(x), V_{j0}(y)] = i\epsilon_{ijk} \delta^3(x - y) V_{k0}(x). \quad (15.31)$$

Taking

$$\mathcal{F}_{i\mu}^{(\pm)} = \frac{1}{2} (V_{i\mu} \pm A_{i\mu}) \quad (15.32)$$

for the density, as discussed in Sect. 13.6, we have the current algebra

$$SU(2)_L \times SU(2)_R. \quad (15.33)$$

Adding further fermions to this model, a representation of the current algebra can be made. Indeed, it leads to chiral dynamics and we can construct a simple representation for the current algebra. However, in the real world there is no particle corresponding to  $\sigma$ , so representations without  $\sigma$  have been proposed. Unfortunately, since the representation without  $\sigma$  is a rather non-linear theory, work eventually shifted to the quark model. In the following, we will consider the model in which  $\sigma$  survives. Coming back to the topic of broken symmetry, we write the field equation in the form

$$\square \varphi_a + \frac{1}{2} (m^2 - f^2 \varphi_b^2) \varphi_a = 0. \quad (15.34)$$

Considering this as a classical equation, the requirement of causality cannot be satisfied unless

$$\frac{1}{2} (m^2 - f^2 \varphi_b^2) \leq 0, \quad (15.35)$$

whence  $\langle \varphi_b \rangle^2 \neq 0$ . Corresponding to the minimum of  $U(\langle \varphi \rangle)$ , we take

$$\langle \hat{\varphi} \rangle = 0, \quad \langle \sigma \rangle = \eta \neq 0, \quad (15.36)$$

where the parity and the isospin are conserved. Hence, if we take

$$\sigma = \sigma_0 + \eta, \quad (15.37)$$

then  $\eta$  is determined by the condition minimizing the potential, viz.,

$$m^2 = f^2 \eta^2 . \quad (15.38)$$

Since we have introduced symmetry breaking, if we investigate what kinds of equations  $\hat{\phi}$  and  $\sigma_0$  satisfy, we find

$$\square \hat{\phi} = \text{non-linear terms} , \quad (\square - m^2) \sigma_0 = \text{non-linear terms} . \quad (15.39)$$

What we have understood from these is that the mass of  $\pi$  vanishes,  $\sigma$  has a finite mass, and  $\pi$  is the Nambu–Goldstone particle. Although the discussion about symmetry breaking has been classical in this section, we will give a more detailed quantum mechanical discussion in the next.

Interpreting  $\pi$  as the pion, since it actually has a small mass, we add a small term to introduce a slight symmetry breaking artificially. This term can be considered to determine the direction of symmetry breaking. If  $a$  is a small quantity,

$$\mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L} + a\sigma . \quad (15.40)$$

In this case,  $\eta$  is shifted by order  $a$  and becomes  $\eta'$ :

$$\sigma = \sigma_0 + \eta' . \quad (15.41)$$

Minimizing the potential classically in this case,

$$m^2 = f^2 \eta'^2 - \frac{2a}{\eta'} . \quad (15.42)$$

Therefore, if  $a$  is small,  $\eta'$  differs from  $\eta$  only by a term of order  $a$ . Although the mass of  $\sigma$  is given once again by  $m$ , the mass of  $\pi$  becomes finite, and if we write it as  $\mu$ , we have

$$\mu^2 = \frac{1}{2} \left( f^2 \eta'^2 - m^2 \right) = \frac{a}{\eta'} . \quad (15.43)$$

Moreover, in this case, the local conservation law of the axial vector current is no longer observed. If we write  $\eta'$  as  $f_\pi$ , then

$$\partial_\mu \hat{A}_\mu = -\mu^2 f_\pi \hat{\phi} . \quad (15.44)$$

This is called the partially conserved axial-vector current (PCAC) hypothesis [129]. Considering  $\pi$  as a composite particle, we define its phenomenological field operator by this equation.

Since the above symmetry breaking argument has been classical and phenomenological, we now formulate this using Green's functions in order to give a rigorous quantum mechanical discussion.

### 15.3 The Mechanism of Spontaneous Symmetry Breaking

As an example of spontaneous symmetry breaking, we previously mentioned the ferromagnet. When there is no magnetic field, the directions of the electron spins are random at high temperature because of their thermal motions, whence there is rotational symmetry. If we turn on a magnetic field, the ferromagnet is magnetized in the direction of the magnetic field. If we now gradually decrease the temperature a phase transition occurs. It turns out that even after removing the magnetic field, it remains magnetized. In order to translate this phenomenon into the case for spontaneous symmetry breaking in field theory, we use a Green's function. The necessary formalism has been given in Sect. 11.2.

The generating function  $\mathcal{R}[J]$  for connected Green's functions is a functional of the external field  $J$ , where  $J$  corresponds to the magnetic field imposed from the outside in the ferromagnet case. The vacuum expectation value of the field  $\varphi(x)$  produced as a consequence corresponds to the magnetization at low temperature. We wish to investigate whether magnetization occurs even in the case where there is no external field. Choosing the magnetic field as an independent variable, we ask whether the magnetization can vanish when the magnetic field takes a suitable value. This suitable value corresponds to the magnetization of the permanent magnet.

Although in Sect. 11.2, we introduced the functional  $\mathcal{R}[J]$  of  $J(x)$  and the functional  $\mathcal{F}[\langle\varphi(x)\rangle]$  of  $\langle\varphi(x)\rangle$ , since it is useful to slightly change the phases, we write

$$\mathcal{R} = -iW, \quad \mathcal{F} = -i\Gamma, \quad (15.45)$$

whence (11.36) becomes

$$\Gamma = W - \int d^4x \langle\varphi(x)\rangle J(x). \quad (15.46)$$

The role of the new variables as generating functionals is expressed by

$$\frac{\delta W}{\delta J(x)} = \langle\varphi(x)\rangle, \quad \frac{\delta \Gamma}{\delta \langle\varphi(x)\rangle} = -J(x). \quad (15.47)$$

$W$  and  $\Gamma$  are generating functionals of the connected Green's function and of the one-particle irreducible Green's function, respectively. The former is the Green's function corresponding to connected Feynman diagrams and the latter is the Green's

function without a pole corresponding to the one-particle state. This is because in the latter the one-particle propagator is absorbed into  $\langle \varphi(x) \rangle$ .

The connected and one-particle irreducible Green's functions in the momentum space are defined by

$$\frac{\delta^n W[J]}{\delta J(x_1) \dots \delta J(x_n)} \Big|_{J=0} = \frac{1}{(2\pi)^{4(n-1)}} \int (d p) \delta(p_1 + \dots + p_n) e^{i p_1 \cdot x_1 + \dots + i p_n \cdot x_n} G^{(n)}(p_1, \dots, p_n) \quad (15.48)$$

and

$$\frac{\delta^n \Gamma[\langle \varphi \rangle]}{\delta \langle \varphi(x_1) \rangle \dots \delta \langle \varphi(x_n) \rangle} \Big|_{\langle \varphi(x) \rangle = \varphi} = \frac{1}{(2\pi)^{4(n-1)}} \int (d p) \delta(p_1 + \dots + p_n) e^{i p_1 \cdot x_1 + \dots + i p_n \cdot x_n} \Gamma^{(n)}(p_1, \dots, p_n), \quad (15.49)$$

respectively, where  $(d p) = d^4 p_1 \dots d^4 p_n$  and  $\varphi$  is a constant. In the theory in which the symmetry is spontaneously broken, it turns out that  $\varphi$  takes a value different from zero. We now define the effective potential  $V(\varphi)$  by

$$V(\varphi) = -\Gamma[\varphi]. \quad (15.50)$$

At the lowest order,  $\Gamma[\varphi]$  coincides with the value when the field operator is replaced by the constant  $\varphi$  in the Lagrangian density. Therefore, it coincides with  $-U(\varphi)$  in the previous section, i.e.,

$$V^{(0)}(\varphi) = -\mathcal{L}(\varphi) = U(\varphi). \quad (15.51)$$

However, since there are differences when we include higher-order quantum corrections, this is called an *effective potential*. In the following, we consider the case where  $\varphi$  is multi-component. We assume that the theory is invariant under a Lie group  $G$ , as considered in Sect. 15.1. Thus, if we have invariance under the transformation (15.2),

$$V(\varphi_a + \varepsilon_\alpha T_{ab}^\alpha \varphi_b) = V(\varphi_a), \quad \text{or} \quad \frac{\partial V}{\partial \varphi_a} T_{ab}^\alpha \varphi_b = 0. \quad (15.52)$$

Corresponding to (11.38) and (11.40), we have

$$\frac{\partial V}{\partial \varphi_a} = J_a, \quad \frac{\partial^2 V}{\partial \varphi_a \partial \varphi_b} = -\Delta_F^{-1}(0)_{ab}, \quad (15.53)$$

where  $J_a$  is the external field corresponding to  $\varphi_a$ , while  $\Delta_{\mathbb{F}}^{-1}(0)$  corresponds to  $p^2 = 0$  and expresses the inverse matrix of the multi-component propagator when the expectation value  $\varphi$  does not necessarily vanish. Differentiating (15.52),

$$\frac{\partial^2 V}{\partial \varphi_a \partial \varphi_c} T_{ab}^\alpha \varphi_b + \frac{\partial V}{\partial \varphi_a} T_{ac}^\alpha = 0 . \quad (15.54)$$

Combining this with (15.53),

$$\Delta_{\mathbb{F}}^{-1}(0)_{ca} T_{ab}^\alpha \varphi_b = J_a T_{ac}^\alpha . \quad (15.55)$$

The conditions for  $\varphi$  to have finite components even if the external field  $J$  vanishes, i.e., the conditions for the symmetry to be spontaneously broken, are

$$J_a = \frac{\partial V}{\partial \varphi_a} = 0 , \quad \Delta_{\mathbb{F}}^{-1}(0)_{ca} T_{ab}^\alpha \varphi_b = 0 . \quad (15.56)$$

The second equation means that there will be massless Nambu–Goldstone particles. It is a necessary condition for this equation to have solutions except for  $\varphi = 0$ . A sufficient condition is that  $V(\varphi)$  should have a minimum for the value of  $\varphi$ .

We consider the  $SO(n)$  symmetry as an example, and take

$$V(\varphi) = F(\varphi_a^2) , \quad (15.57)$$

where  $\varphi_a^2 = \varphi_1^2 + \cdots + \varphi_n^2$ . Assuming that  $F(x)$  has a minimum at  $x = x_0$ ,

$$F'(x_0) = 0 , \quad F''(x_0) > 0 . \quad (15.58)$$

In this case,  $J_a$  vanishes for  $\varphi_a^2 = x_0$ :

$$J_a = 2\varphi_a F'(x_0) = 0 . \quad (15.59)$$

It turns out that the second equation in (15.56) then also holds. The solutions for  $\varphi_a^2 = x_0$  are now degenerate. Thus, the ground states of this system turn out to be degenerate. In this case,

$$\Delta_{\mathbb{F}}^{-1}(0)_{ab} = -\frac{\partial^2 V}{\partial \varphi_a \partial \varphi_b} = -2\delta_{ab} F'(\varphi_a^2) - 4\varphi_a \varphi_b F''(\varphi_a^2) . \quad (15.60)$$

Taking  $\varphi_a^2 = x_0$ ,

$$\Delta_{\mathbb{F}}^{-1}(0)_{ab} = -4\varphi_a \varphi_b F''(x_0) . \quad (15.61)$$

The eigenvalues of this matrix are  $-4x_0 F''(x_0)$  and 0 for all others. The eigenvector for the first of these eigenvalues is parallel to the  $n$ -dimensional vector  $(\varphi_1, \dots, \varphi_n)$ , while the eigenvector for the latter eigenvalue is a vector normal to it. Hence, by a proper rotation in  $SO(n)$ , we can write  $\Delta_F^{-1}(0)$  in diagonal form:

$$\Delta_F^{-1}(0) = \begin{pmatrix} D & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}. \quad (15.62)$$

It turns out that, in this new coordinate system, the components of  $\varphi$  can be written as  $(\sqrt{x_0}, 0, \dots, 0)$ . In this case, if we take  $c = 1$  in (15.56), then from the anti-symmetry of  $T$ , we obtain  $T_{ab}^\alpha \varphi_b = 0$ . If we take  $c \neq 1$ , then we obtain  $\Delta_F^{-1}(0)_{ca} = 0$ , so (15.56) holds in every case. Since the component corresponding to  $c = 1$  has a finite mass and components corresponding to  $c \neq 1$  give massless poles for  $\Delta_F(p^2)$ , Nambu–Goldstone particles turn out to be included. Additionally, since for  $c \neq 1$ ,  $T_{cb}^\alpha \varphi_b \neq 0$  for a proper  $\alpha$ , the discussion in Sect. 15.1 also leads to the conclusion that Nambu–Goldstone particles exist.

The discussion about the sigma model in the previous section shows that the computation above has been carried out explicitly based on the approximation (15.51). The components for  $c = 1$  and for  $c \neq 1$  above correspond to  $\sigma$  and  $\pi$ , respectively.

## 15.4 Higgs Mechanism

So far we have considered the gauge principle. If the symmetry is not broken, there must be quanta of massless gauge fields for the number of conserved quantities. On the other hand, if the symmetry is spontaneously broken, then there must be massless scalar particles. However, in reality, no massless boson has yet been observed, except for the photon. So is it possible for massless bosons to be eliminated by an as-yet-unknown mechanism? It was Englert, Brout, and Higgs in 1964 who gave the first interesting discussion of this issue [144–146]. In the following, we will present their approach.

We consider the simplest symmetric group,  $U(1)$  or  $SO(2)$ . We assume that a given system includes two real scalar fields  $\varphi_1$  and  $\varphi_2$ , and is invariant under the transformation

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \rightarrow \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}. \quad (15.63)$$

If the Lagrangian density of this system is invariant under this transformation, then we can introduce the corresponding Noether current  $j_\mu(x)$ . We obtain the



commutation relations

$$i \int d^3x [j_0(x), \varphi_1(y)] = \varphi_2(y) , \quad (15.64)$$

$$i \int d^3x [j_0(x), \varphi_2(y)] = -\varphi_1(y) , \quad (15.65)$$

where  $x_0 = y_0$ . We assume the following spontaneous symmetry breaking:

$$\langle \varphi_2(x) \rangle \neq 0 . \quad (15.66)$$

In the proof of the Nambu–Goldstone theorem,

$$\text{FT } i \langle \mathbf{0} | [j_\mu(x), \varphi_1(y)] | \mathbf{0} \rangle \propto \langle \varphi_2 \rangle \epsilon(k_0) k_\mu \delta(k^2) , \quad (15.67)$$

where FT indicates the Fourier transform, so we could conclude that there are massless scalar particles. However, in this proof we assumed *explicit Lorentz covariance*. But what happens if this assumption does not hold true, as for example in non-relativistic theories, where this theorem does not necessarily apply. If we do not have explicit covariance and a particular unit vector along the time axis appears in a theory, then the general form of the Fourier transform in (15.67) above becomes

$$\text{FT} \sim k_\mu \rho_1(k^2, n \cdot k) + n_\mu \rho_2(k^2, n \cdot k) + C n_\mu \delta^4(k) . \quad (15.68)$$

The general case within this form which satisfies the conservation law, i.e., the one whose internal product with  $k_\mu$  vanishes, becomes

$$\text{FT} \sim k_\mu \delta(k^2) \rho_4(n \cdot k) + [k^2 n_\mu - k_\mu (n \cdot k)] \rho_5(k^2, n \cdot k) + C n_\mu \delta^4(k) . \quad (15.69)$$

Thus, a counterexample for the Nambu–Goldstone theorem occurs when

$$\rho_4 = 0 . \quad (15.70)$$

But even though there exists no explicit covariance, could the theory be Lorentz covariant? The example given by Higgs is QED in a non-covariant gauge, i.e., with the gauge condition in the form

$$n_\mu A_\mu = 0 . \quad (15.71)$$

Maxwell's equations are then

$$\partial_\mu F_{\mu\nu} = -j_\nu , \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu . \quad (15.72)$$

We consider the following general form of the Fourier transform:

$$\text{FT } i\langle \mathbf{0} | [A_\mu(x), \varphi_1(y)] | \mathbf{0} \rangle = k_\mu \sigma_1(k^2, n \cdot k) + n_\mu \sigma_2(k^2, n \cdot k) + D n_\mu \delta^4(k) . \quad (15.73)$$

Therefore, from the gauge condition (15.71), taking  $n^2 = -1$ , we obtain

$$\sigma_2 = (n \cdot k) \sigma_1 , \quad D = 0 . \quad (15.74)$$

Thus, combining this result with (15.72),

$$\text{FT}; i\langle \mathbf{0} | [j_\mu(x), \varphi_1(y)] | \mathbf{0} \rangle = [k^2 n_\mu - k_\mu (n \cdot k)] \sigma_2(k^2, n \cdot k) . \quad (15.75)$$

Comparing this with (15.69), it turns out that (15.70) is satisfied. Choosing a non-covariant gauge like this, the massless Nambu–Goldstone particle does not appear. The reason why we could do this is that there is a gauge field. Needless to say, the transformations (15.64) and (15.65) appearing at the beginning of this section correspond to global gauge transformations.

Bearing this in mind, we introduce the gauge field corresponding to the  $U(1)$  symmetry and consider what happens when the symmetry is spontaneously broken. The following discussion was given by Higgs in [147]. Since this gauge field can be considered to be the same as the electromagnetic field, we introduce the Lagrangian density

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu} - \frac{1}{2} D_\mu \varphi_a D_\mu \varphi_a + \frac{1}{4} m_0^2 \varphi_a \varphi_a . \quad (15.76)$$

The scalar part is the same as that of the sigma model, but  $\varphi_a$  has two components labelled by  $a = 1, 2$ . The covariant derivative  $D_\mu$  is defined by

$$D_\mu \varphi_a = \partial_\mu \varphi_1 - e A_\mu \varphi_2 , \quad D_\mu \varphi_2 = \partial_\mu \varphi_2 + e A_\mu \varphi_1 . \quad (15.77)$$

In QED, we use complex fields,  $\varphi$  and  $\varphi^\dagger$ , and their covariant derivatives, i.e.,

$$\varphi = \frac{1}{\sqrt{2}}(\varphi_1 - i\varphi_2) , \quad \varphi^\dagger = \frac{1}{\sqrt{2}}(\varphi_1 + i\varphi_2) , \quad (15.78)$$

$$D_\mu \varphi = (\partial_\mu - ie A_\mu) \varphi , \quad D_\mu \varphi^\dagger = (\partial_\mu + ie A_\mu) \varphi^\dagger . \quad (15.79)$$

However, to introduce spontaneous symmetry breaking, it is more useful to use a representation with real scalar fields. The field equations are

$$\partial_\nu F_{\mu\nu} = j_\mu = e(\varphi_2 D_\mu \varphi_1 - \varphi_1 D_\mu \varphi_2) , \quad (15.80)$$

$$D_\mu^2 \varphi_a + \frac{1}{2}(m_0^2 - f^2 \varphi_b \varphi_b) \varphi_a = 0 . \quad (15.81)$$

In this example, as in the case of the sigma model, the Lagrangian density has been chosen so that the symmetry is spontaneously broken. Regarding solutions which give an extremum of the effective potential in the lowest order approximation, in addition to  $\langle A_\mu \rangle = 0$  and  $\langle \varphi_b \rangle = 0$ , there is also

$$\langle A_\mu \rangle = 0, \quad \langle \varphi_b \rangle^2 = \frac{m_0^2}{f^2} \equiv \eta^2. \quad (15.82)$$

This corresponds to (15.38).

Note that the Lagrangian density, and therefore the field equations, are invariant under the gauge transformation

$$A_\mu(x) \rightarrow A_\mu(x) + \frac{1}{e} \partial_\mu \lambda(x), \quad (15.83)$$

$$\begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} \rightarrow \begin{pmatrix} \cos \lambda(x) & \sin \lambda(x) \\ -\sin \lambda(x) & \cos \lambda(x) \end{pmatrix} \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix}. \quad (15.84)$$

If the symmetry is spontaneously broken as in (15.82), this solution is degenerate and its degeneracy is generated by the transformation (15.84), for constant  $\lambda$ , i.e., the solution to (15.82) has the following form which involves an arbitrary parameter  $\alpha$ :

$$\langle \varphi_1 \rangle = \eta \cos \alpha, \quad \langle \varphi_2 \rangle = \eta \sin \alpha. \quad (15.85)$$

In the following discussion, we fix  $\alpha = \pi/2$ . Therefore, in the chosen ground state,

$$\langle A_\mu \rangle = 0, \quad \langle \varphi_1 \rangle = 0, \quad \langle \varphi_2 \rangle = \eta. \quad (15.86)$$

We then expand the field operators around this solution, viz.,

$$\varphi_1 = g, \quad \varphi_2 = \eta + h, \quad (15.87)$$

and introduce a vector field  $B_\mu$  with field strength  $F_{\mu\nu}$ :

$$B_\mu = A_\mu - (e\eta)^{-1} \partial_\mu g, \quad (15.88)$$

$$F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (15.89)$$

Neglecting interactions, we write down the field equations in a linear approximation with respect to the differences with the classical solution:

$$(\square - m_0^2)h = 0, \quad (15.90)$$

$$\partial_\nu F_{\mu\nu} = -m_1^2 B_\mu, \quad \partial_\mu B_\mu = 0, \quad (15.91)$$

where  $m_0$  and  $m_1$  can be expressed using  $\eta$  as

$$m_0 = f\eta, \quad m_1 = e\eta. \quad (15.92)$$

Hence, the scalar field has mass  $m_0$ , the vector field  $B_\mu$  has mass  $m_1$ , and the massless particle disappears. Although the original vector field only has transverse wave components and its mass is 0,  $B_\mu$  absorbs the Nambu–Goldstone field  $g$  as a longitudinal wave and it becomes a vector field with finite mass. This is called the *Higgs mechanism*. Because of this, the massless Nambu–Goldstone particle disappears and the gauge quantum obtains mass.

In order to analyze the interactions, we write down the Lagrangian density. The free part, corresponding to (15.90) and (15.91), is

$$\mathcal{L}_f = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu} - \frac{1}{2}m_1^2 B_\mu B_\mu - \frac{1}{2}(\partial_\mu h \partial_\mu h + m_0^2 h^2). \quad (15.93)$$

The interaction part is

$$\begin{aligned} \mathcal{L}_{\text{int}} = eA_\mu(h\partial_\mu g - g\partial_\mu h) - em_1 h A_\mu A_\mu - \frac{1}{2}f m_0 h(g^2 + h^2) \\ - \frac{1}{2}e^2 A_\mu A_\mu (g^2 + h^2) - \frac{f^2}{8}(g^2 + h^2)^2. \end{aligned} \quad (15.94)$$

The commutation relations in the interaction picture can be derived from the free part as

$$[B_\mu(x), B_\nu(y)] = i \left( \delta_{\mu\nu} - \frac{\partial_\mu \partial_\nu}{m_1^2} \right) \Delta(x - y; m_1^2), \quad (15.95)$$

$$[h(x), h(y)] = i\Delta(x - y; m_0^2). \quad (15.96)$$

This finite-mass scalar field  $h$  is called the *Higgs field* and its quantum is called the *Higgs particle*. To quantize the gauge field, we need gauge conditions. We thus need to impose constraints on the gauge field classically or add additional terms in the Lagrangian to determine the gauge. In this section, we adopt the former approach and impose the Coulomb gauge condition

$$\partial \cdot A + (n \cdot \partial)(n \cdot A) = 0. \quad (15.97)$$

In this gauge, we can obtain the field  $g$  from (15.88):

$$g = -m_1 [\square + (n \cdot \partial)^2]^{-1} [\partial \cdot B + (n \cdot \partial)(n \cdot B)]. \quad (15.98)$$

Thus, we can also express  $A_\mu$  in terms of  $B_\mu$ . Using this,

$$\begin{aligned}
 [A_\mu(x), A_\nu(y)] &= i \left\{ \delta_{\mu\nu} - [(n_\mu \partial_\nu + n_\nu \partial_\mu)(n \cdot \partial) + \partial_\mu \partial_\nu] [\square + (n \cdot \partial)^2]^{-1} \right\} \\
 &\quad \times \Delta(x - y; m_1^2) , \\
 [A_\mu(x), g(y)] &= i m_1 n_\mu (n \cdot \partial) [\square + (n \cdot \partial)^2]^{-1} \Delta(x - y; m_1^2) , \\
 [g(x), g(y)] &= i (n \cdot \partial)^2 [\square + (n \cdot \partial)^2]^{-1} \Delta(x - y; m_1^2) , \\
 [B_\mu(x), g(y)] &= i m_1^{-1} [n_\mu \square - (n \cdot \partial) \partial_\mu] (n \cdot \partial) [\square + (n \cdot \partial)^2]^{-1} \Delta(x - y; m_1^2) .
 \end{aligned} \tag{15.99}$$

The local charge conservation law holds true:

$$\partial_\mu \mathbf{j}_\mu = 0 . \tag{15.100}$$

However, when the  $U(1)$  symmetry is broken, it is clear from (15.80) and (15.91) that

$$\mathbf{j}_\mu(x) = -m_1^2 \mathbf{B}_\mu(x) + \text{non-linear term} . \tag{15.101}$$

In this case, as discussed in Sect. 15.1, since a term proportional to  $1/\sqrt{V}$  is included in  $\mathbf{j}_0(x)$ , its spatial integral  $Q$  diverges when  $V \rightarrow \infty$ :

$$Q = \int d^3x \mathbf{j}_0(x) \sim \sqrt{V} \rightarrow \infty . \tag{15.102}$$

Moreover, the integral itself is the Fourier transform of  $B_0(x)$  and corresponds to the component of  $p = 0$ , so it has the time-dependence  $\exp(\pm i m_1 t)$ . It cannot therefore be a global conserved quantity. This means that the global conservation law is lost.

We shall now check whether the discussion at the beginning of this section is valid, i.e., whether the massless boson disappears under the covariant gauge condition. We compute the following commutator to lowest order:

$$\begin{aligned}
 i\langle 0 | [j_\mu(x), \varphi_1(y)] | 0 \rangle &\approx -i m_1^2 \langle 0 | [B_\mu(x), g(y)] | 0 \rangle \\
 &= m_1 [n_\mu \square - (n \cdot \partial) \partial_\mu] (n \cdot \partial) [\square + (n \cdot \partial)^2]^{-1} \Delta(x - y; m_1^2) .
 \end{aligned}$$

Its Fourier transform is thus proportional to the quantity

$$[(n \cdot k) k_\mu - k^2 n_\mu] (n \cdot k) [k^2 + (n \cdot k)^2]^{-1} \epsilon(k_0) \delta(k^2 + m_1^2) . \tag{15.103}$$

Comparing this with the general form (15.69) given at the beginning of this section, this term corresponds to the case where we only have  $\rho_5$ , while  $\rho_4$  disappears. This is therefore a counterexample to the Nambu–Goldstone theorem (15.70). If the symmetry is not broken spontaneously,  $\eta \rightarrow 0$  and  $m_1 = e\eta \rightarrow 0$ . In this case, (15.103) once more has the form

$$k_\mu \in (k_0 \delta(k^2)) , \quad (15.104)$$

whence the Nambu–Goldstone theorem is valid again, and  $g$  becomes a massless field.

In the treatment above, the explicit Lorentz covariance of the theory has been lost. Additionally, an extra field  $g$  has come in. We shall now present a formalism which conserves the Lorentz covariance and does not include any extra field. Higgs applied a transformation from the Cartesian coordinate system to the polar coordinate system to the field operators:

$$\varphi_1 = R \cos \Theta , \quad \varphi_2 = R \sin \Theta , \quad A_\mu = B_\mu - \frac{1}{e} \partial_\mu \Theta . \quad (15.105)$$

The gauge transformation is

$$\Theta(x) \rightarrow \Theta(x) - \lambda(x) . \quad (15.106)$$

With this transformation, the Lagrangian can be written in the form

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu} - \frac{1}{2} e^2 B_\mu^2 R^2 - \frac{1}{2} \partial_\mu R \partial_\mu R + \frac{1}{4} m_0^2 R^2 - \frac{1}{8} f^2 R^4 . \quad (15.107)$$

The extra degree of freedom has clearly been eliminated.  $R$  is the Higgs field, and the solution which spontaneously breaks the symmetry is

$$\langle \mathbf{B}_\mu \rangle = 0 , \quad \langle \mathbf{R} \rangle = \eta . \quad (15.108)$$

These show that  $\mathbf{B}_\mu$  once again becomes a vector field with mass  $m_1$ , while  $\mathbf{R}$  becomes a scalar field with mass  $m_0$ . Obviously, the massless Nambu–Goldstone particle disappears completely. Although this proof seems to be very clear, it is quite classical. Quantum-mechanically, the transformation (15.105) is quite non-linear and has a strong specificity. Additionally, in the form (15.107), we cannot apply the renormalization theory, and computations of higher-order corrections become impossible. Thus, although we have lots of gauge choices, every case has both advantages and disadvantages. Since an indefinite metric does not appear in the gauge mentioned here, this is called the *unitarity gauge*.

## 15.5 Higgs Mechanism with Covariant Gauge Condition

In the last section, we saw that the Nambu–Goldstone particle does not appear under a non-covariant gauge condition. If the physical contents of the gauge theory do not depend on how we choose the gauge, then under the covariant gauge condition, it turns out that the appearance of this massless particle cannot be observed, even though it is inevitable, as proven in Sect. 15.1. So, by what kind of mechanism is this particle concealed? Finding a solution to this problem will be the main theme of this section.

In order to discuss this issue, we begin with a formal discussion. It sometimes happens that the problem is simplified in the gauge theory by choosing the Landau gauge ( $\alpha = 0$ ). However, the Lagrangian density (5.121) diverges if we take  $\alpha = 0$ . Hence, following Utiyama and Nakanishi, we introduce an auxiliary field  $B$ , and rewrite (5.121) in the form

$$\mathcal{L}_{\text{em}} = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu} - \partial_\mu B A_\mu + \frac{\alpha}{2}B^2. \quad (15.109)$$

Varying  $B$ , we obtain

$$\partial_\mu A_\mu = \alpha B, \quad (15.110)$$

and inserting this into (15.109), we see that it is equivalent to the original (5.121). Taking this form, there is no confusion if we choose the Landau gauge and set  $\alpha = 0$ .

Although in the last section, in order to quantize the vector field, we eliminated the extra components by imposing the Coulomb gauge condition (15.97) in the classical theory, in the present section, we shall add a further term called a *gauge-fixing term* to (15.109) and quantize in a covariant way. Therefore, the free part of the Lagrangian density (15.93) is

$$\mathcal{L}_f = \mathcal{L}_1 + \mathcal{L}_2, \quad (15.111)$$

$$\mathcal{L}_1 = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu} - \frac{1}{2}(m_1 A_\mu - \partial_\mu g)(m_1 A_\mu - \partial_\mu g) + B\partial_\mu A_\mu, \quad (15.112)$$

$$\mathcal{L}_2 = -\frac{1}{2}(\partial_\mu h\partial_\mu h + m_0^2 h^2), \quad (15.113)$$

and the interaction part is again given by (15.94). Hence, quantizing the independent fields  $A_\mu$ ,  $B$ ,  $g$ , and  $h$  appearing in the free part, in the interaction picture, and expressing each of them in terms of the irreducible fields, we have from (15.112),

$$A_\mu^{\text{in}} = U_\mu^{\text{in}} - \frac{1}{m_1}\partial_\mu g_{(-)}^{\text{in}}, \quad B^{\text{in}} = m_1[g_{(+)}^{\text{in}} + g_{(-)}^{\text{in}}], \quad g^{\text{in}} = g_{(+)}^{\text{in}}, \quad (15.114)$$

where  $h$  is already irreducible. These asymptotic fields have the properties

$$\begin{aligned} (\square - m_1^2)H_\mu^{\text{in}} = 0, \quad \partial_\mu U_\mu^{\text{in}} = 0, \\ \langle 0|T^*[U_\mu^{\text{in}}(x), U_\nu^{\text{in}}(y)]|0\rangle = \left(\delta_{\mu\nu} - \frac{\partial_\mu\partial_\nu}{m_1^2}\right)\Delta_F(x-y; m_1^2). \end{aligned} \quad (15.115)$$

This means that the asymptotic field  $U_\mu^{\text{in}}$  has mass  $m_1$  and is a spin-1 field. Moreover, it is clear from

$$\begin{aligned} \square g_{(+)}^{\text{in}} = \square g_{(-)}^{\text{in}} = 0, \\ \langle 0|T[g_{(\pm)}^{\text{in}}(x), g_{(\pm)}^{\text{in}}(y)]|0\rangle = \pm D_F(x-y), \end{aligned} \quad (15.116)$$

that the asymptotic fields  $g_{(\pm)}^{\text{in}}$  are the fields of the Nambu–Goldstone particles, and are a mixture of the one with the positive-definite metric and the one with the indefinite metric. Finally, the asymptotic field  $h^{\text{in}}$  satisfies

$$(\square - m_0^2)h^{\text{in}} = 0, \quad \langle 0|T[h^{\text{in}}(x), h^{\text{in}}(y)]|0\rangle = \Delta_F(x-y; m_0^2). \quad (15.117)$$

This corresponds to the Higgs particle.

What is important here is that

$$\square \mathbf{B} = 0, \quad (15.118)$$

as a consequence of the equations of motion, as in the case of QED. Therefore, we can make the identification

$$\mathbf{B}(x) = B^{\text{in}}(x) = B^{\text{out}}(x). \quad (15.119)$$

In this case, as in the case of QED,  $\mathbf{B}$  commutes with the  $S$ -matrix:

$$[S, \mathbf{B}(x)] = 0. \quad (15.120)$$

Since this theory has an indefinite metric, following the argument in Chap. 5, we introduce the physical state by the subsidiary condition

$$B^{(+)}(x)|\text{phys}\rangle = 0. \quad (15.121)$$

At different space-time points, there is no contradiction here because

$$[\mathbf{B}(x), \mathbf{B}(y)] = 0. \quad (15.122)$$



To make a physical state from a different physical state including the vacuum, we have to act with operators that commute with  $\mathbf{B}$ , which leaves only

$$U_{\mu}^{\text{in}}, \quad h^{\text{in}}, \quad \mathbf{B}. \quad (15.123)$$

If we act with  $\mathbf{B}$ , a zero norm state is created. After all, considering the first two operators in (15.123), only the positive-definite and finite-mass particles survive. This is the Higgs mechanism which conceals the Nambu–Goldstone particle.

Although we have discussed the non-renormalized interaction picture here, the situation is essentially the same as in the renormalized interaction picture. Such an argument has been given by Nakanishi [148], but it has been slightly modified by the author here.

So far in this chapter, we have considered only field operators whose vacuum expectation values do not disappear, such as (15.1). Basically, if (15.10) holds, the Nambu–Goldstone theorem holds for  $\varphi$ , whether it be the field operator itself or one of its local polynomials. If (15.1) does not hold for the field operator itself, but holds for a local polynomial, we call this *dynamical symmetry breaking*. In this case, there are some examples where the Higgs mechanism works. It should be noted that, in 1963, Anderson made the same argument as Higgs for plasma oscillations [149].

## 15.6 Kibble’s Theorem

In the last section, we discussed the gauge theory corresponding to the Abelian group  $U(1)$ . We will now present Kibble’s theorem, which is an extension of the Higgs theorem for the Abelian group to the case of a non-Abelian group.

To investigate the Higgs mechanism, we first considered ways to set the gauge condition within the classical theory and to apply the unitarity gauge. This is more like a classical theoretical method. In contrast, in the last section, we quantized the field in the covariant gauge, and introducing an additional condition, we eliminated the indefinite metric. From the point of view of quantum field theory, the latter method is preferable, but it is then very difficult to quantize the non-Abelian gauge theory in the covariant gauge, something we shall discuss in a later chapter. Therefore, in this section we apply the former method and accept the limitations of a quasi-classical treatment. Here we present the demonstration due to Kibble [150], in which the unitarity gauge is introduced for the non-Abelian gauge field.

### 15.6.1 Adjoint Representation

An infinitesimal transformation of field operators for a given group  $G$  is given by (15.2). In the following, distinguishing upper indices from lower indices, we

shall write

$$\delta\varphi^A = \epsilon^a T_{aB}^A \varphi^B . \quad (15.124)$$

We also define the structure constants  $f_{ab}^c$  by

$$[T_a, T_b]^A{}_B = f_{ab}^c T_{cB}^A . \quad (15.125)$$

Thus, from the anti-commutativity of commutators and the Jacobi identity, the structure constants should satisfy

$$f_{ab}^c = -f_{ba}^c , \quad (15.126)$$

$$f_{ab}^m f_{mc}^l + f_{bc}^m f_{ma}^l + f_{ca}^m f_{mb}^l = 0 . \quad (15.127)$$

We introduce the matrix  $M$  by

$$(M_c)^a{}_b = f_{cb}^a . \quad (15.128)$$

Therefore, from (15.127)

$$[M_a, M_b]^l{}_c = f_{ab}^m (M_m)^l{}_c . \quad (15.129)$$

Hence,  $M$  is a representation of the algebra corresponding to the group  $G$ , called the *adjoint representation*. To prove Kibble's theorem, we need to know certain things about this representation, so we list some relevant formulas:

1. Taking  $\lambda$  as infinitesimal and writing  $\lambda^a T_a$  simply as  $\lambda T$ ,

$$(1 + \lambda T) T_d (1 - \lambda T) = T_b (1 + \lambda M)^b{}_d . \quad (15.130)$$

This is because the coefficient of  $\lambda^c$  on the left-hand side is

$$[T_c, T_d] = f_{cd}^b T_b = T_b (M_c)^b{}_d ,$$

which is equal to the coefficient of  $\lambda^c$  on the right-hand side. Thus, if  $\lambda$  is finite,

$$e^{\lambda T} T_d e^{-\lambda T} = T_b (e^{\lambda M})^b{}_d . \quad (15.131)$$

2. Next, using Feynman's formula (without proof),

$$\begin{aligned}
 \partial_\mu e^{\theta T} &= \int_0^1 d\alpha e^{(1-\alpha)\theta T} T_d e^{\alpha\theta T} \partial_\mu \theta^d \\
 &= e^{\theta T} \int_0^1 d\alpha e^{-\alpha\theta T} T_d e^{\alpha\theta T} \partial_\mu \theta^d \\
 &= e^{\theta T} T_d \left( \int_0^1 d\alpha e^{-\alpha\theta M} \right)^b{}_d \partial_\mu \theta^d \\
 &= e^{\theta T} T_d \Lambda^b{}_d \partial_\mu \theta^d, \tag{15.132}
 \end{aligned}$$

where

$$\Lambda = \frac{1 - e^{-\theta M}}{\theta M}. \tag{15.133}$$

We now define the covariant derivative by

$$D_\mu = \partial_\mu - \mathcal{A}_\mu = \partial_\mu - T A_\mu = \partial_\mu - T_a A_\mu^a. \tag{15.134}$$

Note that the coupling constant has been included in  $T$ , and therefore in the structure constants and also  $M$ , and  $\mathcal{A}$  is equal to what we wrote as  $ig\mathcal{A}_\mu$  in our previous notation. We can now implement the gauge transformation:

$$\mathcal{A}_\mu(x) \rightarrow U(\lambda(x)) \mathcal{A}_\mu(x) U^{-1}(\lambda(x)) - U(\lambda(x)) \partial_\mu U^{-1}(\lambda(x)). \tag{15.135}$$

In the following, replacing  $\exp(\theta T)$  by  $U(\lambda(x))$ , the *inverse* of the transformation above can be written as

$$\begin{aligned}
 \mathcal{A}_\mu &= e^{\theta T} T_a A_\mu^a e^{-\theta T} - e^{\theta T} \partial_\mu e^{-\theta T} \\
 &= T_b (e^{\theta M} A'_\mu)^b - T_b \left( \frac{1 - e^{\theta M}}{\theta M} \partial_\mu \theta \right)^b \\
 &= T_b (e^{\theta M} A'_\mu + e^{\theta M} \Lambda \partial_\mu \theta)^b. \tag{15.136}
 \end{aligned}$$

Thus,

$$A_\mu^A = (e^{\theta M})^A{}_B A'^B_\mu + (e^{\theta M} \Lambda)^A{}_b \partial_\mu \theta^b. \tag{15.137}$$

In the equation above, we have distinguished capital letters and small letters, and we will discuss this difference later.

### 15.6.2 Kibble's Theorem

The Lagrangian of the scalar field which is invariant under a group  $G$  is written in the form

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\varphi_\mu - \partial_\mu\varphi)^2 - \frac{1}{2}(\partial_\mu\varphi)^2 - U(\varphi) \\ &= -\varphi_\mu\partial_\mu\varphi + \frac{1}{2}\varphi_\mu\varphi_\mu - U(\varphi),\end{aligned}\tag{15.138}$$

without specifying the components explicitly.<sup>1</sup> Writing it in this form, the Lagrangian is linear in the derivatives, and the quantity canonically conjugate to  $\varphi$  is  $-\varphi_0$ . Using the effective potential  $V(\varphi)$ , the formula which determines the spontaneous symmetry breaking is

$$\frac{\partial V(\varphi)}{\partial\varphi} = 0.\tag{15.139}$$

To lowest order,  $V(\varphi)$  can be replaced by  $U(\varphi)$ . We assume that one of the solutions is

$$\varphi = \eta.\tag{15.140}$$

Therefore, in general,  $e^{\lambda T}\eta$  is also a solution and the solutions are degenerate. However,  $\lambda$  is a constant. If we write the subgroup of  $G$  which keeps  $\eta$  invariant as  $G_\eta$ , i.e.,

$$e^{\lambda T}\eta = \eta,\tag{15.141}$$

then it turns out that the symmetry corresponding to  $G_\eta$  is not broken for this solution. The number of invariants which are algebraically independent and constructed by  $\eta$  is denoted  $\nu$ , which we call the *canonical number*. For instance, for  $O(3)$ , the invariant

$$\eta^2 = \eta_1^2 + \eta_2^2 + \eta_3^2$$

is the unique invariant for the three-dimensional  $\eta$  and in this case  $\nu = 1$ . We now assume that  $T$  is an  $n$ -dimensional representation and therefore that both  $\varphi$  and  $\eta$  are  $n$ -dimensional vectors. In this case, we choose the canonical form such that only  $\nu$  components of  $\eta$  are non-zero. That is, we eliminate the other components by a transformation in  $G$ . For instance, in the Higgs model,  $G = U(1)$ ,  $n = 2$ , and

---

<sup>1</sup> In this section, the notation  $\phi_\mu$  means  $\partial_\mu\phi$ , following the original article by T. Kibble [150].

$\nu = 1$ . We consider the manifold constructed by the set of solutions, viz.,

$$\{e^{\lambda T} \eta | e^{\lambda T} \in G\} . \quad (15.142)$$

We shall now determine its dimension.

Although  $\eta$  is  $n$ -dimensional, since the values of the  $\nu$  invariants made from  $\eta$  cannot be changed, the number of parameters  $r$  which can be chosen freely is given by  $r = n - \nu$ . Therefore, the dimension of the above set is  $r$ . Note also that the number of components of  $\lambda$  is equal to the number of the generators  $T$ , so that it becomes the order of  $g$ . Since  $r$  parameters among them play the role of designating  $\eta$ , it turns out that  $g - r$  do not change  $\eta$ . Thus, this is also the dimension of  $G_\eta$ .

In the following, fixing the form of  $\eta$ , the general form of the vacuum expectation value that solves (15.139) is

$$\varphi = e^{\underset{(r)}{\mu T}} e^{\underset{(g-r)}{\nu T}} = \eta = e^{\mu T} , \quad e^{\nu T} \in G_\eta , \quad (15.143)$$

where the dimensions are indicated under the two factors. Introducing the angular variable  $\theta$  and the radial variable  $\rho$ , as a generalization of the unitarity gauge in the Higgs model, we take

$$\varphi = e^{\theta T} \rho = e^{\theta T} (\eta + \rho') , \quad (15.144)$$

where  $\theta$  is  $r$ -dimensional. For the canonical form, we have

$$\theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \quad \rho = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \rho_1 \\ \vdots \\ \rho_\nu \end{pmatrix} , \quad (15.145)$$

where the total number of components of  $\theta$  is  $g$  and  $r = n - \nu$ , while  $\rho$  has altogether  $n$  components, of which the first  $r$  are equal to zero.

For the generators  $T_A$ ,  $A = 1, 2, \dots, g - r$ , of  $G_\eta$ , we assume that  $\rho$  satisfies

$$T_A \rho = 0 , \quad (15.146)$$

while for the other  $T_a$ ,  $a = 1, 2, \dots, r$ ,

$$T_a \rho \neq 0 . \quad (15.147)$$

In the following, we express the  $r$  dimensions of  $\theta$  by small Roman letters and  $\nu$  dimensions of  $\rho$  by Greek letters. We define the  $r$ -dimensional matrix  $X$  by

$$X_a{}^b = (T_a\rho)^b = T_{a\alpha}^b\rho^\alpha. \quad (15.148)$$

This matrix is invertible, because if  $\det X = 0$ , then there exist  $c^a$  satisfying

$$c^a T_a\rho = 0,$$

whence  $c^a T_a$  becomes a generator of  $G_\eta$ , which contradicts our assumptions.

We now rewrite  $\varphi$  in the Lagrangian density in terms of  $\theta$  and  $\rho$  using (15.144). With (15.132), we obtain

$$\partial_\mu\varphi^A = (e^{\theta T} T_b\rho)^A \Lambda^b{}_d \partial_\mu\theta^d + (e^{\theta T})^A{}_\beta \partial_\mu\rho^\beta, \quad (15.149)$$

$$\varphi_\mu \partial_\mu\varphi = (\varphi_\mu e^{\theta T} T_b\rho) \Lambda^b{}_d \partial_\mu\theta^d + \varphi_\mu e^{\theta T} \partial_\mu\rho. \quad (15.150)$$

We denote the coefficients of the derivatives on the right-hand side of (15.150) as follows:

$$\theta_{\mu d} \equiv (\varphi_\mu e^{\theta T} T_b\rho) \Lambda^b{}_d, \quad \rho_{\mu\alpha} \equiv \varphi_{\mu A} (e^{\theta T})^A{}_\alpha. \quad (15.151)$$

Now  $\varphi$  has  $n$  components, but  $\rho$  has only  $\nu$  components. Thus,

$$\rho_{\mu\alpha} \rho_\mu{}^\alpha = \varphi_{\mu A} \varphi_\mu{}^A - \varphi_{\mu A} (e^{\theta T})^A{}_a (e^{-\theta T})^a{}_A (e^{-\theta T})^a{}_B \varphi_\mu{}^B. \quad (15.152)$$

However,

$$\begin{aligned} \theta_{\mu d} &= \varphi_{\mu A} (e^{\theta T})^A{}_\alpha T_{b\beta}^\alpha \rho^\beta \Lambda^b{}_d + \varphi_{\mu A} (e^{\theta T})^A{}_a T_{b\beta}^a \rho^\beta \Lambda^b{}_d \\ &= \rho_{\mu\alpha} T_{b\beta}^\alpha \rho^\beta \Lambda^b{}_d + \varphi_{\mu A} (e^{\theta T})^A{}_a X_b{}^a \Lambda^b{}_d. \end{aligned} \quad (15.153)$$

Therefore,

$$\theta_{\mu d} (\Lambda^{-1} X^{-1})^d{}_a = (\rho_{\mu\alpha} T_b\rho) (X^{-1})^b{}_a + \varphi_{\mu A} (e^{\theta T})^A{}_a. \quad (15.154)$$

If we express the difference between  $\rho\rho$  and  $\varphi\varphi$  in (15.152) using (15.154),

$$\begin{aligned} \mathcal{L} &= -\rho_\mu \partial_\mu\rho + \frac{1}{2} \rho_\mu \rho_\mu - U(\rho) - \theta_\mu \partial_\mu\theta \\ &\quad + \frac{1}{2} \left\{ [\theta_{\mu a} (\Lambda^{-1})^a{}_b - \rho^\mu T_b\rho] (X^{-1})^b{}_c \right\}^2. \end{aligned} \quad (15.155)$$

In the Lagrangian density above,  $\theta$  always appears together with  $\theta_\mu$  and never alone, so we see that  $\theta$  becomes a massless Nambu–Goldstone particle. Now,  $\rho$  and  $\rho_\mu$  are gauge invariant, and the gauge transformations of  $\theta$  and  $\theta_\mu$  will be derived below.

From (15.144),  $\varphi$  and  $e^{\theta T}$  transform in the same way under the gauge transformation, so considering  $\delta$  as the gauge transformation corresponding to the infinitesimal gauge function  $\delta\lambda^A$ , we have

$$\delta(e^{\theta T}) = T_A \delta\lambda^A e^{\theta T} . \quad (15.156)$$

Therefore, from (15.132),

$$\delta(e^{\theta T}) = e^{\theta T} T_b \Lambda^b{}_d \delta\theta^d . \quad (15.157)$$

Combining the two equations above,

$$T_b \Lambda^b{}_d \delta\theta^d = e^{-\theta T} T_A e^{\theta T} \delta\lambda^A = T_b (e^{-\theta M})^b{}_A \delta\lambda^A .$$

Therefore,

$$\delta\theta^a = (\Lambda^{-1})^a{}_b (e^{-\theta M})^b{}_A \delta\lambda^A . \quad (15.158)$$

Although the gauge transformation for  $\theta_{\mu a}$  is not necessary for the following discussion, we shall give it for the record. Because  $\theta^a$  and  $-\theta_{0a}$  are canonical conjugates,

$$\delta[\theta_{0a}, \theta^b] = [\delta\theta_{0a}, \theta^b] + [\theta_{0a}, \delta\theta^b] .$$

Alternatively,

$$i \frac{\partial}{\partial\theta_{0b}} \delta\theta_{0a} + i \frac{\partial}{\partial\theta^a} \delta\theta^b = 0 , \quad \delta\theta_{\mu a} = -\theta_{\mu b} \frac{\partial}{\partial\theta^a} \delta\theta^b . \quad (15.159)$$

We now introduce a gauge field into the Lagrangian density in polar form. In this case, the covariant derivative is obtained by the replacement

$$D_\mu \varphi = \partial_\mu \varphi - \delta\varphi(\lambda \rightarrow A_\mu) . \quad (15.160)$$

In other words, we replace the infinitesimal gauge function  $\lambda$  by  $A_\mu$  to obtain the second term on the right-hand side. As already noted, the coupling constant is included in the structure constant. So writing

$$f_{BC}^A = -gf_{ABC} , \quad (15.161)$$

the Lagrangian density including the gauge field is

$$\begin{aligned}
 \mathcal{L} = & -\frac{1}{2}F_{\mu\nu,A}(\partial_\mu A_\nu^A - \partial_\nu A_\mu^A + gf_{ABC}A_\mu^B A_\nu^C) + \frac{1}{4}F_{\mu\nu}F_{\mu\nu} \\
 & -\rho_\mu\partial_\mu\rho + \frac{1}{2}\rho_\mu\rho_\mu - U(\rho) \\
 & + \frac{1}{2}\left\{[\theta_{\mu a}(\Lambda^{-1})^a{}_b - \rho_\mu T_b\rho](X^{-1})^b{}_c\right\}^2 \\
 & -\theta_\mu(\partial_\mu\theta - \Lambda^{-1}e^{-\theta M}A_\mu), \tag{15.162}
 \end{aligned}$$

where the last term comes from introducing the covariant derivative for  $\theta$  by the replacement (15.160).

Now, from the gauge transformation (15.137), we express the gauge field  $A_\mu$  in terms of a new gauge field  $A'_\mu$ :

$$A_\mu^A = (e^{\theta M})^A{}_B A'_\mu{}^B + (e^{\theta M}\Lambda)^A{}_b\partial_\mu\theta^b, \quad F = e^{\theta M}F'. \tag{15.163}$$

Making this replacement, the last term in (15.162) becomes

$$\partial_\mu\theta - \Lambda^{-1}e^{-\theta M}A_\mu = \partial_\mu\theta - \Lambda^{-1}e^{-\theta M}e^{\theta M}(A'_\mu + \Lambda\partial_\mu\theta) = -\Lambda^{-1}A'_\mu, \tag{15.164}$$

whence the term in  $\partial_\mu\theta$  vanishes! This is exactly the same as in the Higgs mechanism. As a result, (15.162) becomes

$$\begin{aligned}
 \mathcal{L} = & -\frac{1}{2}F'_{\mu\nu,A}(\partial_\mu A'_\nu{}^A - \partial_\nu A'_\mu{}^A + gf_{ABC}A'_\mu{}^B A'_\nu{}^C) + \frac{1}{4}F'_{\mu\nu}F'_{\mu\nu} \\
 & -\rho_\mu\partial_\mu\rho + \frac{1}{2}\rho_\mu\rho_\mu - U(\rho) \\
 & + \frac{1}{2}\left\{[\theta_{\mu a}(\Lambda^{-1})^a{}_b - \rho_\mu T_b\rho](X^{-1})^b{}_c\right\}^2 + \theta_\mu\Lambda^{-1}A'_\mu. \tag{15.165}
 \end{aligned}$$

Since  $\partial_\mu\theta$  does not appear here,  $\theta_\mu$  is expressed algebraically in terms of the other quantities. If we now set

$$\theta_{\mu a}(\Lambda^{-1}X^{-1})^a{}_b \equiv h_{\mu b}, \quad \rho_\mu T_a\rho(X^{-1})^a{}_b \equiv k_{\mu b}, \tag{15.166}$$

then only  $\theta_{\mu a}$  involves  $h_{\mu a}$ , and the sum of the last two terms in (15.165) becomes

$$\frac{1}{2}(h_\mu - k_\mu)^2 + h_\mu X A'_\mu = \frac{1}{2}(h_\mu - k_\mu + X A'_\mu)^2 + k_\mu X A'_\mu - \frac{1}{2}(X A'_\mu)^2.$$

Hence, considering  $h_\mu$  as an independent variable and using the variation principle, the first term on the right-hand side in the above equation can be dropped. Thus, the



final Lagrangian density takes the form

$$\begin{aligned} \mathcal{L} = & -\frac{1}{2}F'_{\mu\nu,A}(\partial_\mu A'_\nu{}^A - \partial_\nu A'_\mu{}^A + gf_{ABC}A'^B{}_\mu A'^C{}_\nu) + \frac{1}{4}F'_{\mu\nu}F'_{\mu\nu} \\ & -\rho_\mu\partial_\mu\rho + \frac{1}{2}\rho_\mu\rho_\mu - U(\rho) + (\rho_\mu T_a\rho)A'^a{}_\mu - \frac{1}{2}(X^a{}_b A'^b{}_\mu)^2. \end{aligned} \quad (15.167)$$

This means that  $\theta$  has been completely eliminated and the scalar field is described solely by  $\rho$ . As in the case of the Abelian Higgs model, it turns out that the unitarity gauge has been realized.

Although the number of  $\theta$  components corresponding to Nambu–Goldstone particles was  $r$ , they were absorbed into the massless gauge particles as longitudinal modes and gave finite masses to the gauge particles. Moreover,  $\rho$  became  $\nu$  scalar particles with finite masses, the so-called *Higgs particles*. It turns out that massless gauge particles are left, corresponding to  $g - r$  generators for  $G_\eta$ . We summarize this as follows:

$\rho$ ( $\nu$ components)	Higgs particles with finite masses
$\theta$ ( $r$ components)	gauge particles with finite masses
$A$ ( $r$ components)	gauge particles with finite masses
$A$ ( $g - r$ components)	massless gauge particles

The above argument is the proof of Kibble's theorem, which is an extension of Higgs' theorem to non-Abelian gauge fields. As an example, Kibble discussed the model in which three-component complex scalar fields are introduced for an  $SU(2) \times U(1)$  gauge group. In the Weinberg–Salam model, the three-component fields are replaced by two-component complex scalar fields. This is the subject of the next chapter.

# Chapter 16

## Weinberg–Salam Model



As discussed in Chap. 13, the weak interaction is described phenomenologically by Fermi's universal four-fermion theory. One property of this interaction is that it can be expressed as a product of the current densities which form the Lie algebra. However, since this theory is not renormalizable, the cross-section increases without limit at high energies and also in higher order corrections, i.e., the higher the order, the greater the degree of divergence. Thus, we cannot avoid interpreting Fermi's theory as purely phenomenological, expressing the  $S$ -matrix element only for low energies and only at the lowest order. Many have tried to interpret Fermi's interaction, by introducing intermediate vector fields, as an iteration of Yukawa interactions. The vector field is necessary in order to reproduce a (V – A)-type interaction, but the theory is still not renormalizable. This led to the model by Weinberg and Salam, combining gauge fields with the Higgs–Kibble mechanism and describing the weak interaction and the electromagnetic interaction in a single theory. These authors proposed the model in [151] and in [152], just after the publication of Kibble's work. In fact, as explained at the end of the last section, Kibble himself considered almost the same model as this.

### 16.1 Weinberg–Salam Model

What kind of Lie algebra is needed to describe the weak interaction according to a gauge principle? As mentioned before, the spatial integrations over time-components of the currents  $J_\mu$  and  $J_\mu^\dagger$  carrying charges appeared in the weak interaction from the algebra  $SU(2)_L$ . Thus, we need to introduce the gauge group  $SU(2)$ . Moreover, we would like to introduce the electromagnetic field as a gauge field in order to describe the weak interaction and the electromagnetic interaction

simultaneously, so we need to start with the gauge group

$$SU(2) \times U(1) . \quad (16.1)$$

Recalling the formula,

$$Q = e \left( I_3 + \frac{Y}{2} \right) , \quad (16.2)$$

we consider  $I_i, i = 1, 2, 3$ , and  $Y$  in this equation as the generators for the algebra of the group (16.1). Here,  $I_i$  is the isospin. Originally, the isospin was considered only for hadrons, but this type of isospin was defined for both hadrons and leptons. This is obvious from (13.188), which expresses the universality of the weak interaction. Hence, the quantities  $I_i$  are called *weak isospin*.

The number of generators in (16.1) is  $g = 3 + 1$ . If we choose the Higgs field  $\varphi$  to be a two-component complex scalar field, then  $n = 4$ , and in this case there is only one invariant, viz.,

$$\varphi_1^\dagger \varphi_1 + \varphi_2^\dagger \varphi_2 , \quad (16.3)$$

so  $\nu = 1$ . Therefore, the number of vector particles with finite masses is  $r = n - \nu = 4 - 1 = 3$ , while  $g - r = 1$ , so there is one massless gauge boson, i.e., the photon field. We write the gauge fields corresponding to  $SU(2)$  and  $U(1)$  as  $\hat{W}_\mu$  and  $B_\mu$ , respectively. Therefore, considering that the Higgs field  $\varphi$  has weak isospin and weak hypercharge, the Lagrangian density for this system is

$$\begin{aligned} \mathcal{L} = & -\frac{1}{4} \left( \partial_\mu \hat{W}_\nu - \partial_\nu \hat{W}_\mu + g \hat{W}_\mu \times \hat{W}_\nu \right)^2 - \frac{1}{4} \left( \partial_\mu B_\nu - \partial_\nu B_\mu \right)^2 \\ & - \frac{1}{2} \varphi^\dagger \left( \overleftarrow{\partial}_\mu + ig \frac{\hat{\tau}}{2} \hat{W}_\mu - i \frac{g'}{2} B_\mu \right) \left( \overrightarrow{\partial}_\mu + ig \frac{\hat{\tau}}{2} \hat{W}_\mu - i \frac{g'}{2} B_\mu \right) \varphi \\ & - U(\varphi) , \end{aligned} \quad (16.4)$$

where we have assumed that both the transformation property and the electric charge of  $\varphi$  are the same as those of the K-meson doublet ( $K^+, K^0$ ). Following Kibble's theory, to express the four components of  $\varphi$  using  $\theta_i, i = 1, 2, 3$ , and  $\rho$ , we write

$$\varphi = e^{i\hat{\theta}\hat{\tau}/2} \begin{pmatrix} 0 \\ \rho \end{pmatrix} . \quad (16.5)$$

As mentioned earlier,  $\hat{\theta}$  is absorbed to give masses to the intermediate gauge particles. However, in the approximation which neglects the effects of  $\rho$ , we must

make the replacement

$$\varphi \rightarrow \begin{pmatrix} 0 \\ \eta \end{pmatrix}. \quad (16.6)$$

Then, setting  $\rho = \eta + \rho'$  and rewriting (16.4), in the unitarity gauge,

$$\begin{aligned} \mathcal{L}_U &= -\frac{1}{4} \left( \partial_\mu \hat{W}_\nu - \partial_\nu \hat{W}_\mu + g \hat{W}_\mu \times \hat{W}_\nu \right)^2 - \frac{1}{4} (\partial_\mu B_\nu - \partial_\nu B_\mu)^2 \\ &\quad - \frac{1}{4} g^2 \eta^2 W_\mu^\dagger W_\mu - \frac{1}{8} \eta^2 (g W_\mu^3 + g' B_\mu)^2 - U(\eta) \\ &\quad + \text{terms containing } \rho', \end{aligned} \quad (16.7)$$

where  $W_\mu$  and  $W_\mu^\dagger$  denote the charged intermediate gauge fields, and  $W_\mu^3$  denotes a neutral intermediate gauge field. Diagonalizing the mass terms, we define the combinations

$$Z_\mu = \frac{g W_\mu^3 + g' B_\mu}{\sqrt{g^2 + g'^2}}, \quad m_Z^2 = \frac{1}{4} \eta^2 (g^2 + g'^2), \quad (16.8)$$

$$W_\mu = \frac{W_\mu^1 - i W_\mu^2}{\sqrt{2}}, \quad m_W^2 = \frac{1}{4} \eta^2 g^2. \quad (16.9)$$

Moreover, the combination orthogonal to  $Z_\mu$  is the massless photon field:

$$A_\mu = \frac{-g' W_\mu^3 + g B_\mu}{\sqrt{g^2 + g'^2}}. \quad (16.10)$$

Thus,

$$W_\mu^3 = \frac{g Z_\mu - g' A_\mu}{\sqrt{g^2 + g'^2}}, \quad B_\mu = \frac{g' Z_\mu + g A_\mu}{\sqrt{g^2 + g'^2}}. \quad (16.11)$$

Using the expressions for  $Z_\mu$  and  $A_\mu$ , we rewrite  $\mathcal{L}_U$  in the form

$$\begin{aligned} \mathcal{L}_U &= -\frac{1}{2} W_{\mu\nu}^\dagger W_{\mu\nu} - m_W^2 W_\mu^\dagger W_\mu - \frac{1}{4} A_{\mu\nu} A_{\mu\nu} \\ &\quad - \frac{1}{4} Z_{\mu\nu} Z_{\mu\nu} - \frac{1}{2} m_Z^2 Z_\mu Z_\mu \\ &\quad \left. \begin{aligned} &-ig \sin \theta_W A_\mu (W_{\mu\nu}^\dagger W_\nu - W_{\mu\nu} W_\nu^\dagger - W_{\mu\nu} W_\nu^\dagger) \\ &+ g^2 \sin^2 \theta_W (A_\mu A_\nu W_\mu^\dagger W_\nu - A_\mu A_\nu W_\nu^\dagger W_\mu) \end{aligned} \right\} \text{(Dirac-type EM interaction)} \\ &\quad + ig \sin \theta_W A_{\mu\nu} W_\mu^\dagger W_\nu \quad \text{(anomalous magnetic moment)} \end{aligned}$$

$$\begin{aligned}
& +ig \cos \theta_W Z_\mu (W_{\mu\nu}^\dagger W_\nu - W_{\mu\nu} W_\nu^\dagger) \\
& +g^2 \cos^2 \theta_W (Z_\mu Z_\nu W_\mu^\dagger W_\nu - Z_\mu Z_\nu W_\nu^\dagger W_\mu) \\
& -ig \cos \theta_W Z_{\mu\nu} W_\mu^\dagger W_\nu \\
& +g^2 \sin \theta_W \cos \theta_W (2A_\mu Z_\mu W_\nu^\dagger W_\nu - A_\mu Z_\nu W_\mu^\dagger W_\nu - A_\nu Z_\mu W_\mu^\dagger W_\nu) \\
& +\frac{g^2}{2} (W_\mu^\dagger W_\mu^\dagger W_\nu W_\nu - W_\mu^\dagger W_\nu^\dagger W_\mu W_\nu) \\
& -U(\eta) + \text{terms containing } \rho' .
\end{aligned} \tag{16.12}$$

We have assumed that the first component of  $\varphi$  has an electric charge  $e$ , like  $K^+$ , and that the second component is neutral. Then it turns out that  $W$  has an electric charge  $e$ . From the electromagnetic interaction of  $W$  in the equation above,

$$e = -g \sin \theta_W , \tag{16.13}$$

where  $\theta_W$ , defined by

$$\frac{g'}{\sqrt{g^2 + g'^2}} = \sin \theta_W , \quad \frac{g}{\sqrt{g^2 + g'^2}} = \cos \theta_W , \tag{16.14}$$

is called the *Weinberg angle*. The angle is also called the *weak mixing angle*, introduced by S. Glashow in his 1961 paper [153]. A symbol like  $A_{\mu\nu}$  means  $A_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ . Combining (16.8), (16.9), and (16.14),

$$m_W = m_Z \cos \theta_W . \tag{16.15}$$

Experimentally,

$$\sin^2 \theta_W = 0.22 . \tag{16.16}$$

It turns out that Fermi's weak interaction is produced by relating  $J_\mu$  to  $J_\mu^\dagger$  via  $W$ . Like the nuclear force, this is a combination of two Yukawa interactions. At energies much lower than  $m_W$ , this reduces to the Fermi interaction. Comparing the gauge theory and the Fermi four-fermion interaction at low energies, we obtain

$$\frac{G_F}{\sqrt{2}} = \left( \frac{g}{2\sqrt{2}} \right)^2 \frac{1}{m_W^2} = \frac{e^2}{8m_W^2 \sin^2 \theta_W} . \tag{16.17}$$

From this,

$$m_W^2 = \frac{\sqrt{2}}{8} \frac{g^2}{G_F} > \frac{\sqrt{2}}{8} \frac{e^2}{G_F} = (37.3 \text{ GeV})^2 . \tag{16.18}$$

If the Weinberg angle is determined, it turns out that  $m_W$  and  $m_Z$  are also determined. Indeed, approximate values for these masses were predicted theoretically. In 1983 at CERN, the  $W$  and  $Z$  bosons were detected and their masses were also measured [154–157]:

$$m_W = 80.8 \pm 2.7 \text{ GeV} , \quad m_Z = 92.9 \pm 1.6 \text{ GeV} . \quad (16.19)$$

These values were quite close to the predictions.

According to the gauge theory, in the weak interaction, there are neutral and charged intermediate vector bosons, namely,  $Z^0$  and  $W^\pm$ , respectively. Weak interactions with neutral particles involve phenomena that cannot be described by the Fermi interaction, such as neutrino scattering. Such phenomena were confirmed early on using a large bubble chamber at CERN in [158].

## 16.2 Introducing Fermions

In the last section, we formulated the electroweak unification using only gauge fields and a Higgs field. However, a realistic model cannot be constructed without fermions like quarks and leptons. We thus begin by introducing leptons. According to the  $(V - A)$  theory, what appears in the weak interaction are left-handed leptons. Hence, what interacts with the gauge fields are three generations of left-handed leptons:

$$L_1 = \frac{1}{2}(1 + \gamma_5) \begin{pmatrix} \nu_e \\ e \end{pmatrix} , \quad L_2 = \frac{1}{2}(1 + \gamma_5) \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} , \quad L_3 = \frac{1}{2}(1 + \gamma_5) \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} . \quad (16.20)$$

On the other hand, right-handed leptons do not interact with  $SU(2)$  gauge fields. We therefore choose the Lagrangian density for the left-handed leptons to be

$$- \sum_j \bar{L}_j \gamma_\mu \left( \partial_\mu - i g \frac{\hat{\tau}}{2} \hat{W}_\mu - i \frac{g'}{2} B_\mu \right) L_j . \quad (16.21)$$

Regarding the interaction with  $\hat{W}_\mu$ , since  $L_i$  is a doublet like  $\varphi$ , we have chosen the same form. In contrast,  $Y$  given by (16.2) is  $+1$  for  $\varphi$  carrying the charge  $(+e, 0)$ , and  $-1$  for  $L$  carrying the charge  $(0, -1)$ . Thus, as for the interaction with  $B_\mu$ , we have chosen the sign opposite to that of  $\varphi$ . Regarding the neutrino, there is no right-handed component because the mass vanishes. Therefore, we write the right-handed component as

$$R_1 = \frac{1}{2}(1 - \gamma_5)e , \quad R_2 = \frac{1}{2}(1 - \gamma_5)\mu , \quad R_3 = \frac{1}{2}(1 - \gamma_5)\tau . \quad (16.22)$$

This time, since neutrinos are singlets, they do not interact with  $\hat{W}_\mu$ . However, since  $Y$  becomes  $-2$ , a plausible choice is

$$-\sum_j \bar{R}_j \gamma_\mu (\partial_\mu - ig' B_\mu) R_j . \quad (16.23)$$

The sum of (16.21) and (16.23) yields the Lagrangian density for leptons.

We shall now introduce quarks. Note that the quark doublets appearing in the weak interaction are not necessarily mass eigenstates. They are obtained by the rotation (13.192), introduced by Cabibbo, of the doublet of mass eigenstates. Experimentally, the angle  $\theta_C$  is given by

$$\sin \theta_C = 0.22 . \quad (16.24)$$

This value can be derived using (13.178), (13.150), and (13.180), whence the doublet of  $SU(2)_L$  is

$$q_L = \frac{1}{2}(1 + \gamma_5) \begin{pmatrix} u \\ d' \end{pmatrix} , \quad (16.25)$$

and the others are all singlets of  $SU(2)_L$ , i.e., they are  $u_R$ ,  $d_R$ ,  $s_R$ ,  $s_L$  and so on. However, it will be shown in the next section that this choice yields a constellation of difficulties.

As already mentioned in Sect. 13.6, Glashow, Iliopoulos, and Maiani proposed what is known as the GIM mechanism [130]. They introduced a fourth quark  $c$ , and formed two doublets of  $SU(2)_L$  :

$$\frac{1}{2}(1 + \gamma_5) \begin{pmatrix} u \\ d' \end{pmatrix} , \quad \frac{1}{2}(1 + \gamma_5) \begin{pmatrix} c \\ s' \end{pmatrix} . \quad (16.26)$$

One particular merit of this model will be discussed in the next section. In this model, although the interaction with  $\hat{W}_\mu$  is the same as that of the leptons, if we now compute  $Y$  from (16.2), we obtain  $1/3$ . The Lagrangian density including gauge interactions of the left-handed quarks for each generation is then

$$-\bar{q}_L \gamma_\mu \left( \partial_\mu - ig \frac{\hat{\tau}}{2} \hat{W}_\mu + i \frac{g'}{6} B_\mu \right) q_L . \quad (16.27)$$

On the other hand, since right-handed quarks are all singlets of  $SU(2)_L$ ,  $Y/2 = 2/3$  for quarks with charge  $2e/3$  and  $Y/2 = -1/3$  for quarks with charge  $-e/3$ . Hence, if

$$q_R = \frac{1}{2}(1 - \gamma_5)q , \quad (16.28)$$

then the Lagrangian density is given by

$$- \bar{q}_R \gamma_\mu \left( \partial_\mu + i \frac{g'}{2} Y B_\mu \right) q_R . \quad (16.29)$$

The sum of (16.27) and (16.29) gives the Lagrangian for quarks. Hence, the interactions of fermions and gauge fields are given by:

$$\mathcal{L}_{\text{IG}} = g \hat{W}_\mu \hat{j}_\mu - g' B_\mu (j_\mu^{\text{em}} - j_\mu^3) , \quad (16.30)$$

$$j_\mu^{\text{em}} = -i \bar{e} \gamma_\mu e - i \bar{\mu} \gamma_\mu \mu - i \bar{\tau} \gamma_\mu \tau + i \left( \frac{2}{3} \bar{u} \gamma_\mu u - \frac{1}{3} \bar{d} \gamma_\mu d + \dots \right) , \quad (16.31)$$

$$\hat{j}_\mu = \sum_j i \bar{L}_j \gamma_\mu \frac{\hat{\tau}}{2} L_j + \sum_q i \bar{q}_L \gamma_\mu \frac{\hat{\tau}}{2} q_L . \quad (16.32)$$

We rewrite these interactions using  $A_\mu$  and  $Z_\mu$  :

$$\mathcal{L}_{\text{IG}} = \frac{g}{\sqrt{2}} (W_\mu j_\mu^\dagger + W_\mu^\dagger j_\mu) - g \sin \theta_W j_\mu^{\text{em}} A_\mu + \frac{g}{\cos \theta_W} (j_\mu^3 - \sin^2 \theta_W j_\mu^{\text{em}}) Z_\mu . \quad (16.33)$$

The explicit forms of the various currents are:

$$j_\mu = i \bar{e} \gamma_\mu \frac{1 + \gamma_5}{2} \nu_e + \dots + i \bar{d}' \gamma_\mu \frac{1 + \gamma_5}{2} u + \dots , \quad (16.34)$$

$$j_\mu^\dagger = i \bar{\nu}_e \gamma_\mu \frac{1 + \gamma_5}{2} e + \dots + i \bar{u} \gamma_\mu \frac{1 + \gamma_5}{2} d' + \dots , \quad (16.35)$$

$$j_\mu^3 = \frac{1}{2} \left( i \bar{\nu}_e \gamma_\mu \frac{1 + \gamma_5}{2} \nu_e - i \bar{e} \gamma_\mu \frac{1 + \gamma_5}{2} e \right) + \dots \quad (16.36)$$

$$+ \frac{1}{2} \left( i \bar{u} \gamma_\mu \frac{1 + \gamma_5}{2} u - i \bar{d}' \gamma_\mu \frac{1 + \gamma_5}{2} d' \right) + \dots .$$

Although we have not included the fermion mass terms, these masses are generated via the Higgs mechanism. For this purpose, we introduce Yukawa-type interactions between the Higgs field and fermions. In fact, introducing these interactions contradicts the general idea that all interactions be gauge interactions, so we may have to consider some dynamical symmetry breaking in the future. Considering only the first generation, the interactions invariant under  $SU(2)_L$  are

$$\mathcal{L}_{\text{f}\phi} = -f (\bar{L}_1 \phi) R_1 - f_\uparrow (\bar{q}_L \phi^G) u_R - f_\downarrow (\bar{q}_L \phi) d'_R + \text{h.c.} , \quad (16.37)$$



where  $q_L$  is given by (16.25) and  $\varphi^G$  is the  $G$ -conjugation of  $\varphi$  given by

$$\varphi^G = \begin{pmatrix} \varphi_2^* \\ -\varphi_1^* \end{pmatrix}. \quad (16.38)$$

So, if we make the replacement (16.6) for  $\varphi$ , then

$$\mathcal{L}_{\text{f}\varphi} \rightarrow -f\eta\bar{e}e - f_{\uparrow}\eta\bar{u}u - f_{\downarrow}\eta\bar{d}'d' + \dots. \quad (16.39)$$

From this, we obtain the following formulas for the masses:

$$m_e = f\eta, \quad m_u = f_{\uparrow}\eta, \quad m_{d'} = f_{\downarrow}\eta. \quad (16.40)$$

In this approximation,  $d'$  is still a mass eigenstate. A weak point of this method is that, for each fermion, we need to introduce another parameter  $f$ . Thus, this part is phenomenological, apart from the gauge principle.

### 16.3 GIM Mechanism

In the quark model in which only  $u$ ,  $d$ , and  $s$  are taken into account, the hadron contribution to  $j_{\mu}^3$  is given by

$$\frac{1}{2} \left( i\bar{u}\gamma_{\mu} \frac{1+\gamma_5}{2} u - \bar{d}'\gamma_{\mu} \frac{1+\gamma_5}{2} d' \right). \quad (16.41)$$

This contains a strangeness-changing part, i.e., in the second term there appears

$$-\frac{1}{2} \sin\theta_C \left( i\bar{d}\gamma_{\mu} \frac{1+\gamma_5}{2} s + i\bar{s}\gamma_{\mu} \frac{1+\gamma_5}{2} d \right). \quad (16.42)$$

Such terms can lead to the following decay pattern, where a neutral current changes the strangeness:

$$K_L^0 \rightarrow \mu^+ + \mu^-. \quad (16.43)$$

Experimentally, the branching ratio is very small, viz.,  $(9.1 \pm 1.9) \times 10^{-9}$ . That means that there must be some mechanism that suppresses the process. The GIM mechanism is based on the assumption that there is a fourth quark  $c$ , and each item in (16.26) is an  $SU(2)_L$  doublet. The hadron parts of the charged current are then

given by

$$j_\mu = i\bar{d}'\gamma_\mu \frac{1+\gamma_5}{2}u + i\bar{s}'\gamma_\mu \frac{1+\gamma_5}{2}c + \dots, \quad (16.44)$$

$$j_\mu^\dagger = i\bar{u}\gamma_\mu \frac{1+\gamma_5}{2}d' + i\bar{c}\gamma_\mu \frac{1+\gamma_5}{2}s' + \dots. \quad (16.45)$$

The hadron part of the neutral current then becomes

$$\begin{aligned} j_\mu^3 &= \frac{1}{2} \left( i\bar{u}\gamma_\mu \frac{1+\gamma_5}{2}u + i\bar{c}\gamma_\mu \frac{1+\gamma_5}{2}c - i\bar{d}'\gamma_\mu \frac{1+\gamma_5}{2}d' - i\bar{s}'\gamma_\mu \frac{1+\gamma_5}{2}s' \right) \\ &= \frac{1}{2} \left( i\bar{u}\gamma_\mu \frac{1+\gamma_5}{2}u + i\bar{c}\gamma_\mu \frac{1+\gamma_5}{2}c - i\bar{d}\gamma_\mu \frac{1+\gamma_5}{2}d' - i\bar{s}\gamma_\mu \frac{1+\gamma_5}{2}s \right), \end{aligned} \quad (16.46)$$

so the flavour is conserved. In the neutral current, strangeness is conserved. In fact, the reason why (16.43) still occurs is that the flavour is not strictly conserved, since the flavour symmetry is broken. The above mechanism is called the *GIM mechanism* [130]. As already mentioned in Sect. 13.7, the fourth quark  $c$  was subsequently discovered by Ting and Richter et al. in [132, 133].

## 16.4 Anomalous Terms and Generation of Fermions

So far we have investigated several kinds of symmetries in classical theories or at the lowest order in perturbation. At first, such symmetries were assumed to hold true in general if higher order corrections were included. However, during the 1960s, through research by Adler et al. [159, 160], it gradually became clearer that a certain kind of symmetry is broken by higher order corrections or quantization of a system. In other words, a certain kind of symmetry is not consistent with quantization. Investigating field equations or Green's functions, it sometimes happens that terms are produced which break originally (classically) existing symmetries. These are called *anomalous terms*. Since the details will be discussed in the last chapter, we shall only outline the issue very briefly here. If a certain kind of symmetry is lost by the emergence of anomalous terms, it may happen that renormalizability is also lost. Hence, the requirement that anomalous terms should not appear is sometimes used as a restriction on a model. We shall give an example in this section.

We begin with Adler's anomalous term. In QED, we define the axial vector  $A_\mu$ , the pseudoscalar current  $P$ , and a pseudoscalar term  $C$  by

$$A_\mu = i\bar{\psi}\gamma_\mu\gamma_5\psi, \quad P = i\bar{\psi}\gamma_5\psi, \quad C = \mathbf{E} \cdot \mathbf{H} = \frac{i}{8}\epsilon_{\alpha\beta\gamma\delta}F_{\alpha\beta}F_{\gamma\delta}, \quad (16.47)$$

and we assume that they have been renormalized multiplicatively. In the Heisenberg representation, we have

$$\partial_\lambda A_\lambda = 2m\mathbf{P} - \frac{e^2}{2\pi^2}\mathbf{C} . \quad (16.48)$$

Since in the classical theory, or at the lowest order in perturbation theory, the second term on the right-hand side does not exist, this term is the anomaly. In QED, the axial vector current does not appear in the interaction, so this anomalous term is harmless, but in theories where the axial vector current appears, as in the Weinberg–Salam model, the emergence of such an anomalous term is problematic. It breaks important symmetries, and blocks renormalizability as well. We must therefore choose models in which the anomalous term does not appear. As will be discussed later, the anomalous term appearing here is produced by the contribution from fermion loops, so let us begin by writing down the interaction terms of fermions and gauge fields.

In the Weinberg–Salam model, the representative terms in the Lagrangian density can be written as

$$i\bar{L}\gamma_\mu \left( \frac{g}{2}\hat{C}\hat{W}_\mu + \frac{g'}{2}C_0B_\mu \right) L + i\bar{R}\gamma_\mu g' DB_\mu R . \quad (16.49)$$

For the lepton doublet,  $(\nu_e, e)$ , and the lepton singlet  $e$ , we have

$$\hat{C} = \hat{\tau} , \quad C_0 = 1 , \quad D = \frac{1 - \tau_3}{2} . \quad (16.50)$$

For the quark doublet and right-handed singlet, we have

$$\hat{C} = \hat{\tau} , \quad C_0 = -\frac{1}{3} , \quad D = -\frac{1}{3} + \frac{1 - \tau_3}{2} . \quad (16.51)$$

The relations with the electric charge  $Q$ , in units of  $e$ , are

$$Q = -D = \frac{1}{2}(C_3 - C_0) . \quad (16.52)$$

If we write the field operator of a quark or a lepton as  $\psi$  and we introduce the notation

$$L = \frac{1}{2}(1 + \gamma_5)\psi , \quad R = \frac{1}{2}(1 - \gamma_5)\psi , \quad (16.53)$$

then (16.49) becomes

$$i\bar{\psi}\gamma_\mu \mathcal{Y}_\mu \psi + i\bar{\psi}\gamma_\mu \gamma_5 \mathcal{A}_\mu \psi . \quad (16.54)$$

However, corresponding to (16.49),

$$\begin{aligned}\mathcal{V}_\mu &= \frac{g}{4}\hat{C}\hat{W}_\mu - \frac{g'}{4}C_3B_\mu + \frac{g'}{2}C_0B_\mu, \\ \mathcal{A}_\mu &= \frac{g}{4}\hat{C}\hat{W}_\mu + \frac{g'}{4}C_3B_\mu.\end{aligned}\quad (16.55)$$

According to Bardeen [161], the anomalous terms from the fermion loops are

$$\partial_\lambda(\bar{\psi}\gamma_\lambda\gamma_5C_j\psi) = \text{normal term} + \frac{1}{4\pi^2}\epsilon_{\mu\nu\rho\sigma}\text{Tr}[C_j\{\dots\}], \quad (16.56)$$

where

$$\begin{aligned}\{\dots\} &= \frac{1}{4}\mathcal{V}_{\mu\nu}\mathcal{V}_{\rho\sigma} + \frac{1}{12}\mathcal{A}_{\mu\nu}\mathcal{A}_{\rho\sigma} \\ &\quad + \frac{2}{3}i(\mathcal{A}_\mu\mathcal{A}_\nu\mathcal{V}_{\rho\sigma} + \mathcal{A}_\mu\mathcal{V}_{\nu\rho}\mathcal{A}_\sigma + \mathcal{V}_{\mu\nu}\mathcal{A}_\rho\mathcal{A}_\sigma) \\ &\quad - \frac{8}{3}\mathcal{A}_\mu\mathcal{A}_\nu\mathcal{A}_\rho\mathcal{A}_\sigma,\end{aligned}\quad (16.57)$$

$$\mathcal{V}_{\mu\nu} = \partial_\mu\mathcal{V}_\nu - \partial_\nu\mathcal{V}_\mu - i[\mathcal{V}_\mu, \mathcal{V}_\nu] - i[\mathcal{A}_\mu, \mathcal{A}_\nu], \quad (16.58)$$

$$\mathcal{A}_{\mu\nu} = \partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu - i[\mathcal{V}_\mu, \mathcal{A}_\nu] - i[\mathcal{A}_\mu, \mathcal{V}_\nu].$$

We now consider the condition for such anomalous terms to disappear. This means that the trace should vanish. Expressing the products of the matrices by separating them into commutators and anti-commutators, the trace is eventually described by a coefficient corresponding to  $f_{abc}$  or  $d_{abc}$  in the case of  $SU(3)$ . As discussed by Georgi and Glashow in [162], the condition for the anomalous terms to vanish is

$$d_{abc} = 0. \quad (16.59)$$

This is because at least one factor in any trace becomes  $d$ . So it is required that, for  $C_i$ ,  $i = 1, 2, 3$ , and  $C_0$ , the above equation holds true. Firstly, putting  $C_0$  aside, for  $SU(2)$ ,

$$\text{Tr}\{C_i, C_j\}C_k = 0. \quad (16.60)$$

Note that, since  $\mathcal{A}_\mu$  is not included in  $C_0$ , what matters is

$$\text{Tr}\{C_i, C_j\}C_0 = \text{Tr}\{C_i, C_j\}(C_3 - 2Q). \quad (16.61)$$

Since this becomes non-zero only for  $i = j$ , and  $C_i^2 = 1$ , the condition for the trace to disappear turns out to be

$$\text{Tr } Q = 0 . \quad (16.62)$$

Hence, the sum of the electric charges of all elementary particles with spin 1/2 should vanish. The existence of the  $c$  quark provides a way to satisfy this condition, because the sum of the charges of all the leptons and quarks in each generation is zero. The different generations are:

$$\begin{aligned} \text{leptons} & \quad \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \dots, \\ \text{quarks} & \quad \begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \dots \end{aligned}$$

The sum of the electric charges of the leptons in each generation is  $0 - 1 = -1$ , while the sum of the electric charges of the quarks in each generation is

$$3 \times \left( \frac{2}{3} - \frac{1}{3} \right) = 1 ,$$

assuming that each quark has three possible colours. This would not work without the  $c$  quark. This is why the GIM mechanism is useful for eliminating anomalies.

From the standpoint of cancelling anomalies, it is plausible to consider each generation to consist of a neutral neutrino, a lepton with a negative electric charge, three-coloured quarks with an electric charge  $2e/3$ , and three-coloured quarks with an electric charge  $-e/3$ . Moreover, since there is a third generation of leptons, viz.,  $(\nu_\tau, \tau)$ , we need to consider a third generation of quarks as well. Hence, the quarks  $t$  and  $b$  were postulated. For the third generation, the reader is referred to the discussion in Sect. 13.6.

In addition, in the discussion above we see that the colour degrees of freedom play an important role. There have to be three colour states for each quark for this to work. In the following, we will give another argument which supports this conclusion.

As already discussed in Sect. 8.5, the amplitude for the decay process  $\pi^0 \rightarrow 2\gamma$  is given by (8.103). The effective Hamiltonian density corresponding to this is

$$\mathcal{H}_{\text{eff}} = \frac{\alpha}{\pi} \frac{G}{m} \varphi \mathbf{E} \cdot \mathbf{H} . \quad (16.63)$$

We now derive the effective Hamiltonian density corresponding to (16.63) by using the quark model, and analyze hints of the presence of the colour.

We introduce the electromagnetic interaction to the quark model as a perturbation and check its effects. To do so, we consider the axial vector current

$$A_\lambda^a = \sum_q i\bar{q}\gamma_\lambda\gamma_5\frac{\tau_a}{2}q \quad (q \text{ is only for the first generation}) . \quad (16.64)$$

In QCD, referring to (15.44), the following PCAC relation [129] holds true:

$$\partial_\lambda A_\lambda^a = -f_\pi\mu^2\varphi_a . \quad (16.65)$$

Note that, introducing the electromagnetic interaction as a perturbation, because of Adler's anomalous terms, the equation above transforms to

$$\partial_\lambda A_\lambda^a = -f_\pi\mu^2\varphi_a + b\delta_{a3}\mathbf{E} \cdot \mathbf{H} . \quad (16.66)$$

We will show how to compute the coefficient  $b$  in the last chapter. It turns out to be the same as that of  $\pi^0 \rightarrow 2\gamma$ , viz.,

$$b = -\frac{2\alpha}{\pi}N_C S , \quad (16.67)$$

where  $N_C$  is the number of quark colours and  $S$  is given by

$$S = \sum_{u,d} \left(\frac{\tau_3}{2}\right)_a Q_a^2 = \frac{1}{2} \left(\frac{2}{3}\right)^2 - \frac{1}{2} \left(-\frac{1}{3}\right)^2 = \frac{1}{6} . \quad (16.68)$$

In the sigma model introduced in Sect. 15.2, the term we need to add to the Lagrangian density to change the PCAC relation from (16.65) to (16.66) is

$$\Delta\mathcal{L} = \frac{b}{f_\pi}\varphi_3\mathbf{E} \cdot \mathbf{H} . \quad (16.69)$$

In fact, in this case, the PCAC relation assumes the form

$$\partial_\lambda A_\lambda^a = -f_\pi\mu^2\varphi_a + b\delta_{a3}\frac{\sigma}{f_\pi}\mathbf{E} \cdot \mathbf{H} . \quad (16.70)$$

Here, replacing  $\sigma \rightarrow \eta$  and  $f_\pi \rightarrow \eta$ , we note that (16.70) coincides with (16.66). Then, using the effective Lagrangian density of  $\pi^0 \rightarrow 2\gamma$  in (16.69), based on the quark model, we can compute this decay width and compare it with the experimental results:

$$\Gamma(\pi^0 \rightarrow 2\gamma) = \frac{\mu^3}{64\pi^3} \left(\frac{2\alpha}{f_\pi}N_C S\right)^2 = (7.95 \pm 0.55) \text{ eV} . \quad (16.71)$$

This implies that

$$N_C = 3.06 \pm 0.10 . \quad (16.72)$$

This provides experimental support for the idea that there are three colour states. Moreover, since  $\pi^0$  is a bound state of the form  $q\bar{q}$ , the computation is difficult. To avoid this, we used the sigma model in the computation above, whence the computation itself is phenomenological.

## 16.5 Grand Unified Theory

The Weinberg–Salam theory succeeded in unifying the electromagnetic and weak interactions by applying the gauge principle, thereby producing the *electroweak interaction*. On the other hand, the strong interaction is obtained by applying the gauge principle to the colour group  $SU(3)$ . The dynamics of the strong interaction is called *quantum chromodynamics*, abbreviated as QCD. Since the gauge particles are bosons which form hadrons by gluing quarks together, they are called *gluons*. This symmetry is not broken, so the gluons are assumed to be massless and the strong force should be long range.

Therefore, the gauge group governing the strong interaction and the electroweak interaction is

$$SU(3) \times SU(2) \times U(1) . \quad (16.73)$$

Here we come up with the idea that, if we start with a larger gauge group  $G$ , and consider  $G$  to decompose by spontaneous symmetry breaking into (16.73), then we may obtain a unified description of the strong, weak, and electromagnetic interactions. Such a theory is called a *grand unified theory*. If this gauge theory is to be specified by just one coupling constant, a simple group  $G$  is preferred, i.e., we assume that  $G$  does not have any invariant subgroup except for  $G$  itself and the identity. We thus look for a simple group  $G$  satisfying the condition

$$G \supset SU(3) \times SU(2) \times U(1) \supset SU(3)_{\text{colour}} \times U(1)_{\text{em}} . \quad (16.74)$$

Regarding the group (16.73), due to spontaneous symmetry breaking, the strict symmetries surviving to the end are only the colour  $SU(3)$  and the electromagnetic  $U(1)$ . Since the ranks of  $SU(3)$ ,  $SU(2)$ , and  $U(1)$  are 2, 1, and 1, respectively, the rank of  $G$  must be at least 4. Although there have been many proposals for such a group, we shall consider as an example the group  $SU(5)$ , suggested by Georgi and Glashow in [163].

We note that gravity is not included in this grand unified theory. The energies at which gravity becomes important are

$$E \gtrsim \left( \frac{\hbar c}{G} \right)^{1/2} c^2 \sim 10^{19} \text{ GeV}, \quad (16.75)$$

where  $G$  is the gravitational constant and  $(\hbar c/G)^{1/2}$  is the *Planck mass*. Thus, the grand unified theory without gravity, known also as GUT, should be formulated at energies less than  $10^{19}$  GeV.

In general, the conditions that the group  $G$  must satisfy are:

1.  $G$  is simple and only one coupling constant is introduced.
2. Electric charges of quarks and leptons should satisfy the following inequality, which is established by observation:

$$\left| \frac{Q_e + Q_p}{Q_p} \right| \lesssim \mathcal{O}(10^{-21}). \quad (16.76)$$

3. The rank of  $G$  is greater than or equal to 4:

$$G \supset SU(3) \times SU(2) \times U(1). \quad (16.77)$$

4. To accommodate the Dirac particle, a complex representation should be included.

In the following, we consider  $G = SU(5)$ . First, we consider three generations of quarks and leptons:

$$\begin{aligned} \text{I.} & \quad \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} u \\ d \end{pmatrix}_L, \bar{e}_R, u_R, d_R, \\ \text{II.} & \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \bar{\mu}_R, c_R, s_R, \\ \text{III.} & \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L, \bar{\tau}_R, t_R, b_R. \end{aligned}$$

All quarks have three possible colours. The first generation is thus

$$(\nu_e, e, \bar{e}, u_i, d_i, \bar{u}_i, \bar{d}_i)_L, \quad i = 1, 2, 3, \quad (16.78)$$



where  $\bar{u}_L = \overline{(u_R)}$ . There are 15 components altogether. We use them to express the fundamental representation  $\bar{\mathbf{5}}$  and the anti-symmetric part  $\mathbf{10}$  of  $\mathbf{5} \times \mathbf{5}$ :

$$\bar{\mathbf{5}} \oplus \mathbf{10}, \quad (16.79)$$

$$\bar{\mathbf{5}} = \begin{pmatrix} \bar{d}_R \\ \bar{d}_Y \\ \bar{d}_B \\ e \\ \nu \end{pmatrix}_L, \quad \mathbf{10} = \begin{pmatrix} 0 & \bar{u}_B & -\bar{u}_Y & -u_R & -d_R \\ -\bar{u}_B & 0 & \bar{u}_R & -u_Y & -d_Y \\ \bar{u}_Y & -\bar{u}_R & 0 & -u_B & -d_B \\ u_R & u_Y & u_B & 0 & -\bar{e} \\ d_R & d_Y & d_B & \bar{e} & 0 \end{pmatrix}_L, \quad (16.80)$$

where R, Y, and B indicate the three colours, corresponding to red, yellow, and blue. Expressing them in terms of the transformation properties under  $SU(3)$  and  $SU(2)$ , for instance,

$$\bar{\mathbf{5}} = (\bar{\mathbf{3}}, \mathbf{1}) + (\mathbf{1}, \mathbf{2}) = \bar{d} + \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L,$$

$$(\mathbf{3}, \mathbf{2}) = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad (\mathbf{1}, \mathbf{1}) = \bar{e}_L = \overline{(e_R)}.$$

In this theory, since the charge operator  $Q$  is one of the generators of  $SU(5)$ , the trace vanishes. Hence, the sum of the charges of the members of  $\bar{\mathbf{5}}$  also vanishes:

$$-3Q_d + Q_e = 0, \quad \text{i.e., } Q_d = \frac{1}{3}Q_e = -\frac{1}{3}. \quad (16.81)$$

Since the trace of  $Q$  is zero, there is no anomalous term.

In addition, there are 24 gauge particles in  $SU(5)$ . Half of them are accounted for by 8 gluons and 4 electroweak gauge particles. The rest play the role of connecting leptons with quarks. Considering this kind of gauge particle as a mediator, a nucleon can decay into leptons. We call this gauge particle X, and the width of the nucleon decay becomes the same as the width of the  $\mu$  decay:

$$\Gamma(N \rightarrow \text{lepton}) \propto a_G^2 \frac{m_p^5}{M_X^4}. \quad (16.82)$$

The appearance of the fifth power of the mass of the decaying particle is analogous to what we see in the  $\mu$  decay. The fourth power of the mass of the intermediate gauge particle comes from the squared propagator of this boson. Using the coupling constant  $g$  in the grand unified theory,  $\alpha_G$  is given by

$$\alpha_G = \frac{g^2}{4\pi}. \quad (16.83)$$

Then the lower bound for the proton lifetime is

$$\tau_p > 10^{32} \text{ yr.} \quad (16.84)$$

From this, a lower bound for  $M_X$  is obtained:

$$M_X \gtrsim 10^{15} \text{ GeV} \ll \text{Planck mass.} \quad (16.85)$$

To evaluate  $M_X$ , we use the renormalization group. Even if the symmetry is broken at low energies, we assume that the breaking becomes gradually less important at high energies, until in the end, at an energy  $M_X$ , the symmetry is recovered. This is analogous to the fact that, even if the rotational symmetry is broken in a ferromagnet at low energies, a phase transition occurs at high energies and the rotational symmetry is then recovered. In this case, what corresponds to the temperature is a value of the four-momentum transfer. Writing this as  $Q^2$ , the effective coupling constant becomes a function of  $Q^2$ . This way of thinking is used in the renormalization group method, discussed in Chap. 20. Thus, writing the effective coupling constants at low energies corresponding to the groups  $SU(3)$ ,  $SU(2)$ , and  $U(1)$  as  $\alpha_3(Q^2)$ ,  $\alpha_2(Q^2)$ , and  $\alpha_1(Q^2)$ , respectively, if the symmetry is indeed recovered at high energies, we have

$$\alpha_1(M_X^2) = \alpha_2(M_X^2) = \alpha_3(M_X^2) \equiv \alpha_G(M_X^2). \quad (16.86)$$

However, at low energies, they differ from each other. With increasing  $Q^2$ ,  $\alpha_2(Q^2)$  and  $\alpha_3(Q^2)$  decrease, while  $\alpha_1(Q^2)$  increases. Using the renormalization group, their  $Q^2$ -dependence is

$$SU(3) \quad \frac{1}{\alpha_3(Q^2)} - \frac{1}{\alpha_3(M_X^2)} = -\frac{1}{4\pi} \left( 11 - \frac{2}{3}n_f \right) \ln \frac{M_X^2}{Q^2}, \quad (16.87)$$

$$SU(2) \quad \frac{1}{\alpha_2(Q^2)} - \frac{1}{\alpha_2(M_X^2)} = -\frac{1}{4\pi} \left( \frac{22}{3} - \frac{2}{3}n_f \right) \ln \frac{M_X^2}{Q^2}, \quad (16.88)$$

$$U(1) \quad \frac{1}{\alpha_1(Q^2)} - \frac{1}{\alpha_1(M_X^2)} = \frac{1}{4\pi} \frac{2}{3}n_f \ln \frac{M_X^2}{Q^2}, \quad (16.89)$$

where  $n_f$  is the number of flavours. It turns out that, if  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  are determined experimentally at low energies, then we can find the value of  $M_X$ . Let us therefore determine  $M_X$  from the Weinberg angle.

Using (16.2) in the Weinberg–Salam theory, measuring  $Q$  in units of  $e$ , and writing  $I_i$  as  $T_i$ , we have

$$Q = T_3 + \frac{1}{2}Y. \quad (16.90)$$

We understand that the coupling constants multiplying  $\hat{T}$  and  $Y/2$  are  $g$  and  $g'$ , respectively. Here,  $\hat{T}$  and  $Y$  are connected to  $\hat{W}_\mu$  and  $B_\mu$ , respectively, but in  $SU(5)$ , normalizing  $Y$  again,

$$\text{Tr}(T_i T_j) = \frac{1}{2} \delta_{ij}, \quad i, j = 1, 2, \dots, 24. \quad (16.91)$$

Therefore, considering the representation  $\bar{\mathbf{5}}$  and taking

$$Y = c T_Y, \quad (16.92)$$

we can determine  $c$ . We thus consider the square of the following equation:

$$Q = T_3 + \frac{c}{2} T_Y. \quad (16.93)$$

Therefore, since there exists no quark but  $(\overline{d_R})$ , this does not contribute to  $T_3$ , and from (16.91),

$$\sum_{\bar{\mathbf{5}}} Q^2 = 3 \left( \frac{1}{3} \right)^2 + 1 = \frac{4}{3}, \quad \sum_{\bar{\mathbf{5}}} T_3^2 = \sum_{\bar{\mathbf{5}}} T_Y^2 = \frac{1}{2}, \quad \sum_{\bar{\mathbf{5}}} T_3 T_Y = 0, \quad (16.94)$$

$$\sum_{\bar{\mathbf{5}}} Q^2 = \sum_{\bar{\mathbf{5}}} \left( T_3 + \frac{c}{2} T_Y \right)^2 = \frac{1}{2} + \frac{c^2}{8} = \frac{4}{3}. \quad (16.95)$$

Thus, the value of  $c$  is determined. We obtain

$$Y = 2\sqrt{\frac{5}{3}} T_Y. \quad (16.96)$$

The coefficient of  $T_3$  is  $g$  and that of  $Y$  is  $g'/2$ . Therefore, setting

$$g T_i W_{\mu i} = g_2 T_i W_{\mu i}, \quad \frac{1}{2} g' Y B_\mu = g_1 T_Y B_\mu, \quad (16.97)$$

the coupling constants of  $SU(2)$  and  $U(1)$  are  $g_2$  and  $g_1$ , respectively. In the limit where  $SU(5)$  is exact, we should have  $g_1 = g_2$ . From the equation above, we obtain

$$g' = \sqrt{\frac{3}{5}} g_1. \quad (16.98)$$

We assume that these coupling constants have been defined at low energies. Hence, the Weinberg angle is given by

$$\sin^2 \theta_W = \frac{g'^2}{g^2 + g'^2} = \frac{\frac{3}{5}g_1^2}{g_2^2 + \frac{3}{5}g_1^2}. \quad (16.99)$$

Moreover,  $e^2$  is given by

$$e^2 = \frac{g^2 g'^2}{g^2 + g'^2} = \frac{\frac{3}{5}g_1^2 g_2^2}{g_2^2 + \frac{3}{5}g_1^2}. \quad (16.100)$$

This implies

$$\alpha = \frac{3}{5} \frac{\alpha_1 \alpha_2}{\alpha_2 + \frac{3}{5}\alpha_1} = \alpha_2 \sin^2 \theta_W. \quad (16.101)$$

If we take the difference between (16.88) and (16.89), we obtain

$$\frac{1}{\alpha_2} - \frac{1}{\alpha_1} = -\frac{1}{4\pi} \frac{22}{3} \ln \frac{M_X^2}{Q^2}. \quad (16.102)$$

Additionally, expressing (16.99) in terms of  $\alpha_1$  and  $\alpha_2$ ,

$$\sin^2 \theta_W = \frac{\frac{3}{5}\alpha_1}{\alpha_2 + \frac{3}{5}\alpha_1}. \quad (16.103)$$

Now, eliminating  $\alpha_1$  and  $\alpha_2$  using (16.101), (16.102), and (16.103), we obtain an equation in which all the constants are known:

$$\sin^2 \theta_W = \frac{3}{8} \left( 1 - \frac{\alpha}{4\pi} \frac{110}{9} \ln \frac{M_X^2}{Q^2} \right), \quad (16.104)$$

where  $Q^2$  is of low energy and is assumed to be of order  $(\text{GeV})^2$ . Therefore, if we choose

$$M_X \sim 10^{16} \text{ GeV}, \quad (16.105)$$

the experimental value of  $\theta_W$  is reproduced:

$$\sin^2 \theta_W = 0.21 . \quad (16.106)$$

However, since the lifetime of the proton is about  $10^{31}$  years, it is hard to reach the experimental lower bound of  $10^{32}$  years. Although this model is simple, it does not necessarily agree with experiment. However, it is not until we arrive at the grand unified theory that we recognize that quarks and leptons are originally of the same species. On this point, it seems that a clear understanding has been obtained of the problem of the universality of quarks and leptons discussed in Sect. 13.6.

# Chapter 17

## Path-Integral Quantization Method



Apart from the canonical quantization method, another way to quantize fields is the path-integral method. In many problems, the two methods give the same results, but for the quantization of gauge fields the latter is more effective. It was after the path-integral method had been well understood that the canonical quantization method came to be used. The path-integral method was invented by Feynman [96], and applied to the gauge theory by Faddeev and Popov in [164].

### 17.1 Quantization of a Point-Particle System

We begin with the easiest quantization of a point-particle system. In the Heisenberg picture, the operators depend on the time variable, and an operator  $Q$  is expressed as

$$Q_H(t) . \tag{17.1}$$

Its eigenstate is expressed as  $|q, t\rangle_H$ , whence

$$Q_H(t)|q, t\rangle_H = q|q, t\rangle_H , \tag{17.2}$$

where  $q$  is one of the eigenvalues of  $Q$ . The relationship with the Schrödinger picture is

$$Q_H(t) = e^{iHt} Q_S e^{-iHt} , \tag{17.3}$$

$$|q\rangle = e^{-iHt} |q, t\rangle_H , \tag{17.4}$$

$$Q_S |q\rangle = q |q\rangle . \tag{17.5}$$

We introduce a transformation function:

$$F(q', t'; q, t) = {}_H\langle q', t' | q, t \rangle_H = \langle q' | \exp[-iH(t' - t)] | q \rangle. \quad (17.6)$$

So far the discussion has been based on the canonical formalism. With this as the starting point, we shall now introduce the path integral.

We begin by dividing  $(t - t')$  into  $(n + 1)$  equal intervals:

$$t' = t + (n + 1)\epsilon, \quad t_l = t + l\epsilon. \quad (17.7)$$

Therefore, from the complete set of  $|q_l, t_l\rangle$ , we obtain

$$F(q', t'; q, t) = \int dq_1(t_1) \cdots \int dq_n(t_n) \langle q', t' | q_n, t_n \rangle \langle q_n, t_n | q_{n-1}, t_{n-1} \rangle \cdots \langle q_1, t_1 | q, t \rangle, \quad (17.8)$$

where we have omitted the indices. If we choose  $\epsilon$  small enough, then

$$\langle q', \epsilon | q, 0 \rangle = \langle q' | e^{-i\epsilon H} | q \rangle = \delta(q' - q) - i\epsilon \langle q' | H | q \rangle + \mathcal{O}(\epsilon^2). \quad (17.9)$$

The first term expresses the orthonormality of the eigenstates. We choose the following form for  $H$ :

$$H = \frac{1}{2}P^2 + V(Q). \quad (17.10)$$

Therefore, the second term on the right-hand side of (17.9) becomes

$$\begin{aligned} \langle q' | H(P, Q) | q \rangle &= \int \frac{dp}{2\pi} \exp[ip(q' - q)] \left[ \frac{1}{2}p^2 + V(q) \right] \\ &= \int \frac{dp}{2\pi} \exp[ip(q' - q)] H\left(p, \frac{q + q'}{2}\right). \end{aligned} \quad (17.11)$$

Thus, to first order in  $\epsilon$ ,

$$\langle q_l, t_l | q_{l-1}, t_{l-1} \rangle = \int \frac{dp}{2\pi} \exp \left[ ip(q_l - q_{l-1}) - i\epsilon H \left( p, \frac{1}{2}(q_l + q_{l-1}) \right) \right]. \quad (17.12)$$

Setting  $q_0 = q$  and  $q_{n+1} = q'$  and inserting this into (17.8), we obtain

$$\begin{aligned} F(q', t'; q, t) &= \lim_{n \rightarrow \infty} \int \prod_{j=1}^n dq_j \int \prod_{j=1}^{n+1} \frac{dp_j}{2\pi} \exp \left\{ i \sum_{l=1}^{n+1} \left[ p_l(q_l + q_{l-1}) \right. \right. \\ &\quad \left. \left. - H \left( p_l, \frac{1}{2}(q_l + q_{l-1}) \right) (t_l - t_{l-1}) \right] \right\}. \end{aligned} \quad (17.13)$$

We see from this that the number of integration variables  $p_j$  is one greater than the number of  $q_j$ . Moreover, the normalization condition is determined by unitarity:

$$\int dq' F(q', t'; q_1, t) F^*(q', t'; q_2, t) = \int dq' \langle q', t' | q_1, t \rangle \langle q', t' | q_2, t \rangle^* = \delta(q_1 - q_2). \quad (17.14)$$

Bearing this in mind, without paying too much attention to the normalization, we write the limit (17.13) as

$$F(q', t'; q, t) = \int \mathcal{D}p \mathcal{D}q \exp \left\{ i \int_t^{t'} dt [p\dot{q} - H(p, q)] \right\}. \quad (17.15)$$

Applying this formula to  $H$  as given above, the  $p$ -integral can be carried out:

$$\int_{-\infty}^{\infty} \frac{dp}{2\pi} \exp \left[ i \left( p\dot{q} - \frac{1}{2}p^2 \right) \epsilon \right] = \frac{1}{\sqrt{2\pi i\epsilon}} \exp \left( \frac{1}{2}i\dot{q}^2 \epsilon \right). \quad (17.16)$$

Therefore,

$$\begin{aligned} F(q', t'; q, t) &= \lim_{n \rightarrow \infty} \left( \frac{1}{\sqrt{2\pi i\epsilon}} \right)^{n+1} \int \prod_{j=1}^n dq_j \\ &\quad \times \exp \left\{ i \sum_{l=1}^{n+1} \epsilon \left[ \frac{1}{2} \left( \frac{q_l - q_{l-1}}{\epsilon} \right)^2 - v \left( \frac{q_l + q_{l-1}}{2} \right) \right] \right\} \\ &= \int \mathcal{D}q \exp \left[ i \int_t^{t'} dt L(q, \dot{q}) \right], \end{aligned} \quad (17.17)$$

where

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^2 - V(q). \quad (17.18)$$

Thus, the argument of the exponential is just the action integral:

$$S = \int_t^{t'} dt L(q, \dot{q}). \quad (17.19)$$

As shown above, it turns out that the path integral can be expressed by both the Hamiltonian formalism and the Lagrangian formalism. However, as is clear from the derivation, the Hamiltonian formalism is more fundamental.



In the case of the Lagrangian formalism, the formula above is sometimes deformed to some extent. For example, consider the Lagrangian

$$L = \frac{1}{2} \dot{q}^2 f(q), \quad (17.20)$$

$$p = \frac{\partial L}{\partial \dot{q}} = \dot{q} f(q), \quad H(p, q) = p\dot{q} - L = \frac{1}{2} p^2 [f(q)]^{-1}. \quad (17.21)$$

In this case, the transformation function is

$$\begin{aligned} F(q', t'; q, t) &= \lim_{n \rightarrow \infty} \int \prod_{j=1}^n dq_j \int \prod_{j=1}^{n+1} dp_j \\ &\quad \times \exp \left[ i \left\{ \sum_{l=1}^{n+1} p_l (q_l - q_{l-1}) - \frac{1}{2} \epsilon p_l^2 \left[ f \left( \frac{q_l + q_{l-1}}{2} \right) \right]^{-1} \right\} \right] \\ &= \lim_{n \rightarrow \infty} \int \prod_{j=1}^n dq_j \sqrt{\frac{1}{2\pi i \epsilon}} f \left( \frac{q_j + q_{j-1}}{2} \right) \\ &\quad \times \exp \left\{ i \sum_{l=1}^{n+1} \epsilon \left[ \frac{1}{2} \left( \frac{q_l - q_{l-1}}{\epsilon} \right)^2 f \left( \frac{q_l + q_{l-1}}{2} \right) \right] \right\} \\ &= \lim_{n \rightarrow \infty} \left( \frac{1}{\sqrt{2\pi i \epsilon}} \right)^{n+1} \int \prod_{j=1}^n dq_j \\ &\quad \times \exp \left\{ i \sum_{l=1}^{n+1} \epsilon \left[ \frac{1}{2} \left( \frac{q_l - q_{l-1}}{\epsilon} \right)^2 f \left( \frac{q_l + q_{l-1}}{2} \right) - \frac{i}{2\epsilon} \ln f \left( \frac{q_l + q_{l-1}}{2} \right) \right] \right\} \\ &= \int \mathcal{D}q \exp(iS_{\text{eff}}), \quad (17.22) \end{aligned}$$

where

$$S_{\text{eff}} = \int_t^{t'} dt \left[ L(q, \dot{q}) - \frac{i}{2} \delta(0) \ln f(q) \right]. \quad (17.23)$$

To obtain this result, we made the substitution

$$\sum \epsilon \rightarrow \int dt, \quad \frac{1}{\epsilon} \rightarrow dt^{-1} \rightarrow \delta(0). \quad (17.24)$$

As we have seen above, this method can be used to obtain the effective Lagrangian in the computation of Feynman's equations. What is important here is that, when

we go from the Hamiltonian formalism to the Lagrangian formalism,  $dq$  is replaced by

$$dq \rightarrow dq\sqrt{f(q)}. \quad (17.25)$$

This form is invariant under the general coordinate transformation. In fact, in Riemannian geometry, when the line element

$$ds^2 = g_{ij}dx^i dx^j \quad (17.26)$$

is given, the volume element which is invariant under the general coordinate transformation is

$$dV = \sqrt{|g|}dx^1 \dots dx^n. \quad (17.27)$$

The substitution (17.25) is a special case of this. Here  $g$  is the determinant defined by

$$g = \det(g_{ij}). \quad (17.28)$$

The substitution (17.25) corresponds to (17.27) with  $n = 1$ . This result will be extended to the field theory case later and applied to reproduce the result in Sect. 11.5. Moreover, it is easy to extend the previous result to the system with  $N$  degrees of freedom.

Next, we consider the matrix elements of the operators. We impose the condition

$$t' > t_0 > t. \quad (17.29)$$

In this case, the matrix element of  $Q(t_0)$  is

$$\begin{aligned} \langle q', t' | Q(t_0) | q, t \rangle &= \int dq_1(t_1) \dots \int dq_n(t_n) \langle q', t' | q_n, t_n \rangle \langle q_n, t_n | q_{n-1}, t_{n-1} \rangle \\ &\dots \langle q_{i+1}, t_{i+1} | q_i, t_i \rangle \langle q_i, t_i | Q(t_i) | q_{i-1}, t_{i-1} \rangle \dots \langle q_1, t_1 | q, t \rangle, \end{aligned} \quad (17.30)$$

where  $t_0 = t_i$ . Therefore,

$$\langle q_i, t_i | Q(t_i) | q_{i-1}, t_{i-1} \rangle = q_i \langle q_i, t_i | q_{i-1}, t_{i-1} \rangle. \quad (17.31)$$

Thus,

$$\langle q', t' | Q(t_0) | q, t \rangle = \int \mathcal{D}p \mathcal{D}q q(t_0) \exp \left\{ i \int_t^{t'} dt [p\dot{q} - H(p, q)] \right\}. \quad (17.32)$$

We now consider two operators  $Q(t_1)$  and  $Q(t_2)$ , with the condition

$$t' > t_1 > t_2 > t . \quad (17.33)$$

Then,  $\langle q', t' | Q(t_1) Q(t_2) | q, t \rangle$  can also be written in a form similar to (17.32). In general,

$$\begin{aligned} \langle q', t' | T^* [Q(t_1) \dots Q(t_N)] | q, t \rangle \\ = \int \mathcal{D}p \mathcal{D}q \, q(t_1) \dots q(t_N) \exp \left\{ i \int_t^{t'} dt [p\dot{q} - H(p, q)] \right\}, \end{aligned} \quad (17.34)$$

where  $t_1, \dots, t_N$  are between  $t'$  and  $t$ . In this way, if we compute the right-hand side of (17.34), it naturally becomes a matrix element of the time-ordered operator product.

We now compute the matrix element for the ground state. Assuming

$$T' > t' > t > T , \quad (17.35)$$

we include an external field between  $t'$  and  $t$ , and add a term  $J(t)q(t)$ :

$$\begin{aligned} \langle Q', T' | Q, T \rangle_J &= \int \mathcal{D}p \mathcal{D}q \exp \left\{ i \int_T^{T'} dt [p\dot{q} - H(p, q) - Jq] \right\} \\ &= \int dq' dq \langle Q', T' | q', t' \rangle \langle q', t' | q, t \rangle_J \langle q, t | Q, T \rangle . \end{aligned} \quad (17.36)$$

We consider  $|n\rangle$  to be an energy eigenstate such that

$$H|n\rangle = E_n|n\rangle , \quad \langle q|n\rangle = \varphi_n(q) . \quad (17.37)$$

Therefore,

$$\begin{aligned} \langle q, t | Q, T \rangle &= \langle q | \exp[-iH(t - T)] | Q \rangle \\ &= \sum_n \varphi_n(q) \varphi_n^*(Q) \exp[-iE_n(t - T)] . \end{aligned} \quad (17.38)$$

In order to read off the ground state, identified by  $n = 0$ , from this sum, when we set  $T \rightarrow i\infty$ , only  $n = 0$  survives, according to the third law of thermodynamics:

$$\lim_{T \rightarrow i\infty} \exp(-iE_0 T) \langle q, t | Q, T \rangle = \varphi_0(q, t) \varphi_0^*(Q) , \quad (17.39)$$

where  $\varphi_0(q, t) = \varphi_0(q) \exp(-iE_0 t)$ . Similarly,

$$\lim_{\substack{T' \rightarrow -i\infty \\ T \rightarrow i\infty}} \frac{\langle Q', T' | Q, T \rangle_J}{\exp[-iE_0(T' - T)] \varphi_0^*(Q) \varphi_0(Q')} = \int dq \int dq' \varphi_0^*(q', t') \langle q', t' | q, t \rangle_J \varphi_0(q, t). \quad (17.40)$$

The right-hand side of this equation can be written as  $\langle 0|0 \rangle_J$ . Since this is a generating functional, we write

$$\mathcal{F}[J] = \langle 0|0 \rangle_J. \quad (17.41)$$

Therefore, from the definition,

$$\begin{aligned} \frac{\delta^n \mathcal{F}[J]}{\delta J(t_1) \dots \delta J(t_n)} \Big|_{J=0} &= (-i)^n \langle 0 | T^* [Q(t_1) \dots Q(t_n)] | 0 \rangle \\ &= (-i)^n \int dq \int dq' \varphi_0^*(q', t') \varphi_0(q, t) \\ &\quad \times \int \mathcal{D}p \mathcal{D}q \, q(t_1) \dots q(t_n) \exp \left\{ i \int_t^{t'} dt [p\dot{q} - H(p, q)] \right\}. \end{aligned} \quad (17.42)$$

Moreover, apart from a factor that is independent of  $J$ ,

$$\mathcal{F}[J] \sim \lim_{\substack{T' \rightarrow -i\infty \\ T \rightarrow i\infty}} \langle Q', T' | Q, T \rangle_J. \quad (17.43)$$

Therefore, since the generating functional is normalized by  $\mathcal{F}[0] = \langle 0|0 \rangle = 1$ ,

$$\begin{aligned} \mathcal{F}[J] &= \frac{\lim \langle Q', T' | Q, T \rangle_J}{\lim \langle Q', T' | Q, T \rangle_{J=0}} \\ &= \frac{\lim \int \mathcal{D}q \exp \left\{ i \int_T^{T'} dt [L_{\text{eff}}(q, \dot{q}) - J(t)q(t)] \right\}}{\lim \int \mathcal{D}q \exp \left[ i \int_T^{T'} dt L_{\text{eff}}(q, \dot{q}) \right]}. \end{aligned} \quad (17.44)$$

where  $\lim$  means  $T' \rightarrow -i\infty$  and  $T \rightarrow i\infty$  as above. Since taking this limit corresponds to rotating 90 degrees clockwise with respect to the origin, we can also implement the analytic continuation by replacing  $t$  by  $t(1 - i\epsilon)$  and then taking the limit  $\epsilon \rightarrow +0$ . As an example, we consider the action integral when  $V = \omega^2 q^2/2$ . Replacing  $t$  by  $t(1 - i\epsilon)$  in

$$\exp \left( -\frac{i}{2} \omega^2 \int_T^{T'} dt q^2 \right) \quad (17.45)$$

is equivalent to replacing  $\omega^2$  as follows:

$$\omega^2 \rightarrow \omega^2(1 - i\epsilon) = \omega^2 - i\epsilon. \quad (17.46)$$

In the limit  $\epsilon \rightarrow +0$ , the generating function then becomes a generating function which gives a correct analytic continuation.

## 17.2 Quantization of Fields

In field theory, the field quantity  $\varphi(\mathbf{x})$  at each point in space is an independent dynamical variable. Thus, if we make the replacement

$$q_n(t) \rightarrow \varphi(\mathbf{x}, t), \quad (17.47)$$

then everything is exactly the same as before. In the limit, the generating functional  $\mathcal{T}[J]$  behaves as follows:

$$\mathcal{T}[J] \sim \lim \int \mathcal{D}\varphi \mathcal{D}\pi \exp \left\{ i \int d^4x [\pi(x)\dot{\varphi}(x) - \mathcal{H}(\varphi(x), \pi(x)) - J(x)\varphi(x)] \right\}, \quad (17.48)$$

$$\left. \frac{\delta^n \mathcal{T}[J]}{\delta J(x_1) \dots \delta J(x_n)} \right|_{J=0} = (-i)^n \langle \mathbf{0} | T^*[\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle. \quad (17.49)$$

In particular, in a simple linear theory,

$$\mathcal{T}[J] \sim \lim \int \mathcal{D}\varphi \exp \left\{ i \int d^4x [\mathcal{L}(\varphi(x)) - J(x)\varphi(x)] \right\}. \quad (17.50)$$

In order to implement the analytic continuation, instead of taking the limit, as in the case of a point-particle system, we make the replacement  $m^2 \rightarrow m^2 - i\epsilon$ . As an example, we consider a free scalar field:

$$\mathcal{L} = -\frac{1}{2}[(\partial_\lambda \varphi)^2 + m^2 \varphi^2]. \quad (17.51)$$

We start by evaluating the integral

$$I = \int \prod_j dx_j \exp \left[ -i \left( \sum_{j,k} \frac{1}{2} A_{jk} x_j x_k + \sum_j x_j y_j \right) \right], \quad (17.52)$$

where  $A$  is a real symmetric matrix. To carry out this integration, we make the change of variables

$$x_i = -(A^{-1})_{ij}y_j + z_i . \quad (17.53)$$

Therefore,

$$\frac{1}{2} \sum_{j,k} A_{j,k} x_j x_k + \sum_j x_j y_j = \frac{1}{2} \sum_{j,k} A_{j,k} z_j z_k - \frac{1}{2} \sum_{j,k} (A^{-1})_{j,k} y_j y_k . \quad (17.54)$$

Thus,

$$\begin{aligned} I &= \int \prod_j dz_j \exp \left[ -i \left( \frac{1}{2} \sum_{j,k} A_{j,k} z_j z_k - \frac{1}{2} \sum_{j,k} (A^{-1})_{j,k} y_j y_k \right) \right] \\ &\propto \exp \left[ \frac{i}{2} \sum_{j,k} (A^{-1})_{j,k} y_j y_k \right] . \end{aligned} \quad (17.55)$$

We then make the substitutions

$$x_i \rightarrow \varphi(x) , \quad y_i \rightarrow J(x) , \quad \sum \rightarrow \int d^4x , \quad A_{jk} \rightarrow A(x-y) . \quad (17.56)$$

Hence, for the scalar field described by the Lagrangian (17.51),

$$\begin{aligned} A(x-y) &= -(\square_x - m^2)\delta^4(x-y) , \\ A^{-1}(x-y) &= \frac{1}{(2\pi)^4} \int d^4p \frac{e^{ip \cdot (x-y)}}{p^2 + m^2 - i\epsilon} = i\Delta_F(x-y) . \end{aligned} \quad (17.57)$$

Thus,

$$\mathcal{Z}[J] = \exp \left[ -\frac{1}{2} \int d^4x \int d^4y J(x) \Delta_F(x-y) J(y) \right] . \quad (17.58)$$

In this case, the normalization condition holds true automatically, i.e.,

$$\mathcal{Z}[0] = 1 . \quad (17.59)$$

If there are interactions, we have to expand as a power series in the coupling constants.

We now derive the effective Lagrangian discussed in Sect. 11.5 using the path-integral method. If  $C$  is the inverse matrix of  $D$ ,

$$\mathcal{L} = -\frac{1}{2}D_{ab}(\varphi)\partial_\lambda\varphi_a\partial_\lambda\varphi_b - V(\varphi) , \quad (17.60)$$

$$\pi_a = D_{ab}\dot{\varphi}_b , \quad (17.61)$$

$$\mathcal{H} = \frac{1}{2}C_{ab}\pi_a\pi_b + \frac{1}{2}D_{ab}\nabla\varphi_a\nabla\varphi_b + V . \quad (17.62)$$

The generating functional  $\mathcal{T}[J]$  then has the form

$$\mathcal{T}[J] = \int \mathcal{D}\varphi \mathcal{D}\pi \exp \left\{ i \int d^4x [\pi_a(x)\dot{\varphi}_a(x) - \mathcal{H}(x) - J_a(x)\varphi_a(x)] \right\} . \quad (17.63)$$

To carry out the integral over  $\pi$ , we make the change of variables

$$\pi_a = \Pi_a + D_{ab}\dot{\varphi}_b , \quad (17.64)$$

$$\pi_a\dot{\varphi}_a - \frac{1}{2}C_{ab}\pi_a\pi_b = D_{ab}\dot{\varphi}_a\dot{\varphi}_b - \frac{1}{2}C_{ab}\Pi_a\Pi_b . \quad (17.65)$$

Therefore,

$$\mathcal{T}[J] = \int \mathcal{D}\varphi \mathcal{D}\Pi \exp \left\{ i \int d^4x \left[ \mathcal{L}(x) - \frac{1}{2}C_{ab}\Pi_a(x)\Pi_b(x) - J_a(x)\varphi_a(x) \right] \right\} . \quad (17.66)$$

Integrating over  $\Pi$ , the following factor shows up:

$$\begin{aligned} \prod_x \sqrt{\frac{\det C^{-1}}{2\pi i \Delta}} &\propto \prod_x \sqrt{\det D} , \quad \Delta = d^4x , \\ &= \exp \left[ \frac{1}{\Delta} \sum_x \frac{\Delta}{2} \ln(\det D) \right] \\ &= \exp \left[ \delta^4(0) \int d^4x \frac{1}{2} \ln(\det D) \right] . \end{aligned} \quad (17.67)$$

This corresponds to replacing  $\mathcal{L}(x)$  by the effective Lagrangian density

$$\mathcal{L}_{\text{eff}} = \mathcal{L} - \frac{i}{2} \delta^4(0) \ln(\det D) . \quad (17.68)$$

This is what we used in (11.113). While the derivation based on the canonical formalism was complex, the path-integral method has the advantage of being extremely intuitive and simple.



# Chapter 18

## Quantization of Gauge Fields Using the Path-Integral Method



The quantization of the electromagnetic field, which is the simplest gauge field, was discussed in Chap. 5. The problem arising in that case was that, if we apply the usual quantization discussed in Chap. 4 using the gauge invariant Lagrangian density, then the operator  $D(\partial)$  defined in (4.43) does not have an inverse. A related difficulty should thus arise in the path-integral method. In this chapter, we shall see how we can avoid this difficulty.

### 18.1 Quantization of Gauge Fields

Faddeev and Popov showed for the first time how to quantize gauge fields using the path-integral method [164]. Given a field  $A$ , in order to compute its Green's functions, we need to introduce the action integral  $S[A]$ :

$$S[A] = \int d^4x \mathcal{L}(x). \tag{18.1}$$

The vacuum expectation value of an arbitrary operator  $F[A]$  containing the field  $A$  is given by

$$\langle F[A] \rangle = \frac{\int \mathcal{D}A F[A] \exp \{iS[A]\}}{\int \mathcal{D}A \exp \{iS[A]\}}. \tag{18.2}$$

However, the denominator and numerator are both divergent. The reason is that, if  $A$  is a gauge field, then since all configurations of the gauge field which can be obtained by gauge transformations from any given configuration correspond to exactly the same state physically, the same physical state will appear infinitely many

times. This divergence corresponds to the fact that the differential operator  $D(\partial)$  does not have an inverse. However, if  $F[A]$  is a gauge invariant quantity, then such divergences will cancel between the denominator and the numerator.

Hence, if we introduce a gauge function  $\Omega$  and write the corresponding transformation for the gauge field  $A$  as

$$A(x) \rightarrow A^\Omega(x) , \quad (18.3)$$

we need to divide each path integral in the denominator and in the numerator of (18.2) by

$$\int \mathcal{D}\Omega . \quad (18.4)$$

Faddeev and Popov provided a method to do this. We arrange for one configuration of the gauge field to correspond to one physical state. The condition for picking one configuration is called the *gauge condition*.

### 18.1.1 A Method to Specify the Gauge Condition

We can specify the gauge condition using a functional  $f[A]$  of  $A$  and writing

$$f[A] = 0 . \quad (18.5)$$

Alternatively, for arbitrary  $A$ , we can choose a suitable  $\Omega$  and require

$$f[A^\Omega] = 0 . \quad (18.6)$$

We can then define the gauge-invariant functional  $\Delta_f[A]$  by

$$\Delta_f[A] \int \mathcal{D}\Omega \delta(f[A^\Omega]) = \text{const.} \quad (18.7)$$

Therefore,

$$\frac{\int \mathcal{D}A \exp \{iS[A]\}}{\int \mathcal{D}\Omega} \sim \int \mathcal{D}\mathcal{A} \exp \{iS[A]\} \Delta_f[A] \delta(f[A]) . \quad (18.8)$$

If we now introduce an external field, the generating functional of the Green's function is

$$\mathcal{T}_f[J] \sim \int \mathcal{D}A \exp \left\{ i \int d^4x [\mathcal{L}(x) - J(x)A(x)] \right\} \Delta_f[A] \delta(f[A]) . \quad (18.9)$$

### 18.1.2 The Additional Term Method

Adding a term to  $S$  and setting

$$\varphi[A] \int \mathcal{D}\Omega \exp \{ i \Delta S[A^\Omega] \} = \text{const.} , \quad \Delta S = \int d^4x \Delta \mathcal{L}(x) , \quad (18.10)$$

as in the case above, the equation corresponding to (18.9) is

$$\mathcal{T}[J] \sim \int \mathcal{D}A \exp \left\{ i \int d^4x [\mathcal{L}(x) + \Delta \mathcal{L}(x) - J(x)A(x)] \right\} \varphi[A] . \quad (18.11)$$

This is the customary way of quantizing gauge fields. We now turn to examples.

## 18.2 Quantization of the Electromagnetic Field

We apply the above method for quantizing gauge fields to the case of the electromagnetic field, which is the best known Abelian gauge field. The canonical quantization of the electromagnetic field is well understood. We shall now check whether the same result can be obtained using the path-integral method. The Lagrangian is

$$\begin{aligned} \mathcal{L} &= -\bar{\psi} [\gamma_\mu (\partial_\mu - ie A_\mu) + m] \psi - \frac{1}{4} F_{\mu\nu} F_{\mu\nu} , \\ &= \mathcal{L}_f + \mathcal{L}_{\text{int}} , \end{aligned} \quad (18.12)$$

where  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  and

$$\mathcal{L}_{\text{int}} = ie \bar{\psi} \gamma_\mu \psi A_\mu = j_\mu A_\mu . \quad (18.13)$$

This Lagrangian density is invariant under the gauge transformations

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \lambda(x) , \quad \psi(x) \rightarrow e^{ie\lambda(x)} \psi(x) , \quad \bar{\psi}(x) \rightarrow e^{-ie\lambda(x)} \bar{\psi}(x) . \quad (18.14)$$

### 18.2.1 Specifying the Gauge Condition

Here we consider the Lorenz gauge and the Coulomb (or radiation) gauge:

$$f_L[A] = \partial_\mu A_\mu = 0, \quad f_R[A] = \text{div } \mathbf{A} = 0. \quad (18.15)$$

In both cases, if we implement the gauge transformation with the gauge function  $-\lambda$ , then

$$\Delta_L[A] \int \mathcal{D}\lambda \delta(\partial_\mu A_\mu - \square\lambda) = \text{const.}, \quad (18.16)$$

$$\Delta_R[A] \int \mathcal{D}\lambda \delta(\text{div } \mathbf{A} - \Delta\lambda) = \text{const.}, \quad (18.17)$$

noting that  $\Delta$  does not depend on  $A$  in either case. For this reason, QED remains simple. We thus consider the generating functional in the Lorenz gauge:

$$\mathcal{T}[J, \eta, \bar{\eta}] \sim \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp \left[ i \int d^4x (\mathcal{L} - \bar{\eta}\psi - \bar{\psi}\eta - J_\mu A_\mu) \right] \delta(\partial_\mu A_\mu). \quad (18.18)$$

Setting  $e = 0$ , we carry out the path integral for the free field.

For the fermionic field, we take  $\eta$  and  $\bar{\eta}$  to be anti-commuting  $c$ -numbers and consider

$$\int \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp \left\{ -i \int d^4x [\bar{\psi}(\gamma_\mu \partial_\mu + m)\psi + \bar{\eta}\psi + \bar{\psi}\eta] \right\}. \quad (18.19)$$

In order to evaluate this integral, we generalize the example in Sect. 17.2. We define an inner product by

$$(x, Ax) = \sum_{j,k} x_j^* A_{jk} x_k. \quad (18.20)$$

We then write a generalization of the integral (17.52):

$$I = \int \prod_j dx_j \prod_k dx_k^* \exp \left\{ -i [(x, Ax) + (x, y) + (y, x)] \right\}. \quad (18.21)$$

Introducing the change of variables

$$x_j = -(A^{-1})_{jk} y_k + z_j, \quad (18.22)$$

we can carry out the integral, viz.,

$$I = \int \prod_j dz_j \prod_k dz_k^* \exp \left\{ -i[(z, Az) - (y, A^{-1}y)] \right\} \propto \exp [i(y, A^{-1}y)]. \quad (18.23)$$

Making the same replacement as (17.56),

$$A^{-1}(x-y) = \frac{1}{(2\pi)^4} \int d^4p \frac{e^{ip \cdot (x-y)}}{ip \cdot \gamma + m - i\epsilon} = iS_F(x-y). \quad (18.24)$$

The integral (18.19) thus assumes the form

$$\exp \left[ - \int d^4x \int d^4y \bar{\eta}(x) S_F(x-y) \eta(y) \right]. \quad (18.25)$$

Although we should in fact take into account the anti-commutativity of the variables  $\psi$ ,  $\bar{\psi}$ ,  $\eta$ , and  $\bar{\eta}$ , here we have just given the result by analogy.

Now, for the electromagnetic field, using

$$\delta(\partial_\mu A_\mu) \sim \int \mathcal{D}B \exp \left[ i \int d^4x B(x) \partial_\mu A_\mu(x) \right] \quad (18.26)$$

and integrating by parts in the exponent above, we obtain

$$\int \mathcal{D}A_\mu \mathcal{D}B \exp \left[ -i \int d^4x \left\{ \frac{1}{2} [A_\mu (\partial_\mu \partial_\nu - \delta_{\mu\nu} \square) A_\nu + \partial_\mu B A_\mu - B \partial_\mu A_\mu] + J_\mu A_\mu \right\} \right]. \quad (18.27)$$

For a pair  $(A_\mu, B)$ , the operator corresponding to  $A_{ij}$  is then expressed by the matrix

$$\begin{pmatrix} \partial_\mu \partial_\nu - \delta_{\mu\nu} & \partial_\mu \\ -\partial_\nu & 0 \end{pmatrix}. \quad (18.28)$$

Its inverse matrix appears in the propagator. It can be shown to be

$$\begin{pmatrix} \frac{1}{\square} \left( \frac{\partial_\mu \partial_\nu}{\square} - \delta_{\mu\nu} \right) & -\frac{\partial_\mu}{\square} \\ \frac{\partial_\nu}{\square} & 0 \end{pmatrix}. \quad (18.29)$$

Inserting this, the path integral (18.27) becomes

$$\exp \left[ -\frac{1}{2} \int d^4x \int d^4y J_\mu(x) D_{\mu\nu}(x-y) J_\nu(y) \right], \quad (18.30)$$

where

$$D_{\mu\nu}(x) = \frac{-i}{(2\pi)^4} \int d^4k \frac{1}{k^2 - i\epsilon} \left( \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2 - i\epsilon} \right) e^{ik \cdot x}. \quad (18.31)$$

The propagator appearing here, corresponding to the gauge condition (18.15), is written in the Landau gauge. The introduction of the auxiliary field  $B$  has already been discussed in Sect. 15.5.

### 18.2.2 The Additional Term Method

As an additional term, we choose

$$\Delta S = \int d^4x \Delta \mathcal{L}, \quad \Delta \mathcal{L} = -\frac{1}{2\alpha} (\partial_\mu A_\mu)^2. \quad (18.32)$$

Since  $\varphi[A]$  does not depend on  $A$ , the generating functional has the simple form

$$\mathcal{P}[J, \eta, \bar{\eta}] \sim \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp \left[ i \int d^4x (\mathcal{L} + \Delta \mathcal{L} - J_\mu A_\mu - \bar{\eta} \psi - \bar{\psi} \eta) \right]. \quad (18.33)$$

The  $\psi$ -part is the same as above, but the propagator of the electromagnetic field is

$$D_{\mu\nu} = \frac{-i}{(2\pi)^4} \int d^4k \frac{1}{(k^2 - i\epsilon)^2} (k^2 \delta_{\mu\nu} - k_\mu k_\nu + \alpha k_\mu k_\nu) e^{ik \cdot x}, \quad (18.34)$$

where  $\alpha$  is a gauge parameter. This form coincides with the integral expression already derived in (12.251) with  $\sigma = 0$ .

### 18.2.3 Ward–Takahashi Identity

The path-integral method gives the same result as the canonical quantization. We can use this method to derive other properties, such as the Ward–Takahashi identity [118, 119].

The propagator of the electron in the Landau gauge is

$$\langle \psi(x) \bar{\psi}(y) \rangle_L = \frac{1}{N_L} \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} \psi(x) \bar{\psi}(y) e^{iS[\psi, \bar{\psi}, A]} \delta(\partial_\mu A_\mu), \quad (18.35)$$

where  $N_L$  is a normalization factor given by

$$N_L = \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{iS[\psi, \bar{\psi}, A_\mu]} \delta(\partial_\mu A_\mu) . \quad (18.36)$$

We make the change of variables

$$\psi \rightarrow \psi' = e^{ie\lambda} \psi , \quad \bar{\psi} \rightarrow \bar{\psi}' = e^{-ie\lambda} \bar{\psi} . \quad (18.37)$$

Changing the integration variables to  $\psi'$  and  $\bar{\psi}'$ , and then rewriting them again as  $\psi$  and  $\bar{\psi}$ , the expression (18.35) takes the form

$$\begin{aligned} \langle \boldsymbol{\psi}(x) \bar{\boldsymbol{\psi}}(y) \rangle_L &= \frac{1}{N_L} \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} \boldsymbol{\psi}(x) \bar{\boldsymbol{\psi}}(y) e^{iS[\psi, \bar{\psi}, A]} \delta(\partial_\mu A_\mu) \\ &\quad \times \exp \left\{ ie [\lambda(x) - \lambda(y)] - i \int d^4z j_\mu(z) \partial_\mu \lambda(z) \right\} , \end{aligned} \quad (18.38)$$

where the last term is originated from the electron part of the Lagrangian density. Carrying out the functional differentiation of this equation with respect to  $\lambda(x)$ , and then setting that  $\lambda = 0$ ,

$$ie [\delta^4(x-z) - \delta^4(y-z)] \langle \boldsymbol{\psi}(x), \bar{\boldsymbol{\psi}}(y) \rangle_L + i \partial_\mu \langle \mathbf{j}_\mu(z), \boldsymbol{\psi}(x), \bar{\boldsymbol{\psi}}(y) \rangle_L = 0 . \quad (18.39)$$

Taking the Fourier transform of this equation, we obtain the Ward–Takahashi identity (12.200):

$$-i(p-q)_\mu S'_F(p) \Gamma_\mu(p, q) S'_F(q) = S'_F(p) - S'_F(q) . \quad (18.40)$$

The discussion about the derivation above only refers to the fermionic (electron) part, and not the electromagnetic field, so it turns out that this result holds true for any gauge fields.

Next, we discuss the gauge transformations for Green's functions.

### 18.2.4 Gauge Transformations for Green's Functions

We ask ourselves what kind of relations exist among Green's functions in different gauges. As an example, we investigate the relation between the Landau gauge and the radiation gauge, viz.,

$$\langle \boldsymbol{\psi}(x) \bar{\boldsymbol{\psi}}(y) \rangle_R = \frac{1}{N_R} \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} \boldsymbol{\psi}(x) \bar{\boldsymbol{\psi}}(y) e^{iS[\psi, \bar{\psi}, A]} \delta(\text{div } \mathbf{A}) , \quad (18.41)$$

$$N_R = \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{iS[\psi, \bar{\psi}, A]} \delta(\text{div } \mathbf{A}) . \quad (18.42)$$

We insert the following factor into the denominator and the numerator:

$$\int \mathcal{D}\lambda \delta(\square\lambda - \partial_\mu A_\mu) = \text{const.} \quad (18.43)$$

If we now carry out a gauge transformation, then  $S$  and  $\mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi}$  are invariant. In the integral,

$$\delta(\square\lambda - \partial_\mu A_\mu) \rightarrow \delta(\partial_\mu A_\mu), \quad \delta(\text{div } \mathbf{A}) \rightarrow \delta(\text{div } \mathbf{A} + \Delta\lambda), \quad (18.44)$$

and in the numerator, the following factor shows up:

$$\exp \{ie[\lambda(x) - \lambda(y)]\}. \quad (18.45)$$

From (18.44),

$$\text{div } \mathbf{A} + \Delta\lambda = 0, \quad (18.46)$$

so by solving this equation, we can find the constraint on  $\lambda$  :

$$\lambda(x) = \int d^4z \phi(x-z) \cdot \mathbf{A}(z), \quad (18.47)$$

$$\phi(x) = -\delta(x_0) \nabla \left( \frac{1}{4\pi|\mathbf{x}|} \right). \quad (18.48)$$

When we carry out the functional integration with respect to  $\lambda$ ,  $\delta(\text{div } \mathbf{A} + \Delta\lambda)$  disappears, whence

$$\langle \psi(x) \bar{\psi}(y) \rangle_{\text{R}} = \left\langle \psi(x) \bar{\psi}(y) \exp \left\{ ie \int d^4z [\phi(x-z) - \phi(y-z)] \mathbf{A}(z) \right\} \right\rangle_{\text{L}}. \quad (18.49)$$

Thus the propagator in the radiation gauge has been expressed in terms of the propagator in the Landau gauge.

### 18.3 Quantization of Non-Abelian Gauge Fields

Using the standard path-integral method for quantizing gauge fields, we consider the non-Abelian gauge fields.



### 18.3.1 A Method to Specify the Gauge Condition

We choose the gauge condition

$$f[A] = \partial_\mu A_\mu^a = 0. \quad (18.50)$$

Applying an infinitesimal gauge transformation, from (14.23),

$$\delta A_\mu^a = \frac{1}{g} \partial_\mu \lambda^a + f_{abc} A_\mu^b \lambda^c \equiv \frac{1}{g} (\mathcal{D}_\mu \lambda)^a. \quad (18.51)$$

Thus, under an infinitesimal gauge transformation,

$$(\partial_\mu A_\mu^a)^{\Omega^2} = \partial_\mu (A_\mu^a + \delta A_\mu^a) = \partial_\mu \left[ A_\mu^a + \frac{1}{g} (D_\mu \lambda)^a \right]. \quad (18.52)$$

In the usual way, we compute  $\Delta_f[A]$ . If we use (18.7), we have

$$\begin{aligned} \int \mathcal{D}\Omega \delta[(\partial_\mu A_\mu^a)^{\Omega^2}] &= \int \mathcal{D}\lambda \delta[(\partial_\mu A_\mu^a)^{\Omega^2}] \\ &= \int \mathcal{D}[(\partial_\mu A_\mu^a)^{\Omega^2}] \left\{ \frac{\mathcal{D}[(\partial_\mu A_\mu^a)^{\Omega^2}]}{\mathcal{D}\lambda} \right\}^{-1} \delta[(\partial_\mu A_\mu^a)^{\Omega^2}] \\ &= \left\{ \frac{\mathcal{D}[(\partial_\mu A_\mu^a)^{\Omega^2}]}{\mathcal{D}\lambda} \right\}^{-1}. \end{aligned} \quad (18.53)$$

Thus,  $\Delta_f[A]$  is the functional Jacobian

$$\Delta_f[A] = \frac{\mathcal{D}[(\partial_\mu A_\mu^a)^{\Omega^2}]}{\mathcal{D}\lambda} = \det \left( -\frac{1}{g} \partial_\mu D_\mu \right). \quad (18.54)$$

Normalizing this determinant to unity when  $A_\mu^a = 0$ ,

$$\Delta_f[A] = \det \left( \frac{\partial_\mu D_\mu}{\square} \right). \quad (18.55)$$

To compute this expression, we use the method due to 't Hooft in 1971 [165]. The generating functional for the Green's functions is

$$\mathcal{T}[J] = \int \mathcal{D}A_\mu \exp \{ iS[A, J] \} \Delta_f[A] \delta(\partial_\mu A_\mu), \quad (18.56)$$

where

$$S[A, J] = S[A] - \int d^4x A_\mu^a(x) J_\mu^a(x). \quad (18.57)$$

The  $\delta$ -function in the integral is given by

$$\delta(\partial_\mu A_\mu) \sim \int \mathcal{D}B \exp \left[ i \int d^4x B(x) \partial_\mu A_\mu(x) \right], \quad (18.58)$$

but note that we do not discuss the normalization here. We use

$$\int \prod_j dx_j dy_j \exp [i(x, Ay)] = (2\pi)^n (\det A)^{-1}. \quad (18.59)$$

And so we obtain

$$\det \left( \frac{\partial_\mu D_\mu}{\square} \right)^{-1} \sim \int \mathcal{D}\varphi \mathcal{D}\bar{\varphi} \exp \left[ i \int d^4x \bar{\varphi}(x) \partial_\mu D_\mu \varphi(x) \right]. \quad (18.60)$$

Note also that (18.60) is the inverse of (18.55). Let us therefore consider how to obtain the inverse.

We treat the expression (18.60) as a sum of loop contributions obtained by contractions among the scalar fields  $\varphi$  and  $\bar{\varphi}$ , while the gauge field  $A$  appears as an external line. According to the discussion in Sect. 11.2, this sum is the connected part, so in order to derive the inverse, we need to invert the sign of the connected part. However, the connected part consists of single loops obtained by contracting  $\varphi$  and  $\bar{\varphi}$  before  $A$  is quantized. Thus, we must reverse the sign of each loop. As mentioned in Sect. 8.4, this reversal happens when  $\varphi$  and  $\bar{\varphi}$  are anti-commutative, i.e., when they obey Fermi statistics. For a path integral involving these so-called *Grassmann numbers*, which anti-commute, we need an additional discussion, but for the moment we avoid getting further involved and just write down the result:

$$\Delta_f[A] \sim \int \mathcal{D}\varphi \mathcal{D}\bar{\varphi} \exp \left[ i \int d^4x \bar{\varphi}(x) \partial_\mu D_\mu \varphi(x) \right]. \quad (18.61)$$

Although  $\varphi$  and  $\bar{\varphi}$  are scalar fields, they obey Fermi statistics. It turns out that this introduces an indefinite metric. The effective Lagrangian density in this theory is

$$\mathcal{L}[A] + \bar{\varphi} \partial_\mu D_\mu \varphi + B \partial_\mu A_\mu. \quad (18.62)$$

Here we have summed indices standing for components, although this has not been written explicitly. This Lagrangian corresponds to the one in the Landau gauge in QED. The scalar fields  $\varphi$  and  $\bar{\varphi}$  are called *Faddeev–Popov ghost fields*.

### 18.3.2 The Additional Term Method

We take  $\Delta\mathcal{L}$  to be expressed in terms of  $\partial_\mu A_\mu$  and define  $\varphi[A]$  by

$$\varphi[A] \int \mathcal{D}\Omega \exp \left\{ i\Delta S[(\partial_\mu A_\mu)^{\Omega^2}] \right\} = \text{const.} \quad (18.63)$$

In order to compute this, we use

$$\exp \left\{ i\Delta S[(\partial_\mu A_\mu)^{\Omega^2}] \right\} \sim \int \mathcal{D}C e^{i\Delta S[C]} \delta[(\partial_\mu A_\mu)^{\Omega^2} - C]. \quad (18.64)$$

This computation is the same as the example above, so  $\varphi[A]$  can be readily derived. As a consequence, the effective Lagrangian density, considering the first term as a gauge-invariant term, is

$$\mathcal{L}[A] + \bar{\varphi} \partial_\mu D_\mu \varphi + \Delta\mathcal{L}. \quad (18.65)$$

We choose the following form for  $\Delta\mathcal{L}$ :

$$\Delta\mathcal{L} = -\frac{1}{2\alpha} (\partial_\mu A_\mu)^2. \quad (18.66)$$

Therefore, the total Lagrangian density is

$$\mathcal{L}[A] + \bar{\varphi} \partial_\mu D_\mu \varphi - \frac{1}{2\alpha} (\partial_\mu A_\mu)^2. \quad (18.67)$$

### 18.3.3 Hermitization of the Lagrangian Density

In the discussion so far, we used the effective Lagrangian density to compute the  $S$ -matrix and Green's functions. In the operator formalism, the Lagrangian density should be Hermitian. The Faddeev–Popov ghost term in (18.67) is not Hermitian. Integrating this term by parts,

$$\mathcal{L}_{\text{FP}} \sim -\partial_\mu \bar{\varphi} D_\mu \varphi = -\partial_\mu \bar{\varphi}^a (\partial_\mu \varphi^a + g f_{abc} A_\mu^b \varphi^c). \quad (18.68)$$

If  $\varphi$  and  $\bar{\varphi}$  are Hermitian, then (18.67) is obviously not Hermitian. This is because  $\varphi$  and  $\bar{\varphi}$  are anti-commutative scalar fields. We thus change the phase of this part:

$$\mathcal{L}_{\text{FP}} \rightarrow e^{i\alpha} \mathcal{L}_{\text{FP}}. \quad (18.69)$$

Consequently, the phases of the ghost propagator and the coupling constant for the ghost and the gauge field change according to

$$\langle \varphi(x) \bar{\varphi}(y) \rangle = D_F(x-y) \rightarrow e^{-i\alpha} D_F(x-y), \quad (18.70)$$

$$g \rightarrow g e^{i\alpha}. \quad (18.71)$$

Note that, when  $\varphi$  and  $\bar{\varphi}$  only appear in closed loops, the numbers of  $D_F$  and  $g$  are the same, so the contributions to the  $S$ -matrix or Green's functions are invariant under the above phase transformation. That is, it turns out that the phase  $\alpha$  can be freely chosen. If we choose  $e^{i\alpha} = -i$ , and write  $c$  and  $\bar{c}$  instead of  $\varphi$  and  $\bar{\varphi}$  [see (4.89)], we have

$$c^\dagger = c, \quad \bar{c}^\dagger = \bar{c}, \quad (18.72)$$

$$\mathcal{L}_{\text{FP}} = i \partial_\mu \bar{c} D_\mu c. \quad (18.73)$$

This Lagrangian density is then Hermitian.

The Lagrangian density in a general gauge (also called the  $\alpha$ -gauge) is

$$\mathcal{L} = \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{FP}}, \quad (18.74)$$

where, dropping indices for the gauge field,

$$\mathcal{L}_{\text{inv}} = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu}, \quad (18.75)$$

$$\mathcal{L}_{\text{GF}} = (\partial_\mu B) A_\mu + \frac{\alpha}{2} B \cdot B, \quad \text{or} \quad -\frac{1}{2\alpha} (\partial_\mu A_\mu)^2, \quad (18.76)$$

$$\mathcal{L}_{\text{FP}} = i \partial_\mu \bar{c} D_\mu c. \quad (18.77)$$

The first term in (18.74) is gauge invariant, the second is a gauge-fixing term, and the third is a ghost term. This form was given by Kugo and Ojima in [166].

### 18.3.4 Gauge Transformations of Green's Functions

In the last section, we investigated the relations among Green's functions defined using different gauge conditions in QED. Here we discuss the different relations among Green's functions defined by including an additional term. The gauge-invariant term is the same, and we thus treat two theories which are physically equivalent in different gauges. Hence, we introduce two Lagrangian densities and two action integrals:

$$\mathcal{L}_{II} = \mathcal{L}_I + \Delta \mathcal{L}, \quad S_{II} = S_I + \Delta S. \quad (18.78)$$

Then considering the field operators  $A, B, C, \dots$ , we introduce the Green's function in the second gauge:

$$\langle \mathbf{ABC} \dots \rangle_{II} = \frac{1}{N_{II}} \int \mathcal{D}A_\mu \dots \mathbf{ABC} \dots \exp(iS_{II}), \quad (18.79)$$

$$N_{II} = \int \mathcal{D}A_\mu \dots \exp(iS_{II}). \quad (18.80)$$

Then from (18.78), we decompose  $S_{II}$ , considering  $\Delta S$  as a perturbation and treating  $\exp(i\Delta S)$  like  $A, B, C, \dots$ . Therefore,

$$\begin{aligned} \langle \mathbf{ABC} \dots \rangle_{II} &= \frac{1}{N_{II}} \int \mathcal{D}A_\mu \dots \mathbf{ABC} \dots \exp(i\Delta S) \exp(iS_I) \\ &= \frac{N_I}{N_{II}} \frac{1}{N_I} \int \mathcal{D}A_\mu \dots \mathbf{ABC} \dots \exp(i\Delta S) \exp(iS_I) \\ &= \frac{N_I}{N_{II}} \langle \mathbf{ABC} \dots \exp(i\Delta S) \rangle_I, \end{aligned} \quad (18.81)$$

$$\frac{N_I}{N_{II}} = \frac{1}{N_I} \int \mathcal{D}A_\mu \dots \exp(i\Delta S) \exp(iS_I) = \langle \exp(i\Delta S) \rangle_I. \quad (18.82)$$

We thus obtain

$$\langle \mathbf{ABC} \dots \rangle_{II} = \frac{\langle \mathbf{ABC} \dots \exp(i\Delta S) \rangle_I}{\langle \exp(i\Delta S) \rangle_I}. \quad (18.83)$$

This gives the relation among Green's functions in two different gauges. For example, considering (18.74), we choose the Landau gauge with  $\alpha = 0$  and the gauge with  $\alpha \neq 0$  for  $\mathcal{L}_I$  and  $\mathcal{L}_{II}$ , respectively, and distinguish the Landau gauge by the index  $L$ . Then,

$$\langle \mathbf{ABC} \dots \rangle_\alpha = \frac{\left\langle \mathbf{ABC} \dots \exp \left[ \frac{i\alpha}{2} \int d^4x \mathbf{B}(x) \cdot \mathbf{B}(x) \right] \right\rangle_L}{\left\langle \exp \left[ \frac{i\alpha}{2} \int d^4x \mathbf{B}(x) \cdot \mathbf{B}(x) \right] \right\rangle_L}. \quad (18.84)$$

This equation shows the  $\alpha$ -dependence of an arbitrary Green's function. We may also interpret operators appearing in the discussion above as being *unrenormalized*. The subscript  $L$  indicates that these Green's functions should be evaluated in Heisenberg's picture in the Landau gauge, while the subscript  $\alpha$  indicates that they should be evaluated in the Heisenberg picture in the gauge  $\alpha \neq 0$ . This formula provides a basis for the discussion about the gauge invariance of various kinds of Green's functions.

## 18.4 Axial Gauge

In the last section, we introduced the effective Lagrangian density in the covariant gauge, and thereby understood the need for the Faddeev–Popov ghost. However, if we do not require manifest Lorentz covariance, there is a gauge in which we can quantize without ghosts. This is the axial gauge.

We replace the gauge condition (18.50) discussed in the last section by

$$n_\mu A_\mu^a = 0, \quad (18.85)$$

where  $n_\mu$  is a constant vector. For the infinitesimal gauge transformation (18.51),

$$\delta(n_\mu A_\mu^a) = \frac{1}{g} n_\mu \partial_\mu \lambda^a + f_{abc} (n_\mu A_\mu^b) \lambda^c. \quad (18.86)$$

Under the gauge condition (18.85),

$$\delta(n_\mu A_\mu^a) = \frac{1}{g} n_\mu \partial_\mu \lambda^a. \quad (18.87)$$

This is independent of  $A_\mu$ . Hence,

$$\int \mathcal{D}\Omega \delta[(n_\mu A_\mu^a)^\Omega]$$

does not involve  $A_\mu$  and  $\Delta_f[A]$  is a constant. Therefore, it turns out that the Faddeev–Popov ghost term is not produced here.

Using the additional term method, if we choose

$$\Delta S[A] = -\frac{1}{2\alpha} \int d^4x (n_\mu A_\mu^a)^2, \quad (18.88)$$

and set

$$\varphi[A] \int \mathcal{D}\Omega \exp \{iS[A^\Omega]\} = \text{const.}, \quad (18.89)$$

then once again  $\varphi[A]$  does not involve  $A_\mu$ . Hence, we consider the effective Lagrangian density

$$\mathcal{L} = \mathcal{L}_{\text{inv}} - \frac{1}{2\alpha} (n_\mu A_\mu)^2. \quad (18.90)$$

Now quantizing, leaving only terms in quadratic form and applying the variation principle,

$$D_{\mu\nu}(\partial)A_\nu^a = 0, \quad (18.91)$$

where

$$D_{\mu\nu}(\partial) = (\delta_{\mu\nu} - \partial_\mu \partial_\nu) - \frac{1}{\alpha} n_\mu n_\nu. \quad (18.92)$$

Making the substitution  $\partial_\mu \partial_\nu \rightarrow -k_\mu k_\nu$  in momentum space,

$$D_{\mu\nu}(k) = -k^2 \delta_{\mu\nu} + k_\mu k_\nu - \frac{1}{\alpha} n_\mu n_\nu, \quad (18.93)$$

$$\left[ \delta_{\mu\nu} - \frac{n_\mu k_\lambda + n_\lambda k_\mu}{n \cdot k} + \frac{n^2 + \alpha k^2}{(n \cdot k)^2} k_\mu k_\nu \right] D_{\lambda\nu}(k) = -k^2 \delta_{\mu\nu}. \quad (18.94)$$

Thus, it turns out that the propagator is

$$\frac{1}{k^2 - i\epsilon} \left[ \delta_{\mu\nu} - \frac{n_\mu k_\nu + n_\nu k_\mu}{n \cdot k} + \frac{n^2 + \alpha k^2}{(n \cdot k)^2} k_\mu k_\nu \right]. \quad (18.95)$$

Problems with this gauge include the question of how to treat the pole  $n \cdot k = 0$ , and showing that computations of various physical quantities do not depend on the choice of  $n$ .

## 18.5 Feynman Rules in the $\alpha$ -Gauge

We now introduce the Feynman rules for the Lagrangian density (18.74). First, note that

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g(A_\mu \times A_\nu)^a, \quad (18.96)$$

$$(D_\mu c)^a = \partial_\mu c^a + g(A_\mu \times c)^a. \quad (18.97)$$

For the gauge group indices, we use the inner and the outer product symbols:

$$A \cdot B = \sum_a A^a B_a, \quad (A \times B)^a = \sum_{b,c} f_{abc} A^b B^c. \quad (18.98)$$

We split the Lagrangian density (18.74) into the free part  $\mathcal{L}_f$  and the interaction part  $\mathcal{L}_{\text{int}}$ :

$$\mathcal{L}_f = -\frac{1}{4}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)^2 - \frac{1}{2\alpha}(\partial_\mu A_\mu^a)^2 + i\partial_\mu \bar{c}\partial_\mu c, \quad (18.99)$$

$$\mathcal{L}_{\text{int}} = -g\partial_\mu A_\nu(A_\mu \times A_\nu) - \frac{1}{4}g^2(A_\mu \times A_\nu)(A_\mu \times A_\nu) + ig\partial_\mu \bar{c}(A_\mu \times c). \quad (18.100)$$

Expressing the Feynman rule in the Lagrangian formalism, it turns out that we assign the factors  $i(2\pi)^4$ ,  $-i/(2\pi)^4$ , and  $(-1)$  to each vertex, propagator, and closed ghost loop, respectively. Considering the gauge particle as the gluon, the propagators are

$$\text{gluon} \quad \frac{\delta_{ab}}{k^2 - i\epsilon} \left[ \delta_{\mu\nu} - (1 - \alpha) \frac{k_\mu k_\nu}{k^2 - i\epsilon} \right], \quad (18.101)$$

$$\text{ghost} \quad \frac{i\delta_{ab}}{k^2 - i\epsilon}. \quad (18.102)$$

Moreover, we find the following three types of vertex function:

1. three-gluon vertex,    2. four-gluon vertex,    3. ghost–gluon vertex.

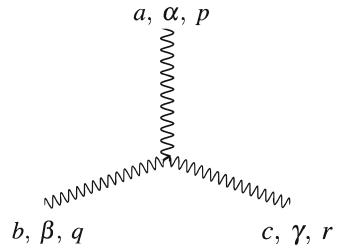
1. Three-gluon vertex (Fig. 18.1). Taking all the momenta of the incoming gluons, the vertex function is

$$-igf_{abc}[\delta_{\beta\gamma}(r - q)_\alpha + \delta_{\gamma\alpha}(p - r)_\beta + \delta_{\alpha\beta}(q - p)_\alpha]. \quad (18.103)$$

2. Four-gluon vertex (Fig. 18.2). In this case, the vertex function is

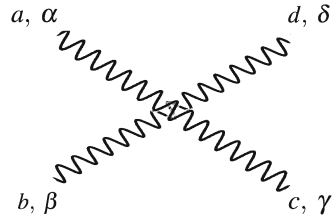
$$\begin{aligned} & -g^2 f_{gac} f_{gbd} (\delta_{\alpha\beta} \delta_{\gamma\delta} - \delta_{\alpha\delta} \delta_{\beta\gamma}) - g^2 f_{gad} f_{gbc} (\delta_{\alpha\beta} \delta_{\gamma\delta} - \delta_{\alpha\gamma} \delta_{\beta\delta}) \\ & - g^2 f_{gab} f_{gcd} (\delta_{\alpha\gamma} \delta_{\delta\beta} - \delta_{\alpha\delta} \delta_{\beta\gamma}). \end{aligned} \quad (18.104)$$

**Fig. 18.1** Three-gluon vertex

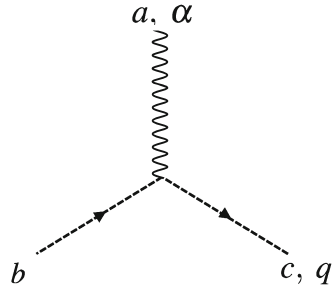




**Fig. 18.2** Four-gluon vertex



**Fig. 18.3** Ghost-gluon vertex



3. Ghost-gluon vertex (Fig. 18.3). Considering the ghost lines to be directed from  $\bar{c}$  to  $c$  and  $q$  to be the outgoing momentum, the vertex function is

$$- g f_{abc} q_\alpha . \tag{18.105}$$

Combining the above propagators and vertex functions, we can compute the  $S$ -matrix elements or Green's functions. Note that the four-momentum is conserved at each vertex. For the total amplitude, we then have conservation of four-momentum, viz., a factor

$$\delta^4(P_f - P_i) . \tag{18.106}$$

Moreover, we have to integrate over all the four-momenta  $k_i$  in closed loops, which are not affected by the overall conservation of four-momentum, i.e., we introduce the integrals

$$\int d^4k_1 \dots \int d^4k_l , \tag{18.107}$$

where  $l$  is the number of closed loops.

# Chapter 19

## Becchi–Rouet–Stora Transformations



We consider a system in which a gauge field and fundamental fermions interact with each other. In this case, local gauge transformations are defined for the gauge field and the fermion field. However, in order to quantize these fields, we need to introduce the gauge-fixing term and the Faddeev–Popov ghost term. Consequently, the invariance under the local gauge transformation is broken. However, a new global invariance shows up in its place. This is invariance under the BRS transformation found by Becchi et al. in [167]. In this chapter, we discuss the properties of this transformation.

### 19.1 BRS Transformations

The Lagrangian density for a system where a gauge field and a fermion field interact is written as

$$\mathcal{L} = \mathcal{L}_{\text{inv}} + \mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{FP}} , \quad (19.1)$$

which is the same as the one in (18.74). Regarding the gauge-fixing term and the ghost term, these are the same as given in (18.76) and (18.77), but the gauge-invariant term is

$$\mathcal{L}_{\text{int}} = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu} - \bar{\psi} (\gamma_{\mu} D_{\mu} + m) \psi , \quad (19.2)$$

where the covariant derivative for the fermion field is

$$D_{\mu} \psi = (\partial_{\mu} - ig t \cdot A_{\mu}) \psi = \left( \partial_{\mu} - ig \sum_a t_a A_{\mu}^a \right) \psi . \quad (19.3)$$

Local gauge transformations are now defined for the gauge field and the fermion field, which are called *fundamental fields*. For a gauge function  $\lambda(x)$ , the infinitesimal gauge transformations for the fundamental fields are

$$\delta A_\mu^a(x) = \frac{1}{g}(D_\mu \lambda(x))^a, \quad \delta \psi(x) = i[\lambda(x) \cdot t]\psi(x). \quad (19.4)$$

The transformations obtained by making the following replacement are called the *BRS transformations*:

$$\lambda^a(x) \rightarrow g c^a(x). \quad (19.5)$$

In the following, we indicate these transformations by the same symbol  $\delta$  used in (19.4).

Since  $c^a(x)$  are anti-commuting fields, these transformations change the statistics of the field operators. Taking this into account, the BRS transformations for fundamental fields are

$$\delta A_\mu = D_\mu c, \quad \delta \psi = ig(c \cdot t)\psi, \quad \delta \bar{\psi} = -ig\bar{\psi}(c \cdot t). \quad (19.6)$$

From these, the following equations are obtained directly:

$$\delta F_{\mu\nu} = g F_{\mu\nu} \times c, \quad (19.7)$$

$$\delta \mathcal{L}_{\text{inv}} = 0. \quad (19.8)$$

Note that the local gauge transformations are not defined for the auxiliary fields  $B$ ,  $c$ , and  $\bar{c}$ , so we cannot define the BRS transformation by the substitution (19.5). We thus introduce the BRS transformations by requiring the total Lagrangian density (19.1) to be invariant under the transformations. To do this, we first write down the field equations derived by the variational principle:

$$D_\mu F_{\mu\nu} + \partial_\nu B - ig\partial_\nu \bar{c} \times c + ig\bar{\psi}\gamma_\nu t\psi = 0, \quad (19.9)$$

$$\partial_\mu A_\mu = \alpha B, \quad (19.10)$$

$$\partial_\mu D_\mu c = D_\mu \partial_\mu \bar{c} = 0. \quad (19.11)$$

Assuming that  $\delta$  and  $\partial_\mu$  commute, Eqs. (19.10) and (19.11) imply

$$\alpha \delta B = \delta(\partial_\mu A_\mu) = \partial_\mu \delta A_\mu = \partial_\mu D_\mu c = 0.$$

Thus,

$$\delta B = 0. \quad (19.12)$$

Moreover, if we rewrite (19.11), we have

$$0 = \delta(\partial_\mu D_\mu c) = \partial_\mu \delta(D_\mu c) = \partial_\mu \delta^2 A_\mu .$$

We can thus set

$$\delta(D_\mu c) = \delta^2 A_\mu = 0 . \quad (19.13)$$

Finally, taking into account (19.8), we require

$$\delta \mathcal{L} = \delta(\mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{FP}}) = 0 . \quad (19.14)$$

Hence, considering (19.12) and (19.13),

$$\begin{aligned} \delta(\mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{FP}}) &= \delta(\partial_\mu B \cdot A_\mu + i\partial_\mu \bar{c} \cdot D_\mu c) \\ &= \partial_\mu B \cdot D_\mu c + i\partial_\mu (\delta \bar{c}) \cdot D_\mu c \\ &= \partial_\mu (B + i\delta \bar{c}) \cdot D_\mu c . \end{aligned}$$

In order for this to vanish, we must set

$$\delta \bar{c} = iB . \quad (19.15)$$

Finally, in order to determine  $\delta c$ , we consider

$$\begin{aligned} \delta(D_\mu c) &= \delta(\partial_\mu c) + g(D_\mu c) \times c + gA_\mu \times \delta c \\ &= D_\mu \left( \delta c + \frac{1}{2} g c \times c \right) = 0 . \end{aligned}$$

From this,

$$\delta c = -\frac{1}{2} g c \times c . \quad (19.16)$$

All the BRS transformations for auxiliary fields are thus determined:

$$\delta B = 0 , \quad \delta \bar{c} = iB , \quad \delta c = -\frac{1}{2} g c \times c , \quad (19.17)$$

and the total Lagrangian density is invariant under the BRS transformation.

## 19.2 BRS Charge

If a theory is BRS invariant, according to Noether's theorem, there should exist a conserved quantity corresponding to this invariance. This is called the *BRS charge*, denoted  $Q_B$ . In the following, we shall derive an expression for this quantity. First, we give the equal-time canonical commutation relations necessary for quantization. Using  $A_j^a$ ,  $j = 1, 2, 3$ , and  $B^a$  as independent variables, the field operators conjugate to each field operator are

$$\pi_j^a = iF_{4j}^a, \quad \pi_B^a = A_0^a = -iA_4^a. \quad (19.18)$$

Therefore, for  $x_0 = y_0$ ,

$$[A_4^a(x), B^b(y)] = \delta_{ab}\delta^3(x-y), \quad [A_j^a(x), F_{4k}^b(y)] = \delta_{ab}\delta_{jk}\delta^3(x-y). \quad (19.19)$$

For the ghost fields, if  $x_0 = y_0$ , we have

$$\{c^a(x), \dot{\bar{c}}^b(y)\} = -\delta_{ab}\delta^3(x-y), \quad \{D_4c^a(x), \bar{c}^b(y)\} = -i\delta_{ab}\delta^3(x-y). \quad (19.20)$$

We can now construct the Noether current:

$$\begin{aligned} J_\mu^B &= \sum_\alpha \delta\varphi_\alpha \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,\mu}} = -D_\nu c \cdot F_{\mu\nu} - ig\bar{\psi}\gamma_\mu(c \cdot t)\psi - B \cdot D_\mu c + \frac{i}{2}g\partial_\mu \bar{c} \cdot (c \times c) \\ &= -\partial_\nu (F_{\mu\nu} \cdot c) - B \cdot D_\mu c + \partial_\mu B \cdot c - \frac{i}{2}g\partial_\mu \bar{c} \cdot (c \times c). \end{aligned} \quad (19.21)$$

This is a conserved quantity. The spatial integral of its time component is  $Q_B$ :

$$\partial_\mu J_\mu^B = 0, \quad (19.22)$$

$$Q_B = \int d^3x J_0^B = \int d^3x \left[ B \cdot \dot{c} - \dot{B} \cdot c - gB \cdot (A_0 \times c) + \frac{i}{2}g\dot{\bar{c}} \cdot (c \times c) \right]. \quad (19.23)$$

Therefore, the BRS transformation can be written

$$\delta\phi = i[Q_B, \phi]_\mp. \quad (19.24)$$

If  $\phi$  is even-ordered with respect to  $c$  and  $\bar{c}$ , we choose  $(-)$ , and if it is odd-ordered, we choose  $(+)$ .

It is also a straightforward matter to show that

$$\delta \left[ -B \cdot D_\mu c + \partial_\mu B \cdot c - \frac{i}{2} g \partial_\mu \bar{c} \cdot (c \times c) \right] = 0 . \quad (19.25)$$

Thus,

$$\delta Q_B = \int d^3x \delta J_0^B = 0 , \quad (19.26)$$

whence

$$\delta Q_B = i \{ Q_B, Q_B \} = 2i Q_B^2 = 0 .$$

This proves the nilpotency of  $Q_B$  :

$$Q_B^2 = 0 . \quad (19.27)$$

This is the most remarkable property of the BRS charge. The BRS transformation itself has this property, i.e., using (19.27),

$$\delta^2 \phi = \begin{cases} -\{ Q_B, [Q_B, \phi] \} = 0 , & \phi \text{ is even-ordered w.r.t. } c \text{ and } \bar{c} , \\ -\{ Q_B, \{ Q_B, \phi \} \} = 0 , & \phi \text{ is odd-ordered w.r.t. } c \text{ and } \bar{c} . \end{cases} \quad (19.28)$$

We thus see that the BRS transformation is nilpotent.

Ghost terms appear now only in the ghost term, which is invariant under the scale transformation

$$c \rightarrow e^\lambda c , \quad \bar{c} \rightarrow e^{-\lambda} \bar{c} . \quad (19.29)$$

Since both  $c$  and  $\bar{c}$  are Hermitian, we cannot introduce a phase transformation. The Noether current corresponding to the above transformation is

$$J_\mu^c = \left( c \frac{\partial \mathcal{L}}{\partial c_{,\mu}} - \bar{c} \frac{\partial \mathcal{L}}{\partial \bar{c}_{,\mu}} \right) = i(\partial_\mu \bar{c} \cdot c - \bar{c} \cdot D_\mu c) , \quad (19.30)$$

$$\partial_\mu J_\mu^c = 0 . \quad (19.31)$$

We thus introduce the conserved quantity  $Q_c$  :

$$Q_c = \int d^3x J_0^c = i \int d^3x [\bar{c} \cdot \dot{c} - \dot{\bar{c}} \cdot c - g \bar{c} \cdot (A_0 \times c)] . \quad (19.32)$$

This quantity satisfies the commutation relations

$$i[Q_c, c(x)] = c(x) , \quad i[Q_c, \bar{c}(x)] = -\bar{c}(x) , \quad (19.33)$$

and  $Q_c$  commutes with all other fields. Therefore, in general,

$$i[Q_c, \phi] = N\phi , \quad (19.34)$$

where  $N$  is the number of  $\bar{c}$  minus the number of  $c$  included in  $\phi$  as factors.  $N$  is called the *ghost number of  $\phi$* . In addition, both  $Q_B$  and  $Q_c$  are Hermitian:

$$Q_B^\dagger = Q_B , \quad Q_c^\dagger = Q_c . \quad (19.35)$$

From the structure of  $Q_B$  given in (19.23), its ghost number is unity. Therefore, from (19.34),

$$i[Q_c, Q_B] = Q_B . \quad (19.36)$$

Combining (19.27) and (19.36) results in the *BRS algebra*. Since commutation relations and anti-commutation relations coexist in this algebra, it is called a *graded Lie algebra*, rather than a Lie algebra.

### 19.3 Another BRS Transformation

The BRS transformation can be obtained by replacing an infinitesimal gauge function by the ghost field  $c$ , as shown in (19.37). So what happens if we then replace an infinitesimal gauge function by  $\bar{c}$ ? This issue was investigated by Curci and Ferrari in [168], Ojima in [169], Bonora and Tonin in [170], and so on. For the fundamental fields, denoting such a transformation by the symbol  $\bar{\delta}$ ,

$$\bar{\delta}A_\mu = D_\mu \bar{c} , \quad \bar{\delta}\psi = ig(\bar{c} \cdot t)\psi , \quad \bar{\psi} = -ig\bar{\psi}(\bar{c} \cdot t) . \quad (19.37)$$

This transformation also leaves invariant the gauge invariant term, in exactly the same way as  $\delta$ . To determine the transformation of the auxiliary field, we note the identity

$$i\partial_\mu \bar{c} \cdot D_\mu c + \partial_\mu B \cdot A_\mu = iD_\mu \bar{c} \cdot \partial_\mu c - \partial_\mu \bar{B} \cdot A_\mu , \quad (19.38)$$

where  $\bar{B}$  is defined by

$$B + \bar{B} - igc \times \bar{c} = 0 . \quad (19.39)$$

It is easily shown that

$$\bar{\delta}\bar{B} = 0, \quad \bar{\delta}c = i\bar{B}, \quad \bar{\delta}\bar{c} = -\frac{1}{2}g\bar{c} \times \bar{c}. \quad (19.40)$$

Writing the BRS charge corresponding to this transformation as  $\bar{Q}_B$ , the BRS algebra mentioned in the previous section is extended as follows:

$$Q_B^2 = \bar{Q}_B^2 = Q_B\bar{Q}_B + \bar{Q}_BQ_B = 0, \quad (19.41)$$

$$i[Q_c, Q_B] = Q_B, \quad i[Q_c, \bar{Q}_B] = -\bar{Q}_B. \quad (19.42)$$

Moreover, from (19.41) we obtain

$$\delta^2 = \bar{\delta}^2 = \bar{\delta}\delta + \delta\bar{\delta} = 0. \quad (19.43)$$

In particular, the last equation is obtained from the identities

$$\begin{aligned} \{Q_B, [\bar{Q}_B, \phi]\} + \{\bar{Q}_B, [Q_B, \phi]\} &= [Q_B, \bar{Q}_B, \phi] \quad (\text{even ordered}), \\ \{Q_B, \{\bar{Q}_B, \phi\}\} + [\bar{Q}_B, \{Q_B, \phi\}] &= [Q_B, \bar{Q}_B, \phi] \quad (\text{odd ordered}), \end{aligned} \quad (19.44)$$

depending on whether  $\phi$  is even-ordered or odd-ordered with respect to  $c$  and  $\bar{c}$ , respectively. If, in addition, the symmetry under the global gauge transformation, i.e., when the gauge function is constant, is not spontaneously broken, the colour charge  $Q^a$  given by

$$J_\mu = -A_\mu \times F_{\mu\nu} - A_\mu \times B + i\bar{\psi}\gamma_\mu t\psi + i(\bar{c} \times D_\mu c) - i(\partial_\mu \bar{c} \times c), \quad (19.45)$$

$$Q^a = \int d^3x J_0^a, \quad (19.46)$$

is conserved. This  $J_\mu$  becomes a source for the gauge field in the sense that

$$\partial_\mu F_{\mu\nu} = -gJ_\nu + i\delta\bar{\delta}A_\nu. \quad (19.47)$$

Moreover,  $Q^a$  satisfies the commutation relations

$$[\psi(x), Q^a] = t^a\psi(x), \dots \quad (19.48)$$

$$[Q^a, Q^b] = if_{abc}Q^c. \quad (19.49)$$

Since none of  $Q_B$ ,  $\bar{Q}_B$ , and  $Q_c$  have colour, they commute with  $Q^a$ , i.e.,

$$[Q^a, Q_B] = [Q^a, \bar{Q}_B] = [Q^a, Q_c] = 0. \quad (19.50)$$



Put another way, the colour charge is a BRS-invariant quantity:

$$\delta Q^a = \bar{\delta} Q^a = 0 . \quad (19.51)$$

## 19.4 BRS Identity and Slavnov–Taylor Identity

For a system including gauge fields, even where the symmetry is spontaneously broken, we can consider that the BRS invariance is not broken, i.e., even if some  $Q^a$  is not conserved, the conservation law for  $Q_B$  may still hold true. An indication of this is the fact that the unitarity of the  $S$ -matrix is based on the BRS invariance, as will be shown later. As a consequence, the following equation holds true for the vacuum  $|\mathbf{0}\rangle$ :

$$Q_B |\mathbf{0}\rangle = 0 . \quad (19.52)$$

Consequently, for arbitrary field operators  $\mathbf{O}_1(x_1), \dots, \mathbf{O}_n(x_n)$ ,

$$\langle \mathbf{0} | \delta T[\mathbf{O}_1(x_1) \dots \mathbf{O}_n(x_n)] | \mathbf{0} \rangle = 0 . \quad (19.53)$$

This is called the *BRS identity*. It will be exemplified in the following. From now on, we shall use the notation

$$\langle \mathbf{0} | T[\mathbf{O}_1(x_1) \dots \mathbf{O}_n(x_n)] | \mathbf{0} \rangle = \langle \mathbf{O}_1(x_1) \dots \mathbf{O}_n(x_n) \rangle .$$

If we have

$$\delta \mathbf{O}_1(x_1) = \dots = \delta \mathbf{O}_n(x_n) = 0 , \quad \mathbf{O}(x) = \delta \mathbf{P}(x) , \quad (19.54)$$

for a set of operators  $\mathbf{O}_1(x_1), \dots, \mathbf{O}_n(x_n)$ , and  $\mathbf{O}(x)$ , then from (19.53),

$$\langle \mathbf{O}(x) \mathbf{O}_1(x_1) \dots \mathbf{O}_n(x_n) \rangle = 0 . \quad (19.55)$$

We thus consider (18.84). In the unrenormalized Landau gauge, since

$$\delta \mathbf{B}(x) = 0 , \quad \mathbf{B}(x) = -i\delta\bar{c}(x) , \quad (19.56)$$

the Green's function including arbitrary numbers of  $\mathbf{B}$  turns out to vanish, so that

$$\left\langle \exp \left[ \frac{i\alpha}{2} \int d^4x \mathbf{B}(x) \dots \mathbf{B}(x) \right] \right\rangle_L = 1 . \quad (19.57)$$

Moreover, if in (18.84) we have

$$\delta \mathbf{A} = \delta \mathbf{B} = \delta \mathbf{C} = \dots = 0, \quad (19.58)$$

then for an unrenormalized Green's function,

$$\langle \mathbf{ABC} \dots \rangle_\alpha = \langle \mathbf{ABC} \dots \rangle_L, \quad (19.59)$$

which means that an unrenormalized Green's function does not depend on the unrenormalized gauge parameter  $\alpha$ . This result will be further generalized.

As in Sect. 18.3, we consider a class of Lagrangian densities with the same gauge-invariant term. We assume that the difference between two arbitrary Lagrangian densities belonging to this class can be written as the BRS transformation of an operator:

$$\Delta \mathcal{L} = \mathcal{L}_{II} - \mathcal{L}_I = \delta \mathcal{M}, \quad (19.60)$$

$$\Delta S = S_{II} - S_I = \delta \int d^4x \mathcal{M}. \quad (19.61)$$

If in (18.83) we now have

$$\langle \exp(i\Delta S) \rangle_I = 1, \quad (19.62)$$

and (18.91) are satisfied once again, then

$$\langle \mathbf{ABC} \dots \rangle_{II} = \langle \mathbf{ABC} \dots \rangle_I, \quad (19.63)$$

for an unrenormalized Green's function. This is important. As will be discussed later, when computing observables such as the  $S$ -matrix using the LSZ reduction formula, we start with the BRS invariant Green's function. Since its normalization is determined by a renormalization condition like (11.122), it turns out that, within the class satisfying the condition (19.61), the  $S$ -matrix is independent of the choice of gauge.

As another application, using

$$-i\delta T[A_\mu^a(x), \bar{c}^b(y)] = T[A_\mu^a(x), B^b(y)] - iT[D_\mu c^a(x), \bar{c}^b(y)], \quad (19.64)$$

we obtain the BRS identity in the form

$$\langle \mathbf{A}_\mu^a(x), \mathbf{B}^b(y) \rangle = i\langle \mathbf{D}_\mu c^a(x), \bar{c}^b(y) \rangle. \quad (19.65)$$

Using the field equation and the canonical commutation relation, we see immediately that the four-divergence of this equation is equal to  $-i\delta_{ab}\delta^4(x-y)$ . Repeating

the discussion in Sect. 15.1,

$$\langle \mathbf{A}_\mu^a(x), \mathbf{B}^b(y) \rangle = -\delta_{ab} \partial_\mu D_F(x-y). \quad (19.66)$$

This is important because this equation assumes the same form in both the renormalized and the unrenormalized formalisms.

A set of relations which hold true among Green's functions can be derived from the BRS identity. This corresponds to the Ward–Takahashi identity in QED, and is called the *Slavnov–Taylor identity* in QCD [171, 172]. We introduce an external field:

$$\begin{aligned} \mathcal{S} = & J_\mu \cdot \mathbf{A}_\mu + \bar{J}_\alpha \cdot \boldsymbol{\psi}_\alpha + J_\alpha \cdot \bar{\boldsymbol{\psi}}_\alpha + \bar{J}_c \cdot \mathbf{c} + J_{\bar{c}} \cdot \bar{\mathbf{c}} + J_B \cdot \mathbf{B} \\ & + K_\mu \cdot \delta \mathbf{A}_\mu + \bar{K}_\alpha \cdot \delta \boldsymbol{\psi}_\alpha + K_\alpha \cdot \delta \bar{\boldsymbol{\psi}}_\alpha + K_c \cdot \delta \mathbf{c}. \end{aligned} \quad (19.67)$$

The fact that  $i\bar{\mathbf{c}}$  is equal to  $i\mathbf{B}$  has already been included in (19.67). If  $\mathcal{S}$  is bosonic, then  $J_\alpha$ ,  $\bar{J}_\alpha$ ,  $K_\alpha$ , and  $\bar{K}_\alpha$  are external fields that anti-commute with the fermionic fields, while  $\bar{J}_c$ ,  $J_{\bar{c}}$ ,  $K_\alpha$ , and  $\bar{K}_\alpha$  are external fields that anti-commute with the ghost fields. These are considered to be anti-commutative external fields. Hence,

$$\exp(-iW) = \langle \mathbf{0} | T \exp \left[ -i \int d^4x \mathcal{S}(x) \right] | \mathbf{0} \rangle. \quad (19.68)$$

The BRS identity in this case is

$$\langle \mathbf{0} | T \left[ \int d^4x \delta \mathcal{S}(x), \exp \left[ -i \int d^4y \mathcal{S}(y) \right] \right] | \mathbf{0} \rangle = 0. \quad (19.69)$$

We introduce the notation

$$\langle \mathcal{O}(x) \rangle \equiv \frac{\langle \mathbf{0} | T \left[ \mathcal{O}(x), \exp \left[ -i \int d^4y \mathcal{S}(y) \right] \right] | \mathbf{0} \rangle}{\langle \mathbf{0} | \exp \left[ -i \int d^4y \mathcal{S}(y) \right] | \mathbf{0} \rangle}. \quad (19.70)$$

Using the nilpotency of the BRS transformation and the fact that  $\bar{J}_c$  and  $J_{\bar{c}}$  anti-commute with  $Q_B$ ,

$$\langle \delta \mathcal{S} \rangle = J_\mu \langle \delta \mathbf{A}_\mu \rangle + \bar{J}_\alpha \langle \delta \boldsymbol{\psi}_\alpha \rangle + J_\alpha \langle \delta \bar{\boldsymbol{\psi}}_\alpha \rangle - \bar{J}_c \langle \delta \mathbf{c} \rangle - J_{\bar{c}} \langle \delta \bar{\mathbf{c}} \rangle = 0. \quad (19.71)$$

Differentiating anti-commuting external fields from the left-hand side,

$$\begin{aligned} \frac{\delta W}{\delta J_\mu} &= \langle \mathbf{A}_\mu \rangle, & \frac{\delta W}{\delta \bar{J}_\alpha} &= \langle \boldsymbol{\psi}_\alpha \rangle, & \frac{\delta W}{\delta J_\alpha} &= \langle \bar{\boldsymbol{\psi}}_\alpha \rangle, \\ \frac{\delta W}{\delta J_B} &= \langle \mathbf{B} \rangle, & \frac{\delta W}{\delta \bar{J}_c} &= \langle \mathbf{c} \rangle, & \frac{\delta W}{\delta J_{\bar{c}}} &= \langle \bar{\mathbf{c}} \rangle. \end{aligned} \quad (19.72)$$

We now introduce the Legendre transformation in order to have  $\langle\varphi\rangle$  and  $K$  as independent variables instead of  $J$  and  $K$  :

$$\Gamma = W - J_\mu \cdot \langle A_\mu \rangle - \bar{J}_\alpha \cdot \langle \psi_\alpha \rangle - J_\alpha \cdot \langle \bar{\psi}_\alpha \rangle - \bar{J}_c \cdot \langle c \rangle - J_{\bar{c}} \cdot \langle \bar{c} \rangle - J_B \cdot \langle B \rangle . \quad (19.73)$$

Differentiating this from the left-hand side,

$$\frac{\delta\Gamma}{\delta\langle\phi\rangle} = \begin{cases} -J_\phi , & \phi = A_\mu , B \\ J_\phi , & \phi = c , \bar{c} , \psi_\alpha , \bar{\psi}_\alpha . \end{cases} \quad (19.74)$$

Further, the derivative with respect to  $K$  is

$$\frac{\delta\Gamma}{\delta K} = \frac{\delta W}{\delta K} . \quad (19.75)$$

From this,

$$\frac{\delta\Gamma}{\delta K_\mu} = \langle \delta A_\mu \rangle , \quad \frac{\delta\Gamma}{\delta \bar{K}_\alpha} = \langle \delta \psi_\alpha \rangle , \quad \frac{\delta\Gamma}{\delta K_\alpha} = \langle \delta \bar{\psi}_\alpha \rangle , \quad \frac{\delta\Gamma}{\delta K_c} = \langle \delta c \rangle , \quad (19.76)$$

while  $\langle \delta \bar{c} \rangle$  is given not by the derivative with respect to  $\Gamma$  but by

$$\langle \delta \bar{c} \rangle = i \langle B \rangle . \quad (19.77)$$

Expressing (19.71) in terms of derivatives with respect to  $\Gamma$ ,

$$\begin{aligned} 0 &= \frac{\delta\Gamma}{\delta\langle A_\mu \rangle} \langle \delta A_\mu \rangle - \frac{\delta\Gamma}{\delta\langle \psi_\alpha \rangle} \langle \delta \psi_\alpha \rangle - \frac{\delta\Gamma}{\delta\langle \bar{\psi}_\alpha \rangle} \langle \delta \bar{\psi}_\alpha \rangle + \frac{\delta\Gamma}{\delta\langle c \rangle} \langle \delta c \rangle + \frac{\delta\Gamma}{\delta\langle \bar{c} \rangle} \langle \delta \bar{c} \rangle \\ &= \langle \delta A_\mu \rangle \frac{\delta\Gamma}{\delta\langle A_\mu \rangle} + \langle \delta \psi_\alpha \rangle \frac{\delta\Gamma}{\delta\langle \psi_\alpha \rangle} + \langle \delta \bar{\psi}_\alpha \rangle \frac{\delta\Gamma}{\delta\langle \bar{\psi}_\alpha \rangle} + \langle \delta c \rangle \frac{\delta\Gamma}{\delta\langle c \rangle} + \langle \delta \bar{c} \rangle \frac{\delta\Gamma}{\delta\langle \bar{c} \rangle} \\ &= \frac{\delta\Gamma}{\delta K_\mu} \frac{\delta\Gamma}{\delta\langle A_\mu \rangle} + \frac{\delta\Gamma}{\delta \bar{K}_\alpha} \frac{\delta\Gamma}{\delta\langle \psi_\alpha \rangle} + \frac{\delta\Gamma}{\delta K_\alpha} \frac{\delta\Gamma}{\delta\langle \bar{\psi}_\alpha \rangle} + \frac{\delta\Gamma}{\delta K_c} \frac{\delta\Gamma}{\delta\langle c \rangle} + i \langle B \rangle \frac{\delta\Gamma}{\delta\langle \bar{c} \rangle} . \end{aligned} \quad (19.78)$$

This is the *Slavnov–Taylor identity* [171, 172].

Furthermore, from the field equations with external fields,

$$-J_B = \partial_\mu \langle A_\mu \rangle - \alpha \langle B \rangle , \quad J_{\bar{c}} = i \partial_\mu \langle D_\mu c \rangle . \quad (19.79)$$

These give the following constraints for  $\Gamma$  with external fields:

$$\frac{\delta\Gamma}{\delta\langle B \rangle} = \partial_\mu \langle A_\mu \rangle - \alpha \langle B \rangle , \quad i \partial_\mu \left( \frac{\delta\Gamma}{\delta K_\mu} \right) = \frac{\delta\Gamma}{\delta\langle \bar{c} \rangle} . \quad (19.80)$$

If  $B$  is not used,  $\langle B \rangle$  should be replaced by

$$\langle B \rangle \rightarrow \frac{1}{\alpha} \partial_\mu \langle A_\mu \rangle . \quad (19.81)$$

This corresponds to setting  $J_B = 0$  at the outset. Differentiating (19.78) with respect to the external fields and then setting the external fields equal to zero, we obtain many identities between the Green's functions.

## 19.5 Representations of the BRS Algebra

To complete the quantization of gauge theories and understand their structure, it is important to know the representation of the BRS algebra. We first investigate the representation of the BRS algebra but excluding  $\tilde{Q}_B$ :

$$i[Q_c, Q_B] = Q_B , \quad Q_B^2 = 0 . \quad (19.82)$$

However, we restrict the representations to those used in gauge theories, rather than considering abstract representations. As mentioned before,  $Q_c$  and  $Q_B$  are both Hermitian operators, and this turns out to be a key fact in obtaining representations.

We first note that all operators in gauge theories can be decomposed into a sum of operators satisfying

$$i[Q_c, \phi] = N\phi . \quad (19.83)$$

The state  $\phi|0\rangle$  satisfies

$$iQ_c\phi|0\rangle = N\phi|0\rangle , \quad N \text{ integer} , \quad (19.84)$$

where we have assumed that the vacuum  $|0\rangle$  is annihilated by  $Q_c$ , i.e.,

$$Q_c|0\rangle = 0 . \quad (19.85)$$

Equation (19.84) indicates that all eigenvalues of the Hermitian operators are pure imaginary. Moreover, the second power of the Hermitian operator  $Q_B$  is zero. All these facts indicate that we must have an indefinite metric. In the following, we will use the result in Chap. 5, and the reader is referred to the discussion there for the above details.

The important point regarding this algebraic representation is to obtain simultaneously the metric matrix  $\eta$  and the representation matrices of the operators  $Q_B$  and  $Q_c$ . Writing the representation matrices of  $Q_B$  and  $iQ_c$  as  $q$  and  $n$ , respectively, the

Hermiticity condition for  $Q_B$  and  $Q_c$  can be written in the form

$$\eta q = q^\dagger \eta, \quad \eta n = -n^\dagger \eta, \quad (19.86)$$

where we have used (5.34). The representation of the BRS algebra satisfies

$$[n, q] = q, \quad q^2 = 0. \quad (19.87)$$

We thus seek a representation which diagonalizes  $n$ . This corresponds to choosing a basis satisfying (19.84) as basis in the state vector space  $\mathcal{V}$ . Thus,

$$n^\dagger = n. \quad (19.88)$$

The diagonal elements of the matrix  $n$  are integers, as already mentioned. According to (19.86),  $n$  anti-commutes with  $\eta$ . Therefore, we decompose  $\mathcal{V}$  in terms of the eigenvalues  $N$  of  $n$ ,

$$\mathcal{V} = \sum_{N=-\infty}^{\infty} \bigoplus \mathcal{V}^{(N)}, \quad (19.89)$$

$$\mathcal{V}^{(N)} = \left\{ |x\rangle / (iQ_c - N)|x\rangle = 0, |x\rangle \in \mathcal{V} \right\}. \quad (19.90)$$

The fact that  $n$  anti-commutes with  $\eta$  means that, multiplying a vector by the matrix  $\eta$ ,  $\mathcal{V}^{(N)}$  turns into  $\mathcal{V}^{(-N)}$ , i.e.,

$$\eta \mathcal{V}^{(N)} = \mathcal{V}^{(-N)}, \quad (19.91)$$

where the manipulation of multiplying a vector by a matrix is defined as follows. If we take  $\{|e_j\rangle\}$  as a set of basis vectors, an arbitrary vector  $|x\rangle$  is expressed as

$$|x\rangle = \sum_j x_j |e_j\rangle. \quad (19.92)$$

Multiplying by  $\eta$  is then defined by

$$\eta |x\rangle = \sum_{j,k} \eta_{jk} x_k |e_j\rangle. \quad (19.93)$$

Moreover, we choose  $\eta$  to be the standard form mentioned in Chap. 5, whence

$$\eta^2 = \mathbf{1}. \quad (19.94)$$

We must now obtain the matrix representing  $q$ . To do this, we first define the BRS singlet and the BRS doublet, following Kugo and Ojima [173], then refine them.

1. **BRS singlet.** If a vector  $|f\rangle$  cannot be written in the form  $|f\rangle = Q_B|g\rangle$  and it satisfies

$$Q_B|f\rangle = 0, \quad (19.95)$$

then  $|f\rangle$  is called a BRS singlet.

2. **BRS doublet.** If there exist vectors  $|f\rangle$  and  $|g\rangle$  satisfying

$$|f\rangle = Q_B|g\rangle \neq 0, \quad \text{and hence,} \quad Q_B|f\rangle = 0, \quad (19.96)$$

we say that  $|f\rangle$  and  $|g\rangle$  form a BRS doublet. Note also that  $|g\rangle$  is called a *parent vector* (parent state) and  $|f\rangle$  is called a *daughter vector* (daughter state).

A weak point of the above definition is the lack of uniqueness. If we consider a singlet state and parent and daughter states of a doublet as  $|s\rangle$ ,  $|p\rangle$ , and  $|d\rangle$ , respectively, then  $|s\rangle + |d\rangle$  also satisfies the first condition (1), while  $|p\rangle + |s\rangle$  also satisfies the condition for the parent in (2). In order to define these uniquely within a given representation, we use the metric matrix  $\eta$ .

We define two subspaces

$$\mathcal{V}_d = q\mathcal{V} = \{q|x\rangle : |x\rangle \in \mathcal{V}\}, \quad (19.97)$$

$$\mathcal{V}_p = q^\dagger\mathcal{V} = \{q^\dagger|x\rangle : |x\rangle \in \mathcal{V}\}. \quad (19.98)$$

Firstly,  $Q_B|x\rangle = q|x\rangle$ , and since  $Q_B$  annihilates the singlet and the daughter state,  $\mathcal{V}_d$  should include every daughter state, so in this sense, it is a set of daughter states. Therefore, considering an arbitrary vector  $|y\rangle \in \mathcal{V}_d$ , it can always be written in the form

$$\mathcal{V}_d \ni |y\rangle = Q_B|x\rangle. \quad (19.99)$$

Setting  $|\tilde{y}\rangle = \eta|y\rangle$  and using the definition of  $\eta$ ,

$$\langle x|Q_B|\tilde{y}\rangle = \langle y|\tilde{y}\rangle = \sum_{j,k,l} y_j^* \eta_{jk} \eta_{kl} y_l = \sum_j |y_j|^2 \neq 0, \quad (19.100)$$

where we have used (19.94). This implies that

$$Q_B|\tilde{y}\rangle = q|\tilde{y}\rangle \neq 0. \quad (19.101)$$

Hence,  $\eta|y\rangle$  is a parent vector. In fact,

$$\eta\mathcal{V}_d = \eta q\mathcal{V} = q^\dagger\eta\mathcal{V} = q^\dagger\mathcal{V} = \mathcal{V}_p, \quad (19.102)$$

where we have used the fact that  $\eta$  is non-degenerate, viz.,

$$\eta \mathcal{V} = \mathcal{V} . \tag{19.103}$$

We thus see that  $\mathcal{V}_p$  is a set of parent states. Alternatively, it can be considered that (19.97) and (19.98) define daughter and parent states, or from (19.102), a parent state is created from a daughter state by multiplying by  $\eta$ .

Furthermore, if  $\eta$  is the standard form, then (16.93) implies

$$\eta \mathcal{V}_p = \mathcal{V}_d . \tag{19.104}$$

Next, we decompose  $\mathcal{V}$  into a direct sum of three subspaces. If  $\mathcal{V}_s$  is the space of singlets,

$$\mathcal{V} = \mathcal{V}_s \oplus \mathcal{V}_p \oplus \mathcal{V}_d . \tag{19.105}$$

Therefore, from (19.102), (19.103), and (19.104),

$$\eta \mathcal{V}_s = \mathcal{V}_s . \tag{19.106}$$

This means that the singlet state is orthogonal to the doublet state. The structure of  $\eta$  is

$$\eta = \begin{array}{ccc|c} & \text{S} & \text{p} & \text{d} \\ \hline & \text{S} & 0 & 0 \\ & \text{p} & 0 & 0 \\ & \text{d} & 0 & 0 \end{array} \tag{19.107}$$

where the shaded regions are non-degenerate matrices. What can be understood from this is that both the parent vectors belonging to  $\mathcal{V}_p$  and the daughter vectors belonging to  $\mathcal{V}_d$  have zero norm. Non-zero inner products appear only between a parent vector and a daughter vector or between singlets.

Now,  $Q_B$  annihilates singlets, so if  $\{0\}$  is the set consisting only of the zero vector,

$$q \mathcal{V}_s = \{0\} . \tag{19.108}$$

Therefore,

$$q^\dagger \mathcal{V}_s = q^\dagger \eta \mathcal{V}_s = \eta q \mathcal{V}_s = \{0\} . \tag{19.109}$$



Thus, singlets turn out to be annihilated by both  $q$  and  $q^\dagger$ . Additionally, since the second power of either  $q$  or  $q^\dagger$  is zero,

$$q\mathcal{V}_d = q^\dagger\mathcal{V}_p = \{0\} . \tag{19.110}$$

Thus,

$$q\mathcal{V}_p = q\mathcal{V}_p \oplus q\mathcal{V}_d \oplus q\mathcal{V}_S = q\mathcal{V} = \mathcal{V}_d , \tag{19.111}$$

$$q^\dagger\mathcal{V}_d = q^\dagger\mathcal{V}_d \oplus q^\dagger\mathcal{V}_p \oplus q^\dagger\mathcal{V}_S = q^\dagger\mathcal{V} = \mathcal{V}_p . \tag{19.112}$$

These relations will be used to determine the matrix  $q$ . Introducing the ghost number  $N$ , we consider the subspace

$$\mathcal{V}_p^{(N)} = \mathcal{V}_p \cap \mathcal{V}^{(N)} . \tag{19.113}$$

From the commutation relation between  $n$  and  $q$ ,

$$q\mathcal{V}_p^{(N)} = \mathcal{V}_d^{(N+1)} . \tag{19.114}$$

For  $\eta$ , we have

$$\eta\mathcal{V}_p^{(N)} = \mathcal{V}_d^{(-N)} , \quad \eta\mathcal{V}_d^{(N+1)} = \mathcal{V}_p^{(-N-1)} . \tag{19.115}$$

Taking into account these relations and the condition (19.86) on  $q$ , in the representation where  $n$  is diagonal,

$$q = \begin{array}{cccc|c} & \begin{matrix} (N+1,d) & (N,p) & (-N,d) & (-N-1,p) \end{matrix} & & & \\ \begin{matrix} (N+1,d) \\ (N,p) \\ (-N,d) \\ (-N-1,p) \end{matrix} & \begin{array}{|c|c|c|c|} \hline 0 & h & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & h^\dagger \\ \hline 0 & 0 & 0 & 0 \\ \hline \end{array} & & & \end{array} \tag{19.116}$$

where  $h$  and  $h^\dagger$  are non-degenerate matrices. If we choose the basis in each subspace properly, then at the end of the day the irreducible representation in the doublet space

has the form

$$\eta = \left( \begin{array}{cc|cc} & & 0 & 1 \\ & & 1 & 0 \\ \hline 0 & 1 & & \\ 1 & 0 & & \end{array} \right), \quad n = \left( \begin{array}{cc|cc} N+1 & 0 & & \\ 0 & N & & \\ \hline & & -N & 0 \\ & & 0 & -N-1 \end{array} \right),$$

$$q = \left( \begin{array}{cc|cc} 0 & a & & \\ 0 & 0 & & \\ \hline & & 0 & a \\ & & 0 & 0 \end{array} \right), \quad (19.117)$$

where  $a$  is a real number which cannot be determined by (19.87) alone, and the blanks are zero matrices. The four states in the basis of this representation form the quartet found by Kugo and Ojima.

What we have discussed above concerns the representation for the BRS doublet. For the singlet, two kinds of representation can be obtained:

- $N \neq 0$

$$n = \begin{pmatrix} N & \\ & -N \end{pmatrix}, \quad \eta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad q = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad (19.118)$$

- $N = 0$

$$n = (0), \quad \eta = (\pm 1), \quad q = (0). \quad (19.119)$$

In order to formulate gauge theories, we introduce the following requirement:

A subspace of the BRS singlet  $\mathcal{V}_S$  belongs to the positive-definite metric.

With this assumption, the BRS singlet representation in a gauge theory is restricted to

$$n = (0), \quad \eta = (1), \quad q = (0). \quad (19.120)$$

## 19.6 Unitarity of the $S$ -Matrix

Since gauge theories have many fields belonging to the indefinite metric, there remains the problem of how to eliminate these from the  $S$ -matrix. We discuss this issue using the method suggested by Kugo and Ojima. Kugo and Ojima defined the

physical state space  $\mathcal{V}_{\text{phys}}$  by [174]

$$\mathcal{V}_{\text{phys}} = \left\{ |x\rangle : Q_{\text{B}}|x\rangle = 0, |x\rangle = 0, |x\rangle \in \mathcal{V} \right\}. \quad (19.121)$$

It is therefore clear from the discussion in the last section that

$$\mathcal{V}_{\text{phys}} = \mathcal{V}_{\text{S}} \oplus \mathcal{V}_{\text{d}}. \quad (19.122)$$

Hence, given the structure of  $\eta$ , if  $|f\rangle$  and  $|g\rangle$  belong to  $\mathcal{V}_{\text{phys}}$ , we have

$$\langle f|g\rangle = \langle f|P(\mathcal{V}_{\text{S}})|g\rangle, \quad (19.123)$$

where  $P(\mathcal{V}_{\text{S}})$  is the projection operator onto the BRS singlet subspace  $\mathcal{V}_{\text{S}}$ .

Since  $Q_{\text{B}}$  is a conserved quantity, it commutes with the  $S$ -matrix, i.e.,

$$[Q_{\text{B}}, S] = 0. \quad (19.124)$$

Hence, if both  $|f\rangle$  and  $|g\rangle$  belong to  $\mathcal{V}_{\text{S}}$ , we have

$$Q_{\text{B}}S|f\rangle = 0, \quad Q_{\text{B}}S|g\rangle = 0. \quad (19.125)$$

Thus,

$$S|f\rangle, S|g\rangle \in \mathcal{V}_{\text{phys}}. \quad (19.126)$$

So, changing the unitarity condition by using (19.123),

$$\langle f|g\rangle = \langle f|S^\dagger S|g\rangle = \langle f|S^\dagger P(\mathcal{V}_{\text{S}})S|g\rangle. \quad (19.127)$$

It thus turns out that the unitarity condition for the  $S$ -matrix is expressed only by states in  $\mathcal{V}_{\text{phys}}$ , which belongs to the positive-definite metric, and the BRS doublets are completely eliminated.

The discussion above can be applied to QED. We use the Lagrangian density in the form (19.1). In this case, since the gauge group is  $U(1)$ , the structure constant  $f$  is zero. The ghost part is

$$\mathcal{L}_{\text{FP}} = i\partial_\mu \bar{c} \partial_\mu c. \quad (19.128)$$

The ghost fields are then free fields:

$$\square c = \square \bar{c} = 0. \quad (19.129)$$

To eliminate the ghost fields, which do not exist in standard QED, we restrict the physical states by

$$c^{(+)}(x)|\text{phys}\rangle = \bar{c}^{(+)}|\text{phys}\rangle = 0 . \quad (19.130)$$

Further, we define them using  $Q_B$ :

$$Q_B|\text{phys}\rangle = 0 . \quad (19.131)$$

This means that consistency between (19.130) and (19.131) is an issue. In QED,

$$\{Q_B, c^{(+)}(x)\} = 0 , \quad \{Q_B, \bar{c}^{(+)}(x)\} = B^{(+)}(x) , \quad (19.132)$$

so it turns out that (19.131) requires the condition

$$B^{(+)}(x)|\text{phys}\rangle = 0 , \quad (19.133)$$

which is nothing but the Lorenz condition introduced in Sect. 5.4. Since  $Q_B$  commutes with  $B^{(+)}(x)$ , it does not give any further condition. Moreover, in this case  $Q_B$  is given by

$$Q_B = \int d^3x (B\dot{c} - \dot{B}c) , \quad (19.134)$$

so if (19.130) and (19.133) are satisfied, then (19.131) is satisfied automatically. Hence, in QED, it turns out that the Kugo–Ojima condition coincides with the Lorenz condition for the states with no ghosts.

## 19.7 Representations of the Extended BRS Algebra

In Sect. 19.5, we investigated representations of the BRS algebra without  $\bar{Q}_B$ . In this section, we add comments on representations of the BRS algebra including  $\bar{Q}_B$ , as defined by (19.41) and (19.42).

We first decompose the state space into the singlet and the doublet:

$$\mathcal{V} = \mathcal{V}_S \oplus \mathcal{V}_D . \quad (19.135)$$

This decomposition is common to both  $Q_B$  and  $\bar{Q}_B$ . The doublet space can be decomposed as

$$\mathcal{V}_D = \mathcal{V}_p \oplus \mathcal{V}_d = \mathcal{V}_{\bar{p}} \oplus \mathcal{V}_{\bar{d}} . \quad (19.136)$$

However, considering  $\bar{q}$  as a representation of  $\bar{Q}_B$ , new subspaces are defined by

$$\mathcal{V}_{\bar{p}} = \bar{q}^\dagger \mathcal{V} = \bar{q}^\dagger \mathcal{V}_D, \quad \mathcal{V}_{\bar{d}} = \bar{q} \mathcal{V} = \bar{q} \mathcal{V}_D. \quad (19.137)$$

For each of  $Q_B$  and  $\bar{Q}_B$ , we introduce decompositions into p and d. A subspace

$$Q_B \bar{Q}_B \mathcal{V} = \bar{Q}_B Q_B \mathcal{V} = \mathcal{V}^{(d,d)} \quad (19.138)$$

is a set of daughter states for both  $Q_B$  and  $\bar{Q}_B$ . Thus,

$$\mathcal{V}^{(d,d)} = \mathcal{V}_d \cap \mathcal{V}_{\bar{d}}. \quad (19.139)$$

Multiplying this subspace by  $\eta$ , it becomes a set of parent states for both  $Q_B$  and  $\bar{Q}_B$ , i.e.,

$$\eta \mathcal{V}^{(d,d)} = \mathcal{V}^{(p,p)} = \mathcal{V}_p \cap \mathcal{V}_{\bar{p}}. \quad (19.140)$$

Starting with this subspace, we introduce the two subspaces

$$Q_B \mathcal{V}^{(p,p)} = \mathcal{V}^{(d,p)}, \quad \bar{Q}_B \mathcal{V}^{(p,p)} = \mathcal{V}^{(p,d)}, \quad (19.141)$$

which are related by

$$\eta \mathcal{V}^{(d,p)} = \mathcal{V}^{(p,d)}. \quad (19.142)$$

Therefore, the doublet space decomposes as follows:

$$\mathcal{V}_D = \mathcal{V}^{(p,p)} \oplus \mathcal{V}^{(p,d)} \oplus \mathcal{V}^{(d,p)} \oplus \mathcal{V}^{(d,d)}. \quad (19.143)$$

In contrast, the singlet space has the following properties:

$$Q_B \mathcal{V}_S = \bar{Q}_B \mathcal{V}_S = \{0\}, \quad \eta \mathcal{V}_S = \mathcal{V}_S. \quad (19.144)$$

The discussion in this section is descriptive and results are given without proof, but we shall give the details in the remainder of this chapter. Note that the decomposition (19.143) shows that there is a one-to-one correspondence between the four subspaces. In the following, we shall express the elements in  $\mathcal{V}_D$  in the form (p,d), using the transformation properties for  $Q_B$  and  $\bar{Q}_B$ . We express relations such as  $Q_B|a\rangle \sim |b\rangle$ ,  $\bar{Q}_B|c\rangle \sim |d\rangle$ , where  $\sim$  stands for equality up to a numerical

coefficient, in the following graphical way:

$$(19.145)$$

Then the graphical representation of the one-to-one correspondence between the elements of the four subspaces in (19.143) is

$$(19.146)$$

These four elements form a quartet.

### 19.8 Representations of BRS Transformations for Auxiliary Fields

We have considered BRS singlets and BRS doublets for state vectors, and this can be extended to operators, i.e., operators can be classified using states obtained by multiplying the vacuum by operators. However, since Heisenberg operators are not irreducible, it is simpler to consider asymptotic fields. As already shown in (19.65) and (19.66), multiplying the vacuum by  $A_\mu$ ,  $B$ ,  $D_\mu c$ , and  $\bar{c}$  produces massless scalar states. Since the operators which create these particles should be included, we consider them as asymptotic fields and write them as follows:

$$A \rightarrow \partial_\mu \chi, \quad B \rightarrow \beta, \quad D_\mu c \rightarrow \partial_\mu \gamma, \quad \bar{c} \rightarrow \bar{\gamma}. \quad (19.147)$$

Similarly, we introduce the following asymptotic fields:

$$\bar{B} \rightarrow \bar{\beta}, \quad D_\mu \bar{c} \rightarrow \partial_\mu \bar{\Gamma}, \quad c \rightarrow \Gamma. \quad (19.148)$$

It is easily checked in the Landau gauge that these create massless particles. Therefore, from the BRS transformations for Heisenberg operators,

$$\delta\chi = \gamma, \quad \delta\bar{\gamma} = i\beta, \quad \delta\gamma = 0, \quad \delta\beta = 0, \tag{19.149}$$

$$\bar{\delta}\chi = \bar{\Gamma}, \quad \bar{\delta}\Gamma = i\bar{\beta}, \quad \bar{\delta}\bar{\Gamma} = 0, \quad \bar{\delta}\bar{\beta} = 0. \tag{19.150}$$

From these asymptotic fields, choosing the Landau gauge, we obtain

$$\begin{aligned} \langle \chi^a(x), \beta^b(y) \rangle &= -\langle \chi^a(x), \bar{\beta}^b(y) \rangle = i\langle \gamma^a(x), \bar{\gamma}^b(y) \rangle = i\langle \Gamma^a(x), \bar{\Gamma}^b(y) \rangle \\ &= -\delta_{ab}D_F(x - y). \end{aligned} \tag{19.151}$$

In the following, we consider two cases:

1. **Quartet representation.** In perturbation theories, the following conditions hold true:

$$\gamma = \Gamma, \quad \bar{\Gamma} = \bar{\gamma}. \tag{19.152}$$

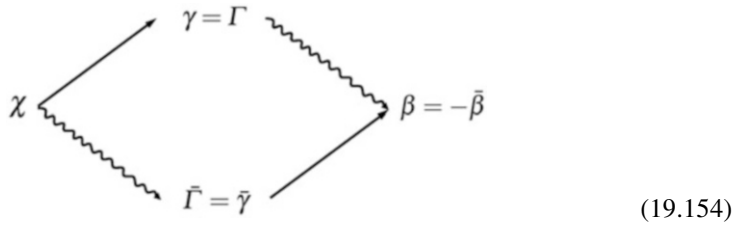
In this case,

$$\bar{\delta}\delta\chi = \bar{\delta}\gamma = \bar{\delta}\Gamma = i\bar{\beta}, \quad \delta\bar{\delta}\chi = \delta\bar{\Gamma} = \delta\bar{\gamma} = i\beta.$$

Thus, taking into account the anti-commutativity of  $\delta$  and  $\bar{\delta}$ ,

$$\bar{\beta} = -\beta. \tag{19.153}$$

In this case, the auxiliary fields form a quartet expressed by the following graph:



This quartet was introduced by Kugo and Ojima.

2. **Octet representation.** Next, we consider what happens if we take  $\gamma$  and  $\bar{\Gamma}$  to be different from  $\Gamma$  and  $\bar{\gamma}$ :

$$\gamma \neq \Gamma, \quad \bar{\Gamma} \neq \bar{\gamma}. \tag{19.155}$$

In this case, if we include  $\beta$ , and taking into account that there are at least five auxiliary fields, it will be impossible to accommodate all the auxiliary fields

unless we introduce at least two quartets. In this case too, we have to write down equations for the BRS transformations twice, as in (19.149) and (19.150). Hence, in addition to the four asymptotic fields in (19.155), we introduce two new asymptotic fields  $\beta, \bar{\beta}$  and  $d, \bar{d}$ :

$$\begin{aligned} \delta\Gamma &= id, & \bar{\delta}\Gamma &= i\beta, & \delta\bar{\beta} &= -\gamma, & \bar{\delta}d &= \gamma, \\ \delta\bar{\gamma} &= i\beta, & \bar{\delta}\bar{\gamma} &= -i\bar{d}, & \delta\bar{d} &= \Gamma, & \bar{\delta}\beta &= \Gamma. \end{aligned} \tag{19.156}$$

These equations include (19.149) and (19.150). In addition, although  $\chi$  has vanished, if we interpret

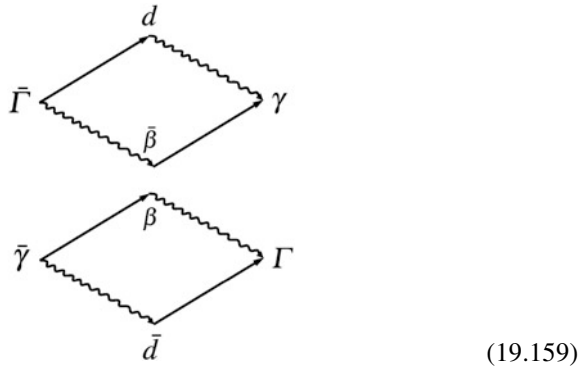
$$\chi = \beta - \bar{\beta} + (\text{singlet}), \tag{19.157}$$

then it does not after all contradict (19.149) and (19.150). Now,  $d$  and  $\bar{d}$  become massless asymptotic fields for  $c \times c$  and  $\bar{c} \times \bar{c}$ , as will be clear from the discussion in the next section.

Additionally, for  $d$  and  $\bar{d}$ , it can be shown immediately that

$$\langle \bar{d}^a(x), d^b(y) \rangle = \delta_{ab} D_F(x - y). \tag{19.158}$$

We use this to make a graphical representation for an octet corresponding to (19.156):



This octet was considered by Kugo.

Now that representations like the quartet and the octet have been obtained, we can investigate what kinds of representations exist more generally. We will discuss this in the next section.



## 19.9 Representations of BRSNO Algebras

We have discussed the fact that, in general gauge theories, an extended BRS invariance holds true and the conserved quantities form the BRS algebras. It was pointed out by Nakanishi and Ojima in [175] that two more conserved charges exist in the Landau gauge. In this gauge, the sum of the gauge-fixing term and the ghost term is

$$\mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{FP}} = \partial_\mu B A_\mu + i\partial_\mu \bar{c} D_\mu c . \quad (19.160)$$

This Lagrangian density is invariant under the following transformations among auxiliary fields:

$$\mathbf{B} \rightarrow \mathbf{B} - ig(\mathbf{c} \times \mathbf{c})_\lambda , \quad \bar{\mathbf{c}} \rightarrow \bar{\mathbf{c}} - 2c\lambda , \quad \mathbf{c} \rightarrow \mathbf{c} , \quad (19.9.2a)$$

where  $\lambda$  is a transformation parameter. Another transformation is

$$\mathbf{B} \rightarrow \mathbf{B} + ig(\bar{\mathbf{c}} \times \bar{\mathbf{c}})_\lambda , \quad \mathbf{c} \rightarrow \mathbf{c} + 2\bar{c}\lambda , \quad \bar{\mathbf{c}} \rightarrow \bar{\mathbf{c}} . \quad (19.9.2b)$$

Writing the Hermitian generators of these transformations as  $Q(c, c)$  and  $Q(\bar{c}, \bar{c})$ , respectively,

$$i[Q(c, c), \mathbf{B}] = -ig\mathbf{c} \times \mathbf{c} , \quad i[Q(c, c), \bar{\mathbf{c}}] = -2c , \quad i[Q(c, c), \mathbf{c}] = 0 , \quad (19.9.3a)$$

$$i[Q(\bar{c}, \bar{c}), \mathbf{B}] = ig\bar{\mathbf{c}} \times \bar{\mathbf{c}} , \quad i[Q(\bar{c}, \bar{c}), \mathbf{c}] = 2\bar{c} , \quad i[Q(\bar{c}, \bar{c}), \bar{\mathbf{c}}] = 0 . \quad (19.9.3b)$$

In contrast, the fundamental fields are invariant under these transformations. Hence, adding  $Q(c, c)$  and  $Q(\bar{c}, \bar{c})$  to the extended BRS algebra (19.41) and (19.42), we obtain the following relations:

$$\begin{aligned} i[Q_c, Q(c, c)] &= 2Q(c, c) , & i[Q_c, Q(\bar{c}, \bar{c})] &= -2Q(\bar{c}, \bar{c}) , \\ [Q(c, c), \bar{Q}_B] &= 2iQ_B , & [Q(\bar{c}, \bar{c}), Q_B] &= -2iQ_B , \\ [Q(c, c), Q_B] &= 0 , & [Q(\bar{c}, \bar{c}), \bar{Q}_B] &= 0 , \\ [Q(c, c), Q(\bar{c}, \bar{c})] &= 4iQ_c . \end{aligned} \quad (19.9.4)$$

Although it is possible to give explicit forms of  $Q(c, c)$  and  $Q(\bar{c}, \bar{c})$ , we omit them here because such expressions will not be necessary. In the Landau gauge, both  $Q(c, c)$  and  $Q(\bar{c}, \bar{c})$  can be conserved quantities, but in other gauges they are not conserved. However, the commutation relations given above still hold true. The algebra that extends the extended BRS algebra by including (19.9.4) is called the *BRSNO algebra*, which we denote by  $\mathcal{A}$ . Moreover, excluding the anti-commuting

operators  $Q_B$  and  $\bar{Q}_B$ , the subspace consisting only of  $Q_c$ ,  $Q(c, c)$ , and  $Q(\bar{c}, \bar{c})$  is denoted by  $\mathcal{B}$ . This can be considered as the bosonic part of the algebra  $\mathcal{A}$ .

To obtain representations of the algebra  $\mathcal{A}$ , we introduce the following notation:

$$\begin{aligned} iQ_c &\rightarrow 2J_x, & Q(c, c) &\rightarrow 2J_+, & Q(\bar{c}, \bar{c}) &\rightarrow 2J_-, \\ iQ_B &\rightarrow \alpha, & \bar{Q}_B &\rightarrow \beta. \end{aligned} \quad (19.9.5)$$

In addition, in the following discussion, we restrict the representations of the algebra  $\mathcal{A}$  to those which can be realized in gauge theories, rather than considering completely abstract representations. First, we obtain a representation of the Lie algebra  $\mathcal{B}$ . We write down the commutation relations which characterize  $\mathcal{B}$  using the symbols above:

$$[J_z, J_+] = J_+, \quad [J_z, J_-] = -J_-, \quad [J_+, J_-] = 2J_z. \quad (19.9.6)$$

These coincide with the commutation relations among the three components of the angular momentum. Hence, in gauge theories and in the Landau gauge, we will prove that the representation of  $\mathcal{B}$  coincides with that of the angular momentum. To do so, we begin with the following requirements:

1. The vacuum state  $|\mathbf{0}\rangle$  must satisfy the condition

$$J_+|\mathbf{0}\rangle = J_-|\mathbf{0}\rangle = J_z|\mathbf{0}\rangle. \quad (19.9.7)$$

2. The vacuum state must be cyclic with respect to the algebra formed by fundamental fields and auxiliary fields, i.e., combining states constructed by multiplying the vacuum by products of field operators, we should be able to generate the entire state vector space  $\mathcal{V}$ .

Since the fundamental fields and  $B - \bar{B}$  commute with the three elements of  $\mathcal{B}$ , these belong to the one-dimensional representation of  $\beta$ . Therefore, to make other representations, we must use  $c$ ,  $\bar{c}$ , and  $B + \bar{B}$  or  $c \times \bar{c}$ . We have the following commutation relations:

$$\begin{aligned} [J_z, c] &= \frac{1}{2}c, & [J_z, \bar{c}] &= -\frac{1}{2}\bar{c}, \\ i[J_+, \bar{c}] &= -c, & i[J_+, c] &= 0, \\ i[J_-, c] &= \bar{c}, & i[J_-, \bar{c}] &= 0. \end{aligned} \quad (19.9.8)$$

We then define a doublet and a triplet by

$$\phi(1/2, 1/2) = i\mathbf{c}, \quad \phi(1/2, -1/2) = \bar{\mathbf{c}}, \quad (19.9.9)$$

$$\phi(1, 1) = \frac{i}{\sqrt{2}}(\mathbf{c} \times \mathbf{c}), \quad \phi(1, 0) = \mathbf{c} \times \bar{\mathbf{c}}, \quad \phi(1, -1) = -\frac{i}{\sqrt{2}}(\bar{\mathbf{c}} \times \bar{\mathbf{c}}). \quad (19.9.10)$$

We immediately obtain

$$\begin{aligned} [J_z, \phi(j, m)] &= m\phi(j, m), \\ [J_+, \phi(j, m)] &= \sqrt{(j-m)(j+m+1)}\phi(j, m+1), \\ [J_-, \phi(j, m)] &= \sqrt{(j+m)(j-m+1)}\phi(j, m-1). \end{aligned} \quad (19.9.11)$$

Since these relations are the same as those for angular momentum, starting with the doublet and the triplet and using the Clebsch–Gordan coefficients for the angular momentum, we can make a product representation:

$$\phi(J, M, c) = \sum_{m_1+m_2=M} \phi(j_1, m_1, a)\phi(j_2, m_2, b) \langle j_1, j_2, m_1, m_2 | j_1, j_2, J, M \rangle, \quad (19.9.12)$$

where  $a$ ,  $b$ , and  $c$  represent other quantum numbers. It is clear that such a product representation again satisfies (19.9.11). Hence, if we take

$$\mathbf{J}^2 = \frac{1}{2}(J_+J_- + J_-J_+) + J_z^2, \quad (19.9.13)$$

then clearly, from (19.9.7) and (19.9.11), for the polynomial of field operators  $\phi(j, m, a)$ ,

$$\mathbf{J}^2\phi(j, m, a)|\mathbf{0}\rangle = j(j+1)\phi(j, m, a)|\mathbf{0}\rangle, \quad J_z\phi(j, m, a)|\mathbf{0}\rangle = m\phi(j, m, a)|\mathbf{0}\rangle. \quad (19.9.14)$$

We thus see that the representation of  $\mathcal{B}$  is exactly the same as that of the angular momentum. Next, we introduce  $\alpha$  and  $\beta$ , with commutation relations

$$[J_z, \alpha] = \frac{1}{2}\alpha, \quad [J_z, \beta] = -\frac{1}{2}\beta, \quad [J_+, \beta] = \alpha, \quad [J_-, \alpha] = \beta. \quad (19.9.15)$$

We consider the product  $\alpha\beta$ :

$$[J_+, \alpha\beta] = [J_-, \alpha\beta] = [J_z, \alpha\beta] = 0. \quad (19.9.16)$$

The subspace  $\mathcal{V}^{(d,d)}$  can now be written as  $\alpha\beta\mathcal{V}$ , and from (19.9.16),

$$\mathbf{J}\mathcal{V}^{(d,d)} \subset \mathcal{V}^{(d,d)}. \quad (19.9.17)$$

This means that irreducible representations of  $\mathcal{B}$  can be constructed within  $\mathcal{V}^{(d,d)}$ . Next, introducing the metric matrix  $\eta$  and expressing the representation matrices of  $J_+$ ,  $J_-$  and  $J_z$  using the same letters, we find that

$$\eta J_+ \eta = J_+^\dagger, \quad \eta J_- \eta = J_-^\dagger, \quad \eta J_z \eta = -J_z^\dagger. \quad (19.9.18)$$

We have chosen  $\eta$  to have the standard form, so

$$\eta\beta\eta = \beta^\dagger, \quad \eta\alpha\eta = -\alpha^\dagger. \quad (19.9.19)$$

In the angular momentum representation, we have

$$J_z^\dagger = J_z, \quad J_+^\dagger = J_-, \quad J_-^\dagger = J_+, \quad (19.9.20)$$

so, combining (19.9.18) with (19.9.20), we find

$$\eta J_z \eta = -J_z, \quad \eta J_+ \eta = J_-, \quad \eta J_- \eta = J_+. \quad (19.9.21)$$

Thus,

$$\eta \mathbf{J}^2 \eta = \mathbf{J}^2, \quad (19.9.22)$$

so we see that the metric matrix connects states having the same value of the quantum number  $j$ . It has been shown in (19.140) that  $\eta$  sends  $\mathcal{V}^{(d,d)}$  to  $\mathcal{V}^{(p,p)}$ , so from (19.9.17),

$$\mathbf{J}\mathcal{V}^{(p,p)} \subset \mathcal{V}^{(p,p)}. \quad (19.9.23)$$

Therefore, it turns out that irreducible representations of  $\mathcal{B}$  can also be constructed within  $\mathcal{V}^{(p,p)}$ . We thus construct within  $\mathcal{V}^{(p,p)}$  an irreducible representation of  $\mathcal{B}$  corresponding to one value of  $j$ , writing its basis as

$$|j, m, a\rangle \in \mathcal{V}^{(p,p)}, \quad m = -j, \dots, j. \quad (19.9.24)$$

Therefore,

$$\alpha|j, m, a\rangle \in \mathcal{V}^{(d,p)}, \quad \beta|j, m, a\rangle \in \mathcal{V}^{(p,d)}. \quad (18.9.25)$$

Moreover,

$$\alpha\beta|j, m, a\rangle \in \mathcal{V}^{(d,d)}. \quad (19.9.26)$$

We thus obtain  $4(2j + 1)$  basis vectors for a representation of the algebra  $\mathcal{A}$ . This set of basis vectors is closed under operations by elements of  $\mathcal{A}$ , i.e.,

$$|j, m, a\rangle, \quad \alpha|j, m, a\rangle, \quad \beta|j, m, a\rangle, \quad \alpha\beta|j, m, a\rangle. \quad (19.9.27)$$

If we fix  $m$  in the above set, they become a BRS quartet. In addition, if we change  $m$ ,  $\{|j, m, a\rangle\}$  and  $\{\alpha\beta|j, m, a\rangle\}$  belong to a representation of the three-dimensional rotation group  $D_j$ . On the other hand,  $\{\alpha|j, m, a\rangle\}$  and  $\{\beta|j, m, a\rangle\}$  become linear combinations of  $D_{j+1/2}$  and  $D_{j-1/2}$ :

$$\begin{aligned} \left|j + \frac{1}{2}, m + \frac{1}{2}, a\right\rangle &= \sqrt{\frac{j+m+1}{2j+1}}\alpha|j, m, a\rangle + \sqrt{\frac{j-m}{2j+1}}\beta|j, m, a\rangle, \\ \left|j - \frac{1}{2}, m - \frac{1}{2}, a\right\rangle &= \sqrt{\frac{j-m}{2j+1}}|j, m, a\rangle + \sqrt{\frac{j+m+1}{2j+1}}\beta|j, m, a\rangle. \end{aligned} \quad (19.9.28)$$

We write the resulting representation of  $\mathcal{A}$  as  $\mathcal{D}_j$ . All representations of  $\mathcal{A}$  except for singlets turn out to be designated by one quantum number  $j$ . We can express the representation  $\mathcal{D}_j$  graphically (see Fig. 19.1). In this case, since  $J_z = iQ_c/2$ ,

$$m = \frac{1}{2}N, \quad (19.9.29)$$

whence  $m$  is one-half of the ghost number.

We now see that the quartet and the octet introduced in the last section belong to the representations  $\mathcal{D}_0$  and  $\mathcal{D}_1$ , respectively. In particular, considering auxiliary

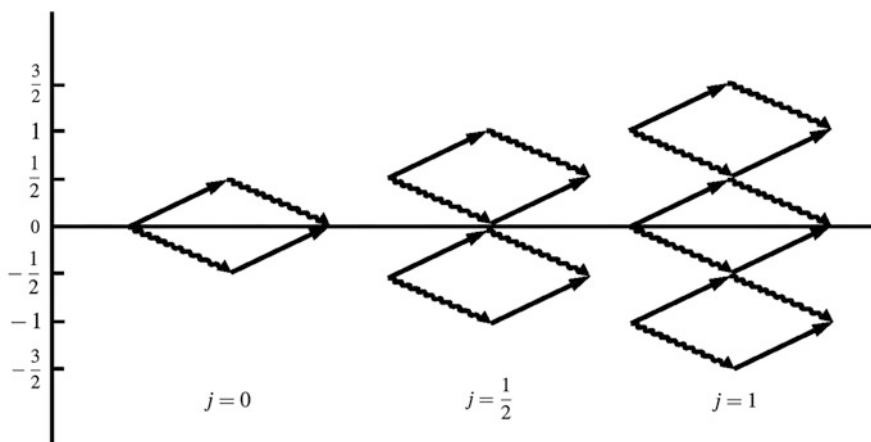


Fig. 19.1 Irreducible representations  $\mathcal{D}_j$  of the algebra  $\mathcal{A}$

fields, depending on whether  $j$  is an integer or a half integer, we find the following difference:

$$j = \text{integer} \quad \beta + \bar{\beta} = 0, \quad (19.9.30)$$

$$j = \text{half integer} \quad \beta + \bar{\beta} \neq 0. \quad (19.9.31)$$

In the latter case,  $c \times c$  and  $\bar{c} \times \bar{c}$  turn out to possess massless asymptotic fields from (19.9.3a) and (19.9.3b). Similarly, from (19.39),  $c \times \bar{c}$  possesses a massless asymptotic field as well. From (19.9.9), it belongs to the triplet of  $\mathcal{B}$ . According to the discussion in Chap. 21, colour confinement occurs when  $j$  is a half integer.

We should also determine the representation of the matrix  $\eta$  of the indefinite metric. Note that the representation  $\mathcal{D}_j$  is not necessarily closed under  $\eta$ . One representation  $\mathcal{D}_j(a)$  may sometimes be transferred to another representation  $\mathcal{D}_j(b)$ . In this case, we need the following two irreducible representations:

$$\eta \mathcal{D}_j(a) = \mathcal{D}_j(b), \quad \eta \mathcal{D}_j(b) = \mathcal{D}_j(a). \quad (19.9.32)$$

We define a phase by

$$\eta |j, m, a\rangle = \alpha \beta |j, -m, b\rangle. \quad (19.9.33)$$

The following relations are then obtained automatically:

$$\begin{aligned} \eta(\alpha |j, m, a\rangle) &= -\beta |j, -m, b\rangle, \\ \eta(\beta |j, m, a\rangle) &= -\alpha |j, -m, b\rangle, \\ \eta(\alpha \beta |j, m, a\rangle) &= |j, -m, b\rangle. \end{aligned} \quad (19.9.34)$$

In this way, we can show that (19.9.20) is satisfied.

For the irreducible representation  $\mathcal{D}_j$ , going back to the notation in Sect. 19.5 and writing

$$\alpha = iq, \quad \beta = \bar{q}, \quad (19.9.35)$$

it can be shown that

$$\begin{aligned} q^2 = \bar{q}^2 = (q^\dagger)^2 = (\bar{q}^\dagger)^2 &= 0, \\ \{q, \bar{q}\} = \{q^\dagger, \bar{q}^\dagger\} = \{q, \bar{q}^\dagger\} = \{q^\dagger, \bar{q}\} &= 0. \end{aligned} \quad (18.9.36)$$

If we choose (19.9.27) as a basis, then for this normalization, we have

$$\Delta = qq^\dagger + q^\dagger q = \bar{q}\bar{q}^\dagger + \bar{q}^\dagger \bar{q} = 1. \quad (19.9.37)$$

This  $\Delta$  is in fact a Casimir operator commuting with all elements in  $\mathcal{L}$ . However, since  $\Delta$  vanishes for singlets,

$$\Delta|f\rangle = \begin{cases} |f\rangle, & |f\rangle \in \mathcal{V}_D, \\ 0, & |f\rangle \in \mathcal{V}_S. \end{cases} \quad (19.9.38)$$

Moreover, it is easily shown that  $\Delta$  also commutes with  $\eta$ .

# Chapter 20

## Renormalization Group



The simple term “renormalization group” has a variety of meanings. This is because various kinds of transformation go by this name. Moreover, it is used in slightly different ways in several different areas of physics. Indeed, different kinds of quantity are renormalized under the renormalization transformations: for instance, free energies in condensed matter physics and Green’s functions in elementary particle physics. If we ask what feature of the renormalization group is the greatest common factor, then we can say that it is a dilation transformation for lengths, momenta, and so on. Note also that the renormalization group really is a group in the mathematical sense. The qualifier “renormalization” refers to the fact that this group is related to manipulations, or renormalizations. In elementary particle theories, two main kinds of renormalization group are considered.

The first kind concerns transformations occurring within one given world. In order to solve the same equations of motion, we define a fundamental quantity, say a Green’s function, under different boundary conditions. The boundary conditions are distinguished by certain parameters. We assume that every Green’s function essentially includes information about solutions to the equations of motion, and any observable can be expressed using any Green’s function. Therefore, observables turn out to be independent of the parameters distinguishing boundary conditions. They are invariant quantities under variations of such parameters. On the other hand, Green’s functions with different values of the parameters are related to one another. We can investigate correspondences between Green’s functions under variations of the parameters. The set of these parameter transformations forms the first kind of renormalization group. This group of transformations is closed in any given world. Observables are invariant quantities under the renormalization group, but quantities which vary under these transformations are not dynamical parameters, so this does not correspond to any conservation law.

The second kind of renormalization group differs from the first kind because it is not restricted to one given world. An example in classical mechanics is Reynolds’ similarity law in fluid dynamics. Even if a viscosity, a density, and other quantities



which characterize a fluid are different, if the Reynolds' number is the same, and if we know the behaviour of one fluid, even a different kind of fluid, then using a suitable dictionary, the behaviours of other fluids can be understood as well. In the realm of elementary particles, if we consider changing masses or coupling constants, this is not restricted to one specific world.

## 20.1 Renormalization Group for QED

A renormalization group of the first kind described above was introduced for QED, first by Stueckelberg and Petermann in [176], and then by Gell-Mann and Low in [177]. There is a pedagogical description in Chap. 8 of the textbook by Bogoliubov and Shirkov, entitled *Introduction to the Theory of Quantized Fields* [178].

In the following, we label unrenormalized field operators and physical quantities with a subscript 0 to distinguish them from renormalized quantities. The Lagrangian density for QED is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^{(0)}F_{\mu\nu}^{(0)} - \frac{1}{2\alpha_0}(\partial_\mu A_\mu^{(0)})^2 - \bar{\psi}^{(0)}(\gamma_\mu D_\mu + m_0)\psi^{(0)}. \quad (20.1)$$

The renormalization is defined by the following equations:

$$A_\mu^{(0)}(x) = Z_3^{1/2}A_\mu(x), \quad \psi^{(0)}(x) = Z_2^{1/2}\psi(x). \quad (20.2)$$

The renormalized parameters are

$$e_0^2 = Z_3^{-1}e^2, \quad \alpha_0 = Z_3\alpha. \quad (20.3)$$

Here we have used Ward's identity, viz.,

$$Z_1 = Z_2. \quad (20.4)$$

Labeling quantities quantized on the mass shell with a subscript C, we introduce the Green's function

$$\langle \mathbf{0} | T[A_\mu(x)A_\nu(y)] | \mathbf{0} \rangle = \frac{-i}{(2\pi)^4} \int d^4k e^{ik \cdot (x-y)} D_{\text{FC}}(k)_{\mu\nu}, \quad (20.5)$$

$$D_{\text{FC}}(k)_{\mu\nu} = \frac{1}{k^2 - i\epsilon} \left[ \left( \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2 - i\epsilon} d_C(k^2) + \alpha \frac{k_\mu k_\nu}{k^2 - i\epsilon} \right) \right]. \quad (20.6)$$

where  $d_C$  has the form

$$d_C(k^2) = 1 + k^2 \int d\kappa^2 \frac{C(\kappa^2)}{k^2 + \kappa^2 - i\epsilon}. \quad (20.7)$$

From (12.249), to lowest order,  $C(\kappa^2)$  has the form

$$C(\kappa^2) = \frac{e^2}{12\pi^2} \frac{1}{\kappa^2} \left( 1 + \frac{2m^2}{\kappa^2} \right) \sqrt{1 - \frac{4m^2}{\kappa^2}} \theta(\kappa^2 - 4m^2). \quad (20.8)$$

When we compute a Green's function, we express contributions corresponding to the Feynman diagrams in terms of combinations of propagators, vertex functions, and coupling constants. When we modify the normalizations of  $S_F$ ,  $D_F$ ,  $\Gamma_\mu$ , and so on, this leads to variations in the coupling constants. However, regarding the Green's function as a whole, these variations cannot be absorbed completely, and the overall normalization turns out to be changed. A special case is the renormalization

$$S'_F(p, e_0^2) = Z_2 S_{FC}(p, e^2), \quad (20.9)$$

$$D'_F(k, e_0^2) = Z_3 D_{FC}(k, e^2), \quad (20.10)$$

$$\Gamma'_\mu(p, q, k, e_0^2) = Z_1^{-1} \Gamma_{\mu C}(p, q, k, e^2), \quad (20.11)$$

where  $S_{FC}$  and  $D_{FC}$  have been normalized at  $p \cdot \gamma = im$  and  $k^2 = 0$ , respectively. We thus take the normalization points to be those for space-like momenta:

$$p \cdot \gamma = \lambda', \quad k^2 = \lambda'^2. \quad (20.12)$$

That is, introducing  $s(\lambda, \lambda', p)$  and  $d(\lambda, \lambda', k)$  by

$$(ip \cdot \gamma + m) S_F(\lambda, \lambda', p) \equiv s(\lambda, \lambda', p), \quad (20.13)$$

$$(\delta_{\mu\nu} k^2 - k_\mu k_\nu) D_F(\lambda, \lambda', k)_{\nu\sigma} \equiv \left( \delta_{\mu\sigma} - \frac{k_\mu k_\nu}{k^2} \right) d(\lambda, \lambda', k), \quad (20.14)$$

we assume that the following normalization has been taken for (20.12):

$$s(\lambda, \lambda', p) = 1, \quad d(\lambda, \lambda', k) = 1. \quad (20.15)$$

$S_F$  and  $D_F$  introduced in (20.13) and (20.14) differ from  $S_{FC}$  and  $D_{FC}$  only by normalizations. Using (20.9) and (20.10), we can now write the equations connecting the unrenormalized Green's functions with the above Green's functions:

$$S'_F(p, e_0^2) = Z_2(\lambda, \lambda'; e_\lambda^2) S_F(\lambda, \lambda'; p, e_\lambda^2), \quad (20.16)$$

$$D'_F(p, e_0^2) = Z_3(\lambda, \lambda'; k, e_\lambda^2), \quad (20.17)$$

$$e_\lambda^2 = Z_3(\lambda, \lambda'; e_\lambda^2) e_0^2, \quad (20.18)$$

where  $e_\lambda^2$  is the parameter for perturbative expansion. The mass shell is specified by

$$\lambda' = im, \quad \lambda = 0. \quad (20.19)$$

Setting  $\lambda' = im$  in (20.16) and  $\lambda = 0$  in (20.17),

$$S'_F(p, e_0^2) = Z_2(\lambda, im, e_\lambda^2) S_{FC}(p, e^2), \quad (20.20)$$

$$D'_F(k, e_0^2) = Z_3(0, \lambda', e^2) D_{FC}(k, e^2). \quad (20.21)$$

We then introduce the ratios

$$z_2(\lambda, \lambda', e^2) = Z_2(\lambda, \lambda'; e_\lambda^2) / Z_2(\lambda, im; e_\lambda^2), \quad (20.22)$$

$$z_3(\lambda, \lambda', e^2) = Z_3(\lambda, \lambda'; e_\lambda^2) / Z_3(0, \lambda'; e^2). \quad (20.23)$$

Therefore, for propagators with pole factors removed,

$$z_2(\lambda, \lambda'; e^2) s(\lambda, \lambda', p, e_\lambda^2) = s_C(p, e^2), \quad (20.24)$$

$$z_3(\lambda, \lambda'; e^2) d(\lambda, \lambda', k, e_\lambda^2) = d_C(k, e^2), \quad (20.25)$$

where  $z_2$  and  $z_3$  are finite factors. Now, from (20.15),

$$z_2(\lambda, \lambda'; e^2) = s_C(\lambda', e^2), \quad z_3(\lambda, \lambda'; e^2) = d_C(\lambda, e^2). \quad (20.26)$$

and

$$e_\lambda^2 = e^2 d_C(\lambda, e^2). \quad (20.27)$$

All the basic renormalization equations are expressed by (20.24)–(20.27).

## 20.2 Approximate Equations for the Renormalization Group

In the renormalization group equations introduced in the previous section, we can obtain  $s$  and  $d$  when  $m^2$  is assumed to be small.

### 20.2.1 Approximation Neglecting Vacuum Polarization

First, an approximation neglecting vacuum polarization is

$$d_C = 1. \quad (20.28)$$

In this case, from (20.24), (20.26), and (20.28), the equation for  $s$  is

$$s(\lambda', p, e^2) = \frac{s_C(p, e^2)}{s_C(\lambda', e^2)}. \quad (20.29)$$

We thus set

$$s_C(p, e^2) = s_C\left(\frac{p^2}{m^2}, e^2\right), \quad s_C(\lambda', e^2) = s_C\left(\frac{\lambda'^2}{m^2}, e^2\right). \quad (20.30)$$

If we assume that  $s(\lambda', p, e^2)$  does not depend on  $m^2$ ,

$$s(\lambda', p, e^2) = s\left(\frac{p^2}{\lambda'^2}, e^2\right). \quad (20.31)$$

This becomes an approximation when  $\lambda'$  and  $p$  are far larger than  $m$  :

$$s\left(\frac{p^2}{\lambda'^2}, e^2\right) = \frac{s_C\left(\frac{p^2}{m^2}, e^2\right)}{s_C\left(\frac{\lambda'^2}{m^2}, e^2\right)}. \quad (20.32)$$

Solutions to this are

$$\begin{aligned} s_C\left(\frac{p^2}{m^2}, e^2\right) &= C(e^2) \left(\frac{p^2}{m^2}\right)^{\gamma(e^2)}, \\ s_C\left(\frac{\lambda'^2}{m^2}, e^2\right) &= C(e^2) \left(\frac{\lambda'^2}{m^2}\right)^{\gamma(e^2)}, \end{aligned} \quad (20.33)$$

where, in perturbation theory,  $C(e^2)$  includes infrared divergences. However, this factor is eliminated from

$$s\left(\frac{p^2}{\lambda'^2}, e^2\right) = \left(\frac{p^2}{\lambda'^2}\right)^{\gamma(e^2)}. \quad (20.34)$$

In fact, the exponent  $\gamma(e^2)$  depends on how we choose the gauge. Therefore, even if the mass is zero, if we normalize at a point off the mass shell, then infrared divergences do not appear. In addition, the reader is referred to (20.102).

### 20.2.2 Approximation Taking into Account Vacuum Polarization

Here we also assume that  $s$  does not depend on  $m^2$ . This is another high-energy approximation. In this case, we write  $s$  in the form

$$s(\lambda, \lambda', p, e_\lambda^2) \approx s\left(\frac{p^2}{\lambda^2}, \frac{p^2}{\lambda'^2}, e_\lambda^2\right). \quad (20.35)$$

Therefore, the renormalization group equation under this approximation is

$$s\left(\frac{p^2}{\lambda^2}, \frac{p^2}{\lambda'^2}, e_\lambda^2\right) = \frac{s_C\left(\frac{p^2}{m^2}, e^2\right)}{s_C\left(\frac{\lambda'^2}{m^2}, e^2\right)}. \quad (20.36)$$

Moreover, assuming that  $d$  does not depend on  $m^2$ , we introduce the approximation

$$d(\lambda, k, e_\lambda^2) = d\left(\frac{k^2}{\lambda^2}, \frac{m^2}{\lambda^2}, e_\lambda^2\right) \approx d\left(\frac{k^2}{\lambda^2}, e_\lambda^2\right). \quad (20.37)$$

Thus, an approximate equation for  $d$  is

$$d\left(\frac{k^2}{\lambda^2}, e_\lambda^2\right) = \frac{d_C\left(\frac{k^2}{m^2}, e^2\right)}{d_C\left(\frac{\lambda^2}{m^2}, e^2\right)}, \quad (20.38)$$

$$e_\lambda^2 = e^2 d_C\left(\frac{\lambda^2}{m^2}, e^2\right). \quad (20.39)$$

Since no gauge parameter has been included here, we can assume that this is in the Landau gauge. In a general gauge, the following parameter is included:

$$\alpha_\lambda e_\lambda^2 = \alpha e^2. \quad (20.40)$$

Combining (20.38) with (20.39),

$$e^2 d_C\left(\frac{k^2}{m^2}, e^2\right) = e^2 d_C\left(\frac{\lambda^2}{m^2}, e^2\right) d\left(\frac{k^2}{\lambda^2}, e^2 d_C\left(\frac{\lambda^2}{m^2}, e^2\right)\right). \quad (20.41)$$

In order to modify this equation, we set

$$e^2 d_C \left( \frac{k^2}{m^2}, e^2 \right) = g \left( \frac{k^2}{m^2}, e^2 \right), \quad (20.42)$$

and the above equation takes the form

$$g \left( \frac{k^2}{m^2}, e^2 \right) = Q \left( \frac{k^2}{\lambda^2}, g \left( \frac{\lambda^2}{m^2}, e^2 \right) \right). \quad (20.43)$$

This functional equation can be solved as we shall now see. We first rewrite (20.43) in the form

$$g(x, e^2) = Q \left( \frac{x}{y}, g(y, e^2) \right). \quad (20.44)$$

Then, using  $g = g(x, e^2)$  and  $g' = g(y, e^2)$ , we invert to express  $x$  and  $y$  in terms of  $g$  and  $g'$ :

$$x = h(g, e^2), \quad y = h(g', e^2). \quad (20.45)$$

Inserting this into (20.44),

$$g = Q \left( \frac{h(g, e^2)}{h(g', e^2)}, g' \right). \quad (20.46)$$

Since this means that the ratio of  $h(g, e^2)$  and  $h(g', e^2)$  does not depend on  $e^2$ ,  $h$  can be written as a product of a function of  $g$  and a function of  $e^2$ , viz.,

$$h(g, e^2) = \frac{G(g)}{\phi(e^2)}, \quad h(g', e^2) = \frac{G(g')}{\phi(e^2)}. \quad (20.47)$$

This means that

$$G(e_\lambda^2) = \frac{\lambda^2}{m^2} \phi(e^2). \quad (20.48)$$

If  $F$  is the inverse function of  $G$ , then

$$e_\lambda^2 = e^2 d_C \left( \frac{\lambda^2}{m^2}, e^2 \right) = F \left( \frac{\lambda^2}{m^2} \phi(e^2) \right), \quad (20.49)$$

or

$$e^2 d_C \left( \frac{\lambda^2}{m^2}, e^2 \right) = g \left( \frac{k^2}{m^2}, e^2 \right) = F \left( \frac{k^2}{m^2} \phi(e^2) \right) = F \left( \frac{k^2 \lambda^2}{\lambda^2 m^2} \phi(e^2) \right). \quad (20.50)$$

Combining (20.48) with (20.50),

$$e^2 d_C \left( \frac{k^2}{m^2}, e^2 \right) = F \left( \frac{k^2}{\lambda^2} G \left( e_\lambda^2 \right) \right). \quad (20.51)$$

Note that, from (20.39),

$$e^2 d_C \left( \frac{\lambda^2}{m^2}, e^2 \right) d \left( \frac{k^2}{\lambda^2}, e^2 d_C \left( \frac{\lambda^2}{m^2}, e^2 \right) \right) = e_\lambda^2 d \left( \frac{k^2}{\lambda^2}, e_\lambda^2 \right). \quad (20.52)$$

Comparing this with (20.41) and (20.51),

$$F \left( \frac{k^2}{\lambda^2} G \left( e_\lambda^2 \right) \right) = e_\lambda^2 d \left( \frac{k^2}{\lambda^2}, e_\lambda^2 \right), \quad (20.53)$$

$$\frac{k^2}{\lambda^2} G \left( e_\lambda^2 \right) = G \left( e_\lambda^2 d \left( \frac{k^2}{\lambda^2}, e_\lambda^2 \right) \right). \quad (20.54)$$

Thus,

$$\ln \frac{k^2}{\lambda^2} = \ln G \left( e_\lambda^2 d \left( \frac{k^2}{\lambda^2}, e_\lambda^2 \right) \right) - \ln G \left( e_\lambda^2 \right). \quad (20.55)$$

We now define the function  $\psi(x)$  introduced by Gell-Mann and Low:

$$\frac{1}{\psi(x)} = \frac{d}{dx} \ln G(x). \quad (20.56)$$

Therefore,

$$\ln \frac{k^2}{\lambda^2} = \int_{e_\lambda^2}^{e_\lambda^2 d \left( \frac{k^2}{\lambda^2}, e_\lambda^2 \right)} \frac{dx}{\psi(x)}. \quad (20.57)$$

Similarly, an equation derived from (20.50), viz.,

$$\frac{k^2}{m^2} \phi(e^2) = G \left( e^2 d_C \left( \frac{k^2}{m^2}, e^2 \right) \right), \quad (20.58)$$

can also be written in the form

$$\ln \frac{k^2}{m^2} = \int_{F(\phi(e^2))}^{e^2 d_C \left( \frac{k^2}{m^2}, e^2 \right)} \frac{dx}{\psi(x)}. \quad (20.59)$$

In perturbation theory,

$$F(\phi(e^2)) = e^2 d_C(1, e^2) = e^2 - \frac{5}{36\pi^2} e^4 + \dots \equiv e'^2. \quad (20.60)$$

It follows immediately from (20.50) that

$$\left[ m \frac{\partial}{\partial m} + \beta(e) \frac{\partial}{\partial e} \right] e^2 d_C \left( \frac{k^2}{m^2}, e^2 \right) = 0, \quad (20.61)$$

where

$$e\beta(e) = \frac{\phi(e^2)}{\phi'(e^2)}. \quad (20.62)$$

Additionally, from (20.59),

$$\left[ m \frac{\partial}{\partial m} + \frac{\psi(e'^2)}{e'} \frac{\partial}{\partial e'} \right] e^2 d_C \left( \frac{k^2}{m^2}, e^2 \right) = 0. \quad (20.63)$$

This implies that

$$\frac{\psi(e'^2)}{e'} \frac{\partial}{\partial e'} = \beta(e) \frac{\partial}{\partial e}. \quad (20.64)$$

Equation (20.61) is the Callan–Symanzik equation [179–181], but without a mass term.

If we compute  $\psi$  and  $\beta$  in perturbation theory, we find

$$\psi(x) = \frac{1}{12\pi^2} \left( x^2 + \frac{3}{16\pi^2} x^3 + \dots \right), \quad (20.65)$$

$$\beta(e) \approx \frac{\psi(e^2)}{e} \approx \frac{e^3}{12\pi^2}. \quad (20.66)$$

In fact, since the function  $d$  does not depend on the gauge parameter  $\alpha$  in QED, the above discussion holds true in arbitrary gauges. However, the function  $s$  does depend on  $\alpha$ .



After the preparation above, we come back to the equation for  $s$ . We rewrite (20.35) using the result for  $d$ :

$$s\left(\frac{p^2}{\lambda^2}, \frac{p^2}{\lambda'^2}, e_\lambda^2\right) = s\left(\frac{p^2}{\lambda^2}, \frac{p^2}{\lambda'^2}, F\left(\frac{\lambda^2}{m^2}\phi(e^2)\right)\right). \quad (20.67)$$

However, according to (20.36), this function should not depend on  $\lambda$ , so the right-hand side of this equation should only be a function of

$$\frac{p^2}{m^2}\phi(e^2) \quad \text{and} \quad \frac{p^2}{\lambda'^2}. \quad (20.68)$$

Therefore,

$$s\left(\frac{p^2}{\lambda^2}, \frac{p^2}{\lambda'^2}, e_\lambda^2\right) = R\left(\frac{p^2}{\lambda'^2}, \frac{p^2}{m^2}\phi(e^2)\right). \quad (20.69)$$

In every gauge except for the Landau gauge, this depends on the variable  $\alpha e^2$  as well. The reason is that, since this equation does not depend on  $\lambda$ ,  $\alpha$  should always be included in the combination  $\alpha_\lambda e_\lambda^2 = \alpha e^2$ . Setting

$$s_C\left(\frac{p^2}{m^2}, e^2, \alpha e^2\right) = A(e^2, \alpha e^2)H\left(\frac{p^2}{\lambda'^2}, \alpha e^2, \frac{p^2}{m^2}\phi(e^2)\right), \quad (20.70)$$

we see that (20.69) holds true. This  $H$  satisfies

$$\mathcal{D}H \equiv \left[ m \frac{\partial}{\partial m} + \beta(e) \frac{\partial}{\partial e} + \delta(e, \alpha) \frac{\partial}{\partial \alpha} \right] H = 0, \quad (20.71)$$

where

$$\delta(e, \alpha) = -\frac{2\alpha}{e}\beta(e). \quad (20.72)$$

We now define  $\gamma_{\text{el}}$  by

$$\mathcal{D} \ln A = -2\gamma_{\text{el}}(e, \alpha). \quad (20.73)$$

Thus,

$$(\mathcal{D} + 2\gamma_{\text{el}})s_C\left(\frac{p^2}{m^2}, e^2, \alpha e^2\right) = 0. \quad (20.74)$$

This equation is also called the Callan–Symanzik equation without the mass term.

In this section, we have treated renormalization groups of the first kind, assuming that the mass  $m^2$  is smaller than the mass corresponding to the four-momentum or the renormalization point. As a result, a derivative with respect to the mass has been included. However, if we think of this as a derivative with respect to the four-momentum, then the story can be considered to be closed in a world with a fixed mass. In this section we have treated an approximation which holds true only if  $m^2$  is small, but what would the renormalization group equations look like if we did not make this approximation? We will discuss this issue in the next section.

### 20.3 Ovsianikov's Equation

We begin with the equation for  $d$  in which  $m$  is not neglected:

$$d\left(\frac{k^2}{\lambda^2}, \frac{m^2}{\lambda^2}, e_\lambda^2\right) = \frac{d_C\left(\frac{k^2}{m^2}, e^2\right)}{d_C\left(\frac{\lambda^2}{m^2}, e^2\right)}. \quad (20.75)$$

Eliminating  $d_C$  from this equation,

$$d\left(\frac{k^2}{\lambda_2^2}, \frac{m^2}{\lambda_2^2}, e_2^2\right) = \frac{d\left(\frac{k^2}{\lambda_1^2}, \frac{m^2}{\lambda_1^2}, e_1^2\right)}{d\left(\frac{\lambda_2^2}{\lambda_1^2}, \frac{m^2}{\lambda_1^2}, e_1^2\right)}, \quad (20.76)$$

$$e_2^2 = e_1^2 d\left(\frac{\lambda_2^2}{\lambda_1^2}, \frac{m^2}{\lambda_1^2}, e_1^2\right), \quad (20.77)$$

where we have used

$$d\left(1, \frac{m^2}{\lambda^2}, e_1^2\right) = 1, \quad (20.78)$$

considering that  $e_1$  and  $e_2$  express  $e_{\lambda_1}$  and  $e_{\lambda_2}$ , respectively. In addition, the combination

$$e_\lambda^2 d\left(\frac{k^2}{\lambda^2}, \frac{m^2}{\lambda^2}, e_\lambda^2\right) = e^2 d_C\left(\frac{k^2}{m^2}, e^2\right) \quad (20.79)$$

is called an *invariant charge* because it does not depend on  $\lambda$ . We now introduce the variables

$$t = \frac{\lambda_2^2}{\lambda_1^2}, \quad x = \frac{k^2}{\lambda_1^2}, \quad y = \frac{m^2}{\lambda_1^2}, \quad e^2 = e_1^2. \quad (20.80)$$

Hence, (20.76) becomes a rigorous equation with the form

$$e^2 d(x, y, e^2) = e^2 d(t, y, e^2) d\left(\frac{x}{t}, \frac{y}{t}, e^2 d(t, y, e^2)\right). \quad (20.81)$$

Similar equations can be found for the other functions. In the expression

$$S_F(\lambda, \lambda', p, e_\lambda^2) = \frac{1}{p^2 + m^2 - i\epsilon} \left[ -i(p \cdot \gamma) a(\lambda, \lambda', p, e_\lambda^2) + mb(\lambda, \lambda', p, e_\lambda^2) \right], \quad (20.82)$$

we write both  $a$  and  $b$  as

$$s(x, y, a, e^2), \quad (20.83)$$

where

$$x = \frac{p^2}{\lambda^2}, \quad y = \frac{m^2}{\lambda^2}, \quad a = \frac{m^2}{\lambda'^2}, \quad e^2 = e_\lambda^2. \quad (20.84)$$

In the following, we do not write  $a$  and  $\alpha e^2$  explicitly. Therefore, the equation for  $s$  is

$$\frac{s(x, y, e^2)}{s(t, y, e^2)} = \frac{s\left(\frac{x}{t}, \frac{y}{t}, e^2 d(t, y, e^2)\right)}{s\left(1, \frac{y}{t}, e^2 d(t, y, e^2)\right)}. \quad (20.85)$$

A similar equation holds true for the vertex function as well. Expanding the vertex function in a covariant polynomial,

$$\Gamma_\mu(p, q, k) = \sum_A c_\mu^A(p, q, \gamma) \Gamma_A(p^2, q^2, k^2), \quad (20.86)$$

and writing the coefficients as  $\Gamma_A(x, y, z, u, e^2)$ , we find

$$\frac{\Gamma_A(x, y, z, u, e^2)}{\Gamma_A(t, y, z, u, e^2)} = \frac{\Gamma_A\left(\frac{x}{t}, \frac{y}{t}, \frac{z}{t}, \frac{u}{t}, e^2 d(t, u, e^2)\right)}{\Gamma_A\left(1, \frac{y}{t}, \frac{z}{t}, \frac{u}{t}, e^2 d(t, u, e^2)\right)}, \quad (20.87)$$

where the variables are

$$x = \frac{p^2}{\lambda^2}, \quad y = \frac{q^2}{\lambda^2}, \quad z = \frac{k^2}{\lambda^2}, \quad u = \frac{m^2}{\lambda^2}, \quad e^2 = e_\lambda^2. \quad (20.88)$$

Since a condition like (20.78) cannot be used in these equations, we can not eliminate expressions like  $s(1, \dots)$  and  $\Gamma_A(1, \dots)$ . We therefore introduce the following functions:

$$\Phi(y, e^2) = \frac{\partial}{\partial x} d(x, y, e^2) \Big|_{x=1}, \quad (20.89)$$

$$\Psi(y, e^2) = \frac{\partial}{\partial x} \ln s(x, y, e^2) \Big|_{x=1}. \quad (20.90)$$

Now, differentiating the functional equations (20.81) and (20.85) with respect to  $x$  and setting  $t = x$ , we obtain

$$\frac{\partial}{\partial x} e^2 d(x, y, e^2) = \frac{e^2 d(x, y, e^2)}{x} \Phi\left(\frac{y}{x}, e^2 d(x, y, e^2)\right), \quad (20.91)$$

$$\frac{\partial}{\partial x} s(x, y, e^2) = \frac{s(x, y, e^2)}{x} \Psi\left(\frac{y}{x}, e^2 d(x, y, e^2)\right). \quad (20.92)$$

These are *Ovsianikov's equations*. Both are highly non-linear, but  $\Phi$  and  $\Psi$  determine the initial conditions for  $d$  and  $s$  at  $x = 1$  and they also determine their differential equations.

If we assume that  $k^2 \gg m^2$ , i.e.,  $x \gg y$ , and we neglect any  $y$ -dependence, then they coincide with the results in the previous section.

### 1. Neglecting $y$

We now have

$$\frac{\partial}{\partial x} e^2 d(x, e^2) = \frac{e^2 d(x, e^2)}{x} \Phi(e^2 d(x, e^2)), \quad (20.93)$$

$$\frac{\partial}{\partial x} s(x, e^2) = \frac{s(x, e^2)}{x} \Psi(e^2 d(x, e^2)). \quad (20.94)$$

Therefore,

$$\ln x = \int_{e^2}^{e^2 d(x, e^2)} \frac{dz}{z \Phi(z)}. \quad (20.95)$$

If we set  $\psi(z) = z\Phi(z)$ , this result coincides with (20.57). Similarly, if we solve the equation for  $s$ , then

$$\ln \frac{s(x, e^2)}{s(x_0, e^2)} = \int_{e^2 d(x_0, e^2)}^{e^2 d(x, e^2)} dz \frac{\Psi(z)}{z\Phi(z)}. \tag{20.96}$$

In perturbation theory,

$$d(x, e^2) = 1 + \frac{e^2}{12\pi^2} \ln x + \frac{e^4}{16\pi^2} \left( \frac{1}{4\pi^2} \ln x + \frac{1}{16\pi} \ln^2 x \right) + \dots, \tag{20.97}$$

$$\Phi(z) = \frac{\partial}{\partial x} d(x, z) \Big|_{x=1} = \frac{z}{12\pi^2} \left( 1 + \frac{3z}{16\pi^2} + \dots \right) \quad (z \ll 1). \tag{20.98}$$

**2. Solution Around the Mass Shell**

Implementing a perturbative expansion of the function  $s$  in the vicinity of the mass shell, we obtain an infrared divergence. We thus derive a solution that is not based on perturbation theory, using the renormalization group equations. In the vicinity of the mass shell, we set

$$x \approx 1, \quad y \approx -1, \tag{20.99}$$

then differentiate (20.85) with respect to  $x$  and set  $t = x'$ , whence

$$\ln \frac{s(x, -1, e^2)}{s(x_0, -1, e^2)} = \int_{x_0}^x \frac{dx'}{x'} \frac{\partial}{\partial \xi} \ln s \left( \xi, -\frac{1}{x'}, e^2 d(x', -1, e^2) \right) \Big|_{\xi=1}.$$

To lowest order in both  $a$  and  $b$ ,

$$\ln \frac{s(x, -1, e^2)}{s(x_0, -1, e^2)} \approx -\frac{3e^2}{8\pi^2} \ln \frac{x-1}{x_0-1}. \tag{20.100}$$

Therefore,

$$s(x, -1, e^2) \propto (x-1)^{-3e^2/8\pi^2}, \tag{20.101}$$

or

$$s\left(\frac{p^2}{m^2}, e^2\right) \propto \left(1 + \frac{p^2}{m^2}\right)^{-3e^2/8\pi^2}. \tag{20.102}$$

In a general gauge, we replace the factor of 3 in the exponent by  $(3 - \alpha)$ . Each term of the perturbative expansion in this equation diverges on the mass shell.

Ovsianikov's equation is exact but non-linear. However, we can derive a linear equation. Although in the above we differentiated the renormalization group functional equation with respect to  $x$  and then set  $t = x$ , this time we differentiate it with respect to  $t$  and set  $t = 1$  :

$$\left[ \Phi(y, e^2) - \left( x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right) + e^2 \Phi(y, e^2) \frac{\partial}{\partial e^2} \right] d(x, y, e^2) = 0. \quad (20.103)$$

This is linear. Up to now  $x = k^2/\lambda^2$  and  $y = m^2/\lambda^2$ , but from now on we shall write  $\lambda$  as  $\mu$ . Therefore,

$$-2 \left( x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right) = \mu \frac{\partial}{\partial \mu}. \quad (20.104)$$

This yields an equation of the form

$$\left[ \mu \frac{\partial}{\partial \mu} + \beta(e, y) \frac{\partial}{\partial e} + 2\gamma_{\text{ph}}(e, y) \right] d(x, y, e^2) = 0, \quad (20.105)$$

$$\beta(e, y) = e\Phi(y, e^2), \quad \gamma_{\text{ph}} = \Phi(y, e^2). \quad (20.106)$$

## 20.4 Linear Equations for the Renormalization Group

We consider a generalization of the linear equation given at the end of the last section. To do this, we start with the renormalization group equation. As an example, we choose the  $\phi^4$ -theory:

$$\mathcal{L} = -\frac{1}{2}[(\partial_\lambda \phi)^2 + m^2 \phi^2] - g \frac{\phi^4}{4!}. \quad (20.107)$$

Indicating the unrenormalized Green's functions with a subscript 0, the renormalized Green's function is given by

$$G_0^{(n)}(p_1, \dots, p_n; g_0) = Z_3^{n/2}(\mu) G^{(n)}(p_1, \dots, p_n; g(\mu)), \quad (20.108)$$

where the renormalized coupling constant and four-point vertex function are given by

$$g_0 = Z_1(\mu) Z_3^{-2}(\mu) g(\mu), \quad (20.109)$$

$$\Gamma(p_i; g(\mu)) = Z_1(\mu) \Gamma_0(p_i; g_0), \quad i = 1, 2, 3, 4, \quad (20.110)$$

and  $Z_1$ ,  $Z_3$ , and so on have been determined by the renormalization conditions

$$(p^2 + m^2)G^{(2)}(p; g(\mu))\big|_{p^2=\mu^2} = -1, \quad (20.111)$$

$$\Gamma(p_i; g(\mu)) = 1, \quad \text{with } p_i \cdot p_j = \frac{\mu^2}{3}(4\delta_{ij} - 1). \quad (20.112)$$

The four-point vertex function  $\Gamma_0$  is defined by the unrenormalized four-point Green's function without the propagators and the coupling constant. For the definition of the  $n$ -point Green's function, the reader is referred to Sect. 15.3.

Since the unrenormalized Green's function is not related to the renormalization point  $\mu$ ,

$$\mu \frac{\partial}{\partial \mu} G_0^{(n)}(p_1, \dots, p_n; g_0) = 0. \quad (20.113)$$

Inserting (20.108) into this equation, we obtain

$$\left( \mu \frac{\partial}{\partial \mu} + \beta \frac{\partial}{\partial g} + n\gamma \right) G^{(n)}(p_1, \dots, p_n; g(\mu)) = 0, \quad (20.114)$$

where

$$\beta = \mu \frac{\partial}{\partial \mu} g(\mu) = \left( 2\mu \frac{\partial}{\partial \mu} \ln Z_3 - \mu \frac{\partial}{\partial \mu} \ln Z_1 \right) g(\mu), \quad (20.115)$$

$$\gamma = \frac{1}{2} \mu \frac{\partial}{\partial \mu} \ln Z_3. \quad (20.116)$$

Both of these are functions of  $m^2/\mu^2$  and  $g$ . From now on, we shall write  $g(\mu)$  simply as  $g$ . Therefore, writing (20.114) in more detail, we have

$$\left[ \mu \frac{\partial}{\partial \mu} + \beta \left( g, \frac{m^2}{\mu^2} \right) \frac{\partial}{\partial g} + n\gamma \left( g, \frac{m^2}{\mu^2} \right) \right] G^{(n)}(p_1, \dots, p_n; \mu, g) = 0. \quad (20.117)$$

In this equation,  $g$  and  $\mu$  are treated as independent variables.

Similarly, we can write down a linear renormalization equation in QED. If the numbers of photon legs and electron legs in the Green's function are  $n$  and  $l$ , respectively, we have

$$\begin{aligned} G_0^{(n,l)}(k_1, \dots, k_n; p_1, \dots, p_l; e_0) \\ = Z_3^{n/2}(\mu) Z_2^{l/2}(\mu) G^{(n,l)}(k_1, \dots, k_n; p_1, \dots, p_l; e(\mu)), \end{aligned} \quad (20.118)$$

where

$$e_0^2 = Z_3^{-1}(\mu)e^2(\mu). \quad (20.119)$$

This implies

$$\left[ \mu \frac{\partial}{\partial \mu} + \beta \left( e, \frac{m^2}{\mu^2} \right) \frac{\partial}{\partial e} + n\gamma_{\text{ph}} \left( e, \frac{m^2}{\mu^2} \right) + l\gamma_{\text{ph}} \left( e, \frac{m^2}{\mu^2} \right) \right] G^{(n,l)}(\dots; \mu, e) = 0. \quad (20.120)$$

This equation is in the Landau gauge and a generalization of (20.93). In a general gauge  $\alpha$ ,

$$\left[ \mu \frac{\partial}{\partial \mu} + \beta \left( e, \frac{m^2}{\mu^2} \right) \frac{\partial}{\partial e} + \delta \left( e, \frac{m^2}{\mu^2}, \alpha \right) \frac{\partial}{\partial \alpha} + \gamma^{(n,l)} \right] G^{(n,l)}(\dots; \mu, e, \alpha) = 0, \quad (20.121)$$

where

$$\gamma^{(n,l)} = n\gamma_{\text{ph}} \left( e, \frac{m^2}{\mu^2} \right) + l\gamma_{\text{el}} \left( e, \frac{m^2}{\mu^2}, \alpha \right). \quad (20.122)$$

As stated in the last section,

$$\gamma_{\text{ph}} \left( e, \frac{m^2}{\mu^2} \right) = \Phi \left( \frac{m^2}{\mu^2}, e^2 \right), \quad (20.123)$$

$$\beta \left( e, \frac{m^2}{\mu^2} \right) = e\gamma_{\text{ph}} \left( e, \frac{m^2}{\mu^2} \right), \quad \delta \left( e, \frac{m^2}{\mu^2}, \alpha \right) = -2\alpha\gamma_{\text{ph}} \left( e, \frac{m^2}{\mu^2} \right). \quad (20.124)$$

The last expression was derived from (12.98). The renormalization equations derived in this section are linear and homogeneous, but the coefficients are functions of two variables, so they are hard to deal with. In contrast, the coefficients in the Callan–Symanzik equation introduced in the next section are single-variable functions, while the equation itself becomes inhomogeneous.

## 20.5 Callan–Symanzik Equation

So far we have considered a renormalization group of the first kind. In this section, we move on to the second kind, i.e., we introduce the idea of changing the particle mass.



We first define a Fourier transform of the Green's function. Using the notation  $(dp) = d^4 p_1 \dots d^4 p_n$ ,

$$\begin{aligned} \langle \mathbf{0} | T[\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle_{\text{conn}} & \quad (20.125) \\ &= \left[ \frac{i}{(2\pi)^4} \right]^{n-1} \int (dp) e^{i(p_1 \cdot x_1 + \dots + p_n \cdot x_n)} \delta^4(p_1 + \dots + p_n) G^{(n)}(p_1, \dots, p_n), \end{aligned}$$

$$\begin{aligned} \langle \mathbf{0} | T[A(x)\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle_{\text{conn}} & \quad (20.126) \\ &= \left[ \frac{i}{(2\pi)^4} \right]^n \int d^4 q (dq) e^{i(q \cdot x + p_1 \cdot x_1 + \dots + p_n \cdot x_n)} \delta^4(q + p_1 + \dots + p_n) A^{(n)}(q; p_1, \dots, p_n). \end{aligned}$$

We start with an unrenormalized theory, then introduce the quantity

$$T^{(0)}(x) = m_0^2 [\varphi^{(0)}(x)]^2. \quad (20.127)$$

To lowest order,

$$G_0^{(2)}(p) = -\frac{1}{p^2 + m_0^2 - i\epsilon}. \quad (20.128)$$

Therefore,

$$m_0 \frac{\partial}{\partial m_0} \left( -\frac{1}{p^2 + m_0^2 - i\epsilon} \right) = \frac{1}{p^2 + m_0^2 - i\epsilon} 2m_0^2 \frac{1}{p^2 + m_0^2 - i\epsilon}. \quad (20.129)$$

This equation tells us that a derivative with respect to  $m_0$  is equivalent to inserting  $T^{(0)}$  in an unrenormalized form, i.e.,

$$m_0 \frac{\partial}{\partial m_0} G_0^{(n)}(p_1, \dots, p_n) = T_0^{(n)}(0; p_1, \dots, p_n) \equiv \Delta G_0^{(n)}(p_1, \dots, p_n). \quad (20.130)$$

Next, we consider a renormalization on the mass shell:

$$G_0^{(n)}(p_1, \dots, p_n) = Z_3^{n/2} G^{(n)}(p_1, \dots, p_n; g). \quad (20.131)$$

From now on, we consider all renormalized quantities like  $m$  and  $g$  to be functions of  $m_0$ ,  $g_0$ , and a cut-off momentum  $\Lambda$ . We think about the derivative with respect to  $m_0$  in the case where  $g_0$  and  $\Lambda$  are fixed. First, we define  $Z$  by

$$Z m_0 \frac{\partial m}{\partial m_0} = m. \quad (20.132)$$

Therefore, for functions of  $g$  and  $m$ ,

$$Zm_0 \frac{\partial}{\partial m_0} f = \left( Zm_0 \frac{\partial m}{\partial m_0} \frac{\partial}{\partial m} + Zm_0 \frac{\partial g}{\partial m_0} \frac{\partial}{\partial g} \right) f \equiv \mathcal{D} f, \quad (20.133)$$

where

$$\mathcal{D} = m \frac{\partial}{\partial m} + \beta(g) \frac{\partial}{\partial g}, \quad \beta(g) = Zm_0 \frac{\partial g}{\partial m_0}. \quad (20.134)$$

Thus, we redefine  $T$  by

$$T(x) = Zm_0^2 [\varphi^{(0)}(x)]^2. \quad (20.135)$$

This yields the Callan–Symanzik equation [179–181]:

$$\left[ m \frac{\partial}{\partial m} + \beta(g) \frac{\partial}{\partial g} + n\gamma_\varphi(g) \right] G^{(n)}(p_1, \dots, p_n) = T^{(n)}(0; p_1, \dots, p_n), \quad (20.136)$$

where  $\gamma_\varphi$  is called the *anomalous dimension of the field  $\varphi$* , defined by

$$\gamma_\varphi(g) = \frac{1}{2} Zm_0 \frac{\partial}{\partial m_0} \ln Z_3. \quad (20.137)$$

Denoting the multiplicative renormalization of  $\varphi^2$  by  $S$ , we rewrite  $T$  as

$$T(x) = m^2 S(x). \quad (20.138)$$

Focusing on the pole at  $p^2 + m^2 = 0$  when  $n = 2$ , multiplying (20.136) corresponding to  $n = 2$  by  $(p^2 + m^2)^2$ , and taking the limit  $p^2 + m^2 \rightarrow 0$ , the left-hand side becomes  $2m^2$ . Therefore, the right-hand side should be  $2m^2$  as well. Using the LSZ reduction formula,  $S$  turns out to satisfy the following renormalization condition:

$$\sqrt{2p_0} \langle \mathbf{p} | S(x) | \mathbf{p} \rangle \sqrt{2p_0} = 1. \quad (20.139)$$

We now define a multiple mass insertion term. Inserting  $T$  into  $G^{(n)}$  a total of  $k$  times and writing it as  $\delta^k G^{(n)}$ ,

$$\left\{ \mathcal{D} + n\gamma_\varphi(g) - k[2 - \gamma_S(g)] \right\} \Delta^k G^{(n)} = \Delta^{k+1} G^{(n)}, \quad (20.140)$$

where

$$\gamma_S(g) = Z m_0 \frac{\partial}{\partial m_0} \ln Z_S, \quad [\varphi^{(0)}(x)]^2 = Z_S S(x). \quad (20.141)$$

On the other hand, an unrenormalized multiple mass insertion term is

$$\Delta^k G_k^{(n)} = (2m_0^2)^k \left( \frac{\partial}{\partial m_0^2} \right)^k G_0^{(n)}, \quad (20.142)$$

and its renormalization is

$$\Delta^k G^{(n)} = Z^k Z_3^{-n/2} \Delta^k G_0^{(n)}. \quad (20.143)$$

Inserting this into the Callan–Symanzik equation,

$$\gamma_S(g) = 2(1 - Z) - m_0 \frac{\partial}{\partial m_0} Z. \quad (20.144)$$

According to perturbation theory,

$$\beta(g) = \frac{3g^2}{16\pi^2}, \quad \gamma_\varphi(g) = \frac{g^2}{12(16\pi^2)^2}, \quad \gamma_S(g) = -\frac{g}{16\pi^2}. \quad (20.145)$$

An advantage compared to the renormalization group of the first kind is that  $\beta$  and  $\gamma_\varphi$  become functions of  $g$  alone. A disadvantage is that the equation becomes inhomogeneous. However, if  $p_i^2 \gg m^2$ , since the momentum dependence of  $\Delta G^{(n)}$  decreases by a factor of momentum squared compared to  $G^{(n)}$ , we can neglect  $\Delta G^{(n)}$  in comparison to  $G^{(n)}$ . We then obtain a homogeneous equation:

$$[\mathcal{D} + n\gamma_\varphi(g)]G^{(n)} \approx 0. \quad (20.146)$$

Similarly, if  $\mu^2 \gg m^2$ , even in the renormalization group of the first kind, the  $m^2/\mu^2$ -dependence of the coefficient function is dropped, and it becomes a function of  $g$  alone, so the equations coincide.

We now consider QED. The renormalizations are

$$G_0^{(n,l)} = Z_3^{n/2} Z_2^{l/2} G^{(n,l)}, \quad (20.147)$$

$$\Delta^k G_0^{(n,l)} = m_0^k \left( \frac{\partial}{\partial m_0} \right)^k G_0^{(n,l)} = Z^{-k} Z_3^{n/2} Z_2^{l/2} \Delta^k G^{(n,l)}, \quad (20.148)$$

$$Z m_0 \frac{\partial}{\partial m_0} m = m. \quad (20.149)$$

When we differentiate with respect to  $m_0$ , we fix  $e_0$ ,  $\alpha_0$ , and  $\Lambda$ , and we define the following quantities:

$$\begin{aligned}\beta(e) &= Z m_0 \frac{\partial e}{\partial m_0} , \\ \gamma_{\text{el}}(e, \alpha) &= \frac{1}{2} Z m_0 \frac{\partial}{\partial m_0} \ln Z_2 , \\ \gamma_{\text{ph}}(e) &= \frac{1}{2} Z m_0 \frac{\partial}{\partial m_0} \ln Z_3 , \\ \delta(e, \alpha) &= Z m_0 \frac{\partial}{\partial m_0} \alpha .\end{aligned}\tag{20.150}$$

Combining these with the Ward identities,

$$e^2 = Z_3 e_0^2 , \quad e^2 \alpha = e_0^2 \alpha_0 ,\tag{20.151}$$

we obtain the following equations, which correspond to (20.124):

$$\beta(e) = e \gamma_{\text{ph}}(e) , \quad \delta(e, \alpha) = -2\alpha \gamma_{\text{ph}}(e) .\tag{20.152}$$

We thus introduce the differential operator

$$\mathcal{D} = m \frac{\partial}{\partial m} + \beta(e) \frac{\partial}{\partial e} - 2\alpha \gamma_{\text{ph}}(e) \frac{\partial}{\partial \alpha} .\tag{20.153}$$

Therefore, the Callan–Symanzik equation in QED becomes

$$[\mathcal{D} + \gamma^{(n,l)} - k(1 - \gamma_S)] \Delta^k G^{(n,l)} = \Delta^{k+1} G^{(n,l)} .\tag{20.154}$$

In this case, it turns out that  $m(\bar{\psi}\psi)_R$  is inserted as  $T$ .  $(\bar{\psi}\psi)_R$  is the quantity obtained when  $\bar{\psi}\psi$  has been properly renormalized in a multiplicative way. The renormalization condition is

$$\langle \mathbf{p} | [\bar{\psi}(x)\psi(x)]_R | \mathbf{p} \rangle = \bar{u}(p)u(p) ,\tag{20.155}$$

where  $u$  has been normalized so that  $u^\dagger(p)u(p) = 1$ .

In general, the renormalization may be not only multiplicative, but also a linear transformation:

$$O_i^{(0)} = \sum_j Z_{ij} O_j .\tag{20.156}$$

For the Green's function,

$$O_{i(0)}^{(n,l)} = \sum_j Z_3^{n/2} Z_2^{l/2} Z_{ij} O_j^{(n,l)}. \quad (20.157)$$

Therefore,

$$[\mathcal{D} + \gamma^{(n,l)} - k(1 - \gamma_S)] \Delta^k O_i^{(n,l)} + \sum_j \gamma_{ij} \Delta^k O_j^{(n,l)} = \Delta^{k+1} O_i^{(n,l)}, \quad (20.158)$$

where

$$\gamma_S = 1 - Z - m_0 \frac{\partial Z}{\partial m_0}, \quad \gamma_{ij} = \sum_k Z_{ik}^{-1} \left( Z m_0 \frac{\partial}{\partial m_0} Z_{kj} \right). \quad (20.159)$$

To lowest order in QED,

$$\beta = \frac{e^3}{12\pi^2}, \quad \gamma_{ph} = \frac{e^2}{12\pi^2}, \quad \gamma_{el} = \frac{e^2}{16\pi^2} \alpha, \quad \gamma_S = -\frac{3e^2}{8\pi^2}. \quad (20.160)$$

Next, let us consider renormalizing the electron mass multiplicatively. This is possible only if the divergence of the self-energy becomes logarithmic:

$$Z_m = \frac{m}{m_0}, \quad \gamma_m = Z m_0 \frac{\partial}{\partial m_0} \ln Z_m. \quad (20.161)$$

Thus, immediately from the definition,

$$\gamma_m = 1 - Z. \quad (20.162)$$

Thus, taking  $\gamma_m$  as a function of  $e$  alone,

$$\begin{aligned} \gamma_S &= 1 - Z - Z m_0 \frac{\partial}{\partial m_0} \ln Z \\ &= 1 - Z - Z m_0 \frac{\partial}{\partial m_0} \ln(1 - \gamma_m) \\ &= \gamma_m - \mathcal{D} \ln(1 - \gamma_m) \\ &= \gamma_m - \beta(e) \frac{\partial}{\partial e} \ln(1 - \gamma_m), \end{aligned} \quad (20.163)$$

or

$$\beta \frac{d\gamma_m}{de} = (1 - \gamma_m)(\gamma_S - \gamma_m). \quad (20.164)$$

Therefore, the equation for  $Z$  is

$$\left[ \beta \frac{d}{de} + (1 - \gamma_S) \right] Z^{-1}(e) = 1, \quad Z(0) = 1. \tag{20.165}$$

From this, we see that  $Z$  is a finite function of  $e$  alone.

## 20.6 Homogeneous Callan–Symanzik Equation

The Callan–Symanzik equation introduced in the previous section is inhomogeneous. It is in fact a system of equations for an infinite number of functions  $G^{(n)}$ ,  $\Delta G^{(n)}$ , and so on. We shall now rewrite them in a homogeneous and closed form. This is actually possible, but the price to pay for this is that the number of independent variables increases by one.

In order to write this infinite number of functions collectively, we introduce a generating function:

$$G^{(n)}(p_1, \dots, p_n; m^2, g, K) = \sum_{j=0}^{\infty} \frac{K^j}{j!} \Delta^j G^{(n)}(p_1, \dots, p_n; m^2, g). \tag{20.166}$$

$$\tilde{\mathcal{D}} = \mathcal{D} - \left\{ 1 + [2 - \gamma_S(g)]K \right\} \frac{\partial}{\partial K}. \tag{20.167}$$

Therefore, from (20.140),

$$[\tilde{\mathcal{D}} + n\gamma_\varphi(g)]G^{(n)}(p_1, \dots, p_n; m, g, K) = 0. \tag{20.168}$$

As for the renormalization group of the first kind, we have thus obtained a homogeneous form. Note that the number of independent variables also coincides. In the case of the first kind, those variables are  $\mu$ ,  $\nu$ , and  $m$ , while in the case of the second kind, they are  $m$ ,  $g$ , and  $K$ . Therefore, by a suitable variable transformation, each can be transformed into the other. This was proven by Higashijima and the author in [182].

Similarly in QED, if we take

$$G^{(n,l)}(p_i; m, e, K) = \sum_{j=0}^{\infty} \frac{K^j}{j!} \Delta^j G^{(n,l)}(p_j; m, e), \tag{20.169}$$

$$\tilde{\mathcal{D}} = \mathcal{D} - \left\{ 1 + [1 - \gamma_S(e)]K \right\} \frac{\partial}{\partial K}, \tag{20.170}$$

then

$$[\tilde{\mathcal{G}} + \gamma^{(n,l)}]G^{(n,l)}(p_i; m, e, K) = 0. \quad (20.171)$$

Now, an infinitesimal renormalization group transformation varies a function  $Q$  of  $m$ ,  $g$ , and  $K$  via an infinitesimal parameter  $\delta\rho$  (scalar theory):

$$\delta Q = (\tilde{\mathcal{G}}Q)\delta\rho, \quad (20.172)$$

where

$$\delta m = m\delta\rho, \quad \delta g = \beta(g)\delta\rho, \quad \delta K = -\left\{1 + [2 - \gamma_S(g)]K\right\}\delta\rho. \quad (20.173)$$

Clearly,  $m$  is not fixed, so this is a property of the second kind. From (20.168), the response of the Green's function is

$$\delta G^{(n)} = -n\gamma_\varphi(g)G^{(n)}\delta\rho. \quad (20.174)$$

The operators yielding finite transformations are then

$$R_\rho = \exp(\rho\tilde{\mathcal{G}}). \quad (20.175)$$

These operators form a group in the sense that

$$R_\rho R_{\rho'} = R_{\rho+\rho'}. \quad (20.176)$$

We thus set

$$R_\rho[g] = \bar{g}(\rho), \quad R_\rho[K] = \bar{K}(\rho), \quad R_\rho[m] = me^\rho. \quad (20.177)$$

This tells us, for instance, that  $\bar{g}(\rho)$  is the solution of (20.173) for the initial condition  $\bar{g}(0) = g$ . Therefore, from (20.173),

$$G^{(n)}(p_i; m, g, K) = \exp\left[n \int_0^\rho d\rho' \gamma_\varphi(\bar{g}(\rho'))\right] G^{(n)}(p_i; me^\rho, \bar{g}(\rho), \bar{K}(\rho)). \quad (20.178)$$

Note that, from a dimensional analysis,

$$G^{(n)}(p_i; me^\rho, \bar{g}(\rho), \bar{K}(\rho)) = e^{(4-3n)\rho} G^{(n)}(p_i e^{-\rho}, m, \bar{g}(\rho), \bar{K}(\rho)). \quad (20.179)$$

Combining these two equations, we obtain the response of the Green's function for a large momentum.

Since  $m$  is not an invariant quantity under the renormalization group, it is not a physical mass appearing as a pole in the Green's function. In particular,  $K$  is a

parameter which shifts the mass. We thus eliminate  $K$  by a suitable renormalization group transformation. Hence, we use

$$\bar{K}(\rho) = 0 \tag{20.180}$$

to define  $\rho$ . For this  $\rho$ ,  $\bar{m}(\rho) = m e^\rho$  becomes a physical mass. In this case, (20.178) becomes

$$G^{(n)}(p_i; m, g, k) = Z^n G^{(n)}(p_i; \bar{m}, \bar{g}, 0) . \tag{20.181}$$

Multiplying the left-hand side by  $(\tilde{\mathcal{D}} + n\gamma_\varphi)$ , we obtain zero, so the right-hand side should also vanish under this operation:

$$\left[ (\tilde{\mathcal{D}}\bar{m}) \frac{\partial}{\partial \bar{m}} + (\tilde{\mathcal{D}}\bar{g}) \frac{\partial}{\partial \bar{g}} + n\bar{\gamma}_\varphi \right] G^{(n)}(p_i; \bar{m}, \bar{g}, 0) = 0 , \tag{20.182}$$

$$\bar{\gamma}_\varphi = \gamma_\varphi + \tilde{\mathcal{D}}(\ln Z) . \tag{20.183}$$

However, since the homogeneous equation cannot be expressed in terms of just two variables  $\bar{m}$  and  $\bar{g}$ , we deduce that

$$\tilde{\mathcal{D}}\bar{m} = 0 , \quad \tilde{\mathcal{D}}\bar{g} = 0 , \quad \bar{\gamma}_\varphi = 0 . \tag{20.184}$$

Hence, in the physical quantity,  $\bar{m}$  becomes invariant under the renormalization group transformation, and if we take

$$\bar{m}^2 = m^2 F(g, K) , \quad F(g, 0) = 1 , \tag{20.185}$$

then

$$\left[ \beta(g) \frac{\partial}{\partial g} - \left\{ 1 + [2 - \gamma_S(g)] K \right\} \frac{\partial}{\partial K} + 2 \right] F(g, K) = 0 . \tag{20.186}$$

This equation can be used to determine the dependence of the physical mass on  $g$  and  $K$ .

So far we have given the renormalization group equation for  $G$ , but it can sometimes be given for the one-particle irreducible  $\Gamma$  introduced in Sect. 15.3. We shall thus obtain an equation for  $\Gamma$ . We first introduce a generating functional including a mass-insertion term:

$$\mathcal{T}[J] = \left\langle \mathbf{0} \left| T \exp \left\{ -i \int d^4x \left[ J(x)\varphi(x) + Km^2\mathcal{S}(x) \right] \right\} \right| \mathbf{0} \right\rangle , \tag{20.187}$$

$$\mathcal{T}[J] = \exp(-iW[J]) , \quad \Gamma = W - \int d^4x J(x)\langle\varphi(x)\rangle . \tag{20.188}$$



Since  $W[J]$  is a generating functional for  $G$ , (20.168) can be written in the form

$$\left[ \tilde{\mathcal{D}} + \gamma_\varphi(g) \int d^4x J(x) \frac{\delta}{\delta J(x)} \right] W[J] = 0. \quad (20.189)$$

The variation of  $W$  is now

$$\delta W = \frac{\partial W}{\partial m} \Big|_J \delta m + \frac{\partial W}{\partial g} \Big|_J \delta g + \frac{\partial W}{\partial K} \Big|_J \delta K + \int d^4x \frac{\delta W}{\delta J(x)} \delta J(x). \quad (20.190)$$

As we have seen in Sect. 15.3,

$$\frac{\delta W}{\delta J(x)} = \langle \varphi(x) \rangle, \quad \frac{\delta \Gamma}{\delta \langle \varphi(x) \rangle} = -J(x). \quad (20.191)$$

Thus, using (20.188), we compute  $\delta \Gamma$  :

$$\delta \Gamma = \frac{\partial W}{\partial m} \Big|_J \delta m + \frac{\partial W}{\partial g} \Big|_J \delta g + \frac{\partial W}{\partial K} \Big|_J \delta K + \int d^4x \delta \langle \varphi(x) \rangle \frac{\delta \Gamma}{\delta \langle \varphi(x) \rangle}. \quad (20.192)$$

Therefore,

$$\frac{\partial \Gamma}{\partial m} \Big|_{\langle \varphi \rangle} = \frac{\partial W}{\partial m} \Big|_J, \quad \frac{\partial \Gamma}{\partial g} \Big|_{\langle \varphi \rangle} = \frac{\partial W}{\partial g} \Big|_J, \quad \frac{\partial \Gamma}{\partial K} \Big|_{\langle \varphi \rangle} = \frac{\partial W}{\partial K} \Big|_J. \quad (20.193)$$

Additionally,

$$\int d^4x J(x) \frac{\delta W}{\delta J(x)} = - \int d^4x \langle \varphi(x) \rangle \frac{\delta \Gamma}{\delta \langle \varphi(x) \rangle}. \quad (20.194)$$

If we rewrite (20.189) using (20.193) and (20.194), then

$$\left[ \tilde{\mathcal{D}} - \gamma_\varphi(g) \int d^4x \langle \varphi(x) \rangle \frac{\delta}{\delta \langle \varphi(x) \rangle} \right] \Gamma[\langle \varphi(x) \rangle] = 0. \quad (20.195)$$

A renormalization group equation for each function  $\Gamma$  is

$$\left[ \tilde{\mathcal{D}} - n\gamma_\varphi(g) \right] \Gamma^{(n)}(p_1, \dots, p_n; m, g, K) = 0. \quad (20.196)$$

Comparing this equation with (20.168), the coefficient of  $\gamma_\varphi$  has the opposite sign.

So far we have investigated renormalization groups of the first and second kind for the scalar theory and QED. However, it is for non-Abelian gauge theories that the renormalization group shows its real worth. The first reason is this: as mentioned in Chap. 13, it is believed that strong interactions among quarks can

be described by the non-Abelian gauge theory with the gauge group  $SU(3)$ . A second reason is that non-Abelian gauge theories have a crucially different property which other interactions do not possess: at high energies, effective interactions become weak, whence perturbation theory can be applied. This property is called *asymptotic freedom*. Its renormalization group equation facilitates comparison with experiments.

It will turn out that a distinguishing feature of quarks mentioned in Chap. 13 can be explained using the renormalization group equation.

## 20.7 Renormalization Group for Non-Abelian Gauge Theories

Combining (19.10) and (19.66),

$$\frac{\partial}{\partial x_\mu} \langle \mathbf{0} | T[A_\mu^a(x), A_\nu^b(y)] | \mathbf{0} \rangle = a \delta_{ab} \frac{\partial}{\partial x_\nu} D_F(x-y). \quad (20.197)$$

This equation has the same form for both renormalized quantities and unrenormalized quantities. In particular, the longitudinal part of the gauge field is not affected by the renormalization. We write the propagators for the transverse and ghost parts in unrenormalized forms:

$$D_0^{\text{tr}}(k)_{\mu\nu} = \frac{1}{k^2 - i\epsilon} \left( \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2 - i\epsilon} \right) d_0(k^2), \quad (20.198)$$

$$\langle \mathbf{0} | T[\mathbf{c}^a(x), \bar{\mathbf{c}}^b(y)] | \mathbf{0} \rangle = \delta_{ab} \frac{1}{(2\pi)^4} \int d^4k e^{ik \cdot (x-y)} G_0(k^2), \quad (20.199)$$

and we introduce the renormalization constants

$$Z_3 = d_0(\mu^2), \quad \tilde{Z}_3 = G_0(\mu^2). \quad (20.200)$$

We also introduce the renormalization constants for vertex functions:

$$\text{gauge field} \quad \Gamma_0^{abc}(p, q, r)_{\lambda\mu\nu} \Big|_{p^2=q^2=r^2=\mu^2} = Z_1^{-1} \Gamma_{\text{bare}}^{abc}(p, q, r)_{\lambda\mu\nu} \Big|_{p^2=q^2=r^2=\mu^2}, \quad (20.201)$$

$$\text{ghost field} \quad \Gamma_0^{abc}(p, q, r)_\lambda \Big|_{p^2=q^2=r^2=\mu^2} = \tilde{Z}_1^{-1} \Gamma_{\text{bare}}^{abc}(p; q, r)_\lambda \Big|_{p^2=q^2=r^2=\mu^2}. \quad (20.202)$$

The subscript “bare” indicates the vertex function in the Born approximation, where  $g$  has been subtracted from the one given in Sect. 18.4. Taking functional derivatives with respect to  $\langle c \rangle$ ,  $\langle A \rangle$ , and  $\langle B \rangle$  of the Slavnov–Taylor identity [171, 172]

introduced in Sect. 19.5 and setting the external fields equal to zero, we can show that

$$Z_3/Z_1 = \tilde{Z}_3/\tilde{Z}_1 . \quad (20.203)$$

We skip the details of the proof. This equation is called the *Slavnov–Taylor identity*. The renormalizations of the gauge parameter and the coupling constant are

$$\alpha_0 = Z_3\alpha , \quad g_0 = Z_1Z_3^{-3/2}g . \quad (20.204)$$

If we now consider an  $n$ -point Green's function for the gauge field, its renormalization is

$$G_0^{(n)}(p_1, \dots, p_n) = Z_3^{n/2}G^{(n)}(p_1, \dots, p_n) , \quad (20.205)$$

where the vector indices have been omitted. The left-hand side is a function of  $g_0$ ,  $\alpha_0$ , and  $\Lambda$ , and the right-hand side is a function of  $g$ ,  $\alpha$ , and  $\mu$ . Since the left-hand side does not depend on the renormalization point  $\mu$ , we have

$$\mu \frac{\partial}{\partial \mu} G_0^{(n)}(p_1, \dots, p_n) = 0 . \quad (20.206)$$

Using this and referring to the discussion in Sect. 20.4,

$$\left[ \mu \frac{\partial}{\partial \mu} + \beta(g, \alpha) \frac{\partial}{\partial g} + \delta(g, \alpha) \frac{\partial}{\partial \alpha} + n\gamma_V(g, \alpha) \right] G^{(n)}(p_1, \dots, p_n) = 0 , \quad (20.207)$$

$$\begin{aligned} \beta(g, \alpha) &= \mu \frac{\partial g}{\partial \mu} \Big|_{g_0, \alpha_0, \Lambda} , & \gamma_V(g, \alpha) &= \mu \frac{\partial}{\partial \mu} \ln Z_3^{1/2} \Big|_{g_0, \alpha_0, \Lambda} , \\ \delta(g, \alpha) &= \mu \frac{\partial \alpha}{\partial \mu} \Big|_{g_0, \alpha_0, \Lambda} = -2\alpha \gamma_V(g, \alpha) . \end{aligned} \quad (20.208)$$

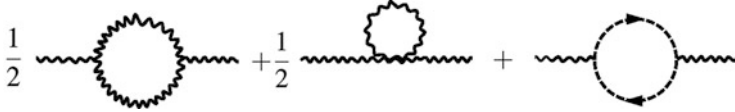
Among the parameters appearing above,  $\mu$  and  $\Lambda$  are the only ones with physical dimensions. Thus,  $g$ ,  $Z_3$ ,  $\alpha$ , and so on must be functions of  $\Lambda/\mu$ , whence

$$\mu \frac{\partial}{\partial \mu} f(\Lambda/\mu) = -\frac{\partial}{\partial \ln \Lambda} f(\Lambda/\mu) .$$

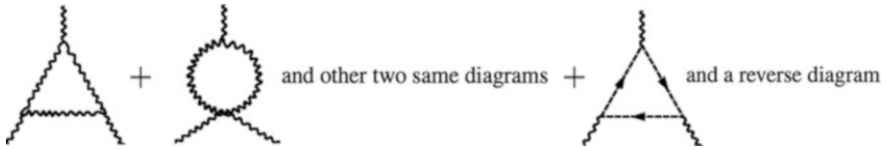
Since the  $Z$  factors become unity to lowest order,

$$\beta = -g \frac{\partial}{\partial \ln \Lambda} \left( \frac{Z_3^{3/2}}{Z_1} \right) , \quad \gamma_V = -\frac{1}{2} \frac{\partial}{\partial \ln \Lambda} Z_3 . \quad (20.209)$$

Therefore, we need to know the coefficients of  $\ln \Lambda$  in  $Z_1$  and  $Z_3$ .



**Fig. 20.1** Feynman diagram for calculating  $Z_3$ . Gluons are represented by *wavy lines* and ghosts by *dashed lines*



**Fig. 20.2** Feynman diagram for calculating  $Z_1$

To compute  $Z_3$ , we consider the Feynman diagram in Fig. 20.1. Note that the wavy line stands for the gluon and the dashed line with the arrow stands for the ghost. The result is

$$Z_3 = 1 + \frac{g^2}{16\pi^2} \left( \frac{13}{3} - \alpha \right) C_2(G) \ln \Lambda , \tag{20.210}$$

where

$$\sum_{c,d} f_{acd} f_{bcd} = \delta_{ab} C_2(G) . \tag{20.211}$$

In particular, for  $SU(N)$ ,  $C_2(G) = N$ . In order to compute  $Z_1$ , we consider the Feynman diagrams in Fig. 20.2. The result is

$$Z_1 = 1 + \frac{g^2}{16\pi^2} \left( \frac{17}{6} - \frac{3}{2}\alpha \right) C_2(G) \ln \Lambda . \tag{20.212}$$

Similarly, the  $Z$  factors for the ghost are

$$\tilde{Z}_3 = 1 + \frac{g^2}{16\pi^2} \left( \frac{3}{2} - \frac{1}{2}\alpha \right) C_2(G) \ln \Lambda , \tag{20.213}$$

$$\tilde{Z}_1 = 1 + \frac{g^2}{16\pi^2} (-\alpha) C_2(G) \ln \Lambda . \tag{20.214}$$

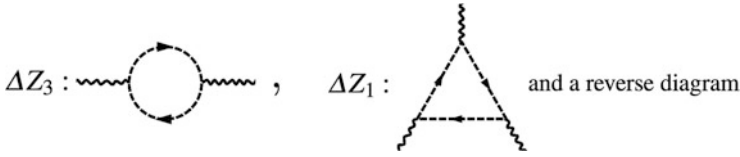


Fig. 20.3 Feynman diagrams for  $\Delta Z_1$  and  $\Delta Z_3$

Thus,

$$\beta(g, \alpha) = -\frac{g^3}{16\pi^2} \left(\frac{11}{3}\right) C_2(G), \tag{20.215}$$

$$\gamma_V(g, \alpha) = -\frac{g^2}{32\pi^2} \left(\frac{13}{3} - \alpha\right) C_2(G). \tag{20.216}$$

If there are fermions, their contributions to  $Z_1$  and  $Z_3$  are

$$\Delta Z_1 = \Delta Z_3 = -\frac{g^2}{8\pi^2} \frac{4}{3} T(R) \ln \Lambda, \tag{20.217}$$

where

$$\text{Tr}(t^a t^b) = \delta_{ab} T(R), \quad r T(R) = d(R) C_2(R), \tag{20.218}$$

with  $r$  the number of creation operators,  $d(R)$  the dimension of the representation, and  $C_2(R)$  as defined in Sect. 13.2. The corresponding Feynman diagrams are shown in Fig. 20.3. For the  $N$ -dimensional fundamental representation of  $SU(N)$ ,

$$T(R) = \frac{1}{2}, \quad C_2(R) = \frac{N^2 - 1}{2N}. \tag{20.219}$$

Therefore, assuming that all fermions obey the fundamental representation and denoting the number of their species, i.e., flavours, by  $n_f$ ,

$$\beta = -\frac{g^3}{16\pi^2} \left[ \frac{11}{3} C_2(G) - \frac{4}{3} n_f T(R) \right], \tag{20.220}$$

$$\gamma_V = -\frac{g^2}{32\pi^2} \left[ \left(\frac{13}{3} - \alpha\right) C_2(G) - \frac{8}{3} n_f T(R) \right]. \tag{20.221}$$

This result was obtained in 1973 by Gross and Wilczek [183] and by Politzer [184].

## 20.8 Asymptotic Freedom

In Sect. 13.8, we discussed the parton model, concluding that interactions among quarks become weak at short range, i.e., the quarks in a hadron are almost free. In this section, we will prove that QCD based on the colour  $SU(3)$  group has this property.

For both the first kind and the second kind, elements of the renormalization group can be written as  $R_\rho$ . If  $\rho$  is taken as a transformation parameter, infinitesimal transformations can be expressed as follows:

$$\delta\mu = \mu\delta\rho, \quad \delta g = \beta(g)\delta\rho. \quad (20.222)$$

In the limits  $\rho \rightarrow \pm\infty$ , we then ask whether  $g$  approaches a constant limiting value  $g_{\pm\infty}$ . If it approaches constant values, then  $g_\infty$  and  $g_{-\infty}$  are called the *ultraviolet fixed point* and the *infrared fixed point*, respectively. In this case,

$$\beta(g_\infty) = 0, \quad \text{or} \quad \beta(g_{-\infty}) = 0. \quad (20.223)$$

In particular, if  $\beta(g) = -bg^3$  ( $b > 0$ ) when  $g$  is small, then  $g = 0$  in the limit  $\rho \rightarrow \infty$ . Such a theory is said to be *asymptotically free*. Examples are non-Abelian gauge theories, the  $g\phi^4$  theory for  $g < 0$ , the six-dimensional  $\phi^3$  theory, and so on.

In a non-Abelian gauge theory, we take  $g = g_0$  when  $\rho = 0$  and solve the equation for small  $g$ :

$$\delta g = -bg^3\delta\rho, \quad b > 0, \quad (20.224)$$

then

$$\frac{1}{\bar{g}^2(\rho)} = \frac{1}{g_0^2} + 2b\rho. \quad (20.225)$$

We can now investigate what can be expected if QCD is asymptotically free.

### 20.8.1 Electron–Positron Collision

An electron–positron collision provides an opportunity to investigate asymptotic freedom. We take  $j_\mu$  as the four-momentum current density for hadrons. Then, based on the computation given in Sect. 9.4, we can use this  $j_\mu$  to define  $\Pi(s)$  from (20.123) and (20.124). If we consider the lowest order for  $e$  and assume that  $s$  is large, we obtain (20.124).

When  $\Pi$  includes leptons, the function  $d$  appearing in the renormalization group for the electromagnetic field can be written as

$$d = 1 - \Pi . \quad (20.226)$$

We note that, in QCD + QED,

$$(\mathcal{D} + 2\gamma_{\text{ph}})d = 0 . \quad (20.227)$$

Applying the Landau gauge for both QCD and QED,

$$\mathcal{D} = \mu \frac{\partial}{\partial \mu} + \beta_e \frac{\partial}{\partial e} + \beta_g \frac{\partial}{\partial g} . \quad (20.228)$$

Note that both  $\Pi$  and  $\gamma_{\text{ph}}$  are of order  $e^2$ . If we drop terms of order  $e^4$ , then  $\gamma_{\text{ph}}\Pi$  and  $\beta\partial/\partial e$  can be omitted. Then, applying this approximation to (20.227) and using (20.226),

$$\left[ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} \right] \Pi = 2\gamma_{\text{ph}} . \quad (20.229)$$

In particular, taking the imaginary part, we obtain

$$\left[ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} \right] \text{Im } \Pi \left( \frac{s}{\mu^2}, \frac{m^2}{\mu^2}, g \right) = 0 . \quad (20.230)$$

To lowest order in QED,

$$\begin{aligned} \text{Im } \Pi &= \frac{1}{3} \sum_i \frac{Q_i^2}{4\pi} \left( 1 + \frac{2m_i^2}{s} \right) \sqrt{1 - \frac{4m_i^2}{s}} \\ &\rightarrow \frac{1}{3} \sum_i \frac{Q_i^2}{4\pi} \quad (\text{fors } \rightarrow \infty) . \end{aligned} \quad (20.231)$$

We introduce the following quantity, called the *Drell ratio*:

$$R(s) = \frac{\sigma(e^+ + e^- \rightarrow \text{hadrons})}{\sigma(e^+ + e^- \rightarrow \mu^+ + \mu^-)} .$$

To lowest order again, for  $s \rightarrow \infty$ , the following scaling law is obtained:

$$R \rightarrow \sum_i Q_i^2 / e^2 . \quad (20.232)$$

Now, at high energies, we set  $m^2/\mu^2 \rightarrow 0$  for  $s \gg \mu^2 \gg m^2$ . Therefore, since  $\Pi$  is dimensionless,

$$\left[ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} \right] \text{Im } \Pi \left( \frac{s}{\mu^2}, g \right) = 0. \quad (20.233)$$

This implies

$$\text{Im } \Pi \left( \frac{s}{\mu^2}, g \right) = \text{Im } \Pi \left( \frac{s_0}{\mu^2}, \bar{g}(\rho) \right), \quad s = s_0 e^{2\rho}. \quad (20.234)$$

If  $\rho$  is large, then in QCD  $\bar{g}\rho \rightarrow 0$ , and

$$\text{Im } \Pi \left( \frac{s_0}{\mu^2}, \bar{g}(\rho) \right) = \text{Im } \Pi \left( \frac{s_0}{\mu^2}, 0 \right) [1 + B \bar{g}^2(\rho) + \dots]. \quad (20.235)$$

We write a solution to (20.225) in the form

$$\bar{g}^2(\rho) = \left( b \ln \frac{s}{\Lambda^2} \right)^{-1}, \quad \Lambda^2 = s_0 \exp \left( -\frac{1}{b g_0^2} \right). \quad (20.236)$$

Moreover, in  $SU(3)$ ,

$$C_2(G) = 3, \quad T(R) = \frac{1}{2}, \quad C_2(R) = \frac{4}{3}, \quad (20.237)$$

such that

$$b = \frac{1}{16\pi^2} (11 - 2n_f), \quad B = \frac{3}{16\pi^2} C_2(R) = \frac{1}{4\pi^2}. \quad (20.238)$$

If we assume that  $s_0/\mu^2 \gg 1$ , then at high energies,

$$\text{Im } \Pi \left( \frac{s}{\mu^2}, g \right) = \frac{1}{12\pi} \sum_i Q_i^2 \left[ 1 + \frac{C_i}{\ln(s/\Lambda^2)} + \dots \right], \quad (20.239)$$

where

$$C_i = \begin{cases} \frac{B}{b} = \frac{12}{33 - 2n_f} & \text{(quark)}, \\ 0 & \text{(lepton)}. \end{cases} \quad (20.240)$$



Thus, assuming that  $\delta_{iq}$  is 1 for quarks and 0 for leptons,

$$R(s) = \sum_i \frac{Q_i^2}{e^2} \left[ 1 + \frac{12}{33 - 2n_f} \frac{\delta_{iq}}{\ln(s/\Lambda^2)} + \dots \right]. \quad (20.241)$$

Hence, for  $s \rightarrow \infty$ ,  $R$  gets closer to a constant. This is called a *scaling relation*. In addition, the total cross-section to hadrons turns out to be inversely proportional to  $s$ . We rewrite the sum over  $i$  to take into account the fact that there are three colours for each flavour:

$$\sum_i Q_i^2 = 3 \sum_f Q_f^2 + \sum_l Q_l^2, \quad (20.242)$$

where indices  $f$  and  $l$  range over flavours and leptons, respectively, and we assume that the correction term in (20.241) applies only to quarks. This scaling relation agrees well with experiment, supporting the idea that there are three colours.

The branching ratio for  $\tau^-$  decay is quite similar. Taking into account the large mass of  $\tau$ , viz., 1784 MeV, the scaling law is considered to apply, so it can decay into each of the following with the same probability:

$$\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau, \quad \tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau, \quad \tau^- \rightarrow d' + \bar{u} + \nu_\tau \quad (3\text{colours}). \quad (20.243)$$

Thus, the branching ratio for  $\tau$  to decay into hadrons is expected to be 3/5. The experimental value is 0.64.

## 20.8.2 Bjorken Scaling Law

We shall now interpret the Bjorken scaling law [105] discussed in Sect. 13.8 in relation to the question of the asymptotic freedom. However, a rigorous argument requires a complicated procedure involving Wilson's operator expansion, so here we shall limit ourselves to a sketch of the basic scenario.

Around  $x^2 = 0$ , current densities can be expanded in the following way:

$$j_\mu \left( \frac{x}{2} \right) j_\nu \left( -\frac{x}{2} \right) \sim \sum_n C^{(n)}(x^2, g) x_{\mu_1} \dots x_{\mu_n} \mathbf{O}_{\mu\nu; \mu_1 \dots \mu_n}^{(n)}. \quad (20.244)$$

This is *Wilson's operator expansion*. Taking the expectation value of this equation in a one-nucleon state and carrying out the Fourier expansion, we obtain a structure function  $F_i(x, q^2)$ . Here,  $\mathbf{O}^{(n)}$  is a local operator that is symmetric with respect to  $\mu_1, \dots, \mu_n$ , and contractions over arbitrary indices have been chosen so that they vanish. Writing  $q^2$  instead of  $Q^2$ , we have  $x = q^2/2Mv$ .

If we choose a large enough value for  $q^2$ , the following equation holds:

$$\int_0^1 dx F_i(x, q^2) x^{n-2} \sim \bar{C}_i^{(n)}(q^2, g) \langle P | \mathbf{O}^{(n)} | P \rangle, \quad (20.245)$$

where  $\bar{C}^{(n)}$  is the Fourier transform of  $C^{(n)}$ .

If we consider that, in a commutation relation such as (12.180),  $e$  is a fixed parameter, the anomalous dimension of  $j_\mu$  becomes zero. Hence, taking the anomalous dimension of  $\mathbf{O}^{(n)}$  to  $\gamma_n$ , the anomalous dimension of  $\bar{C}^{(n)}$  becomes  $-\gamma_n$ :

$$[\mathcal{D} - \gamma_n(g)] \bar{C}_i^{(n)}\left(\frac{q^2}{\mu^2}, g\right) = 0. \quad (20.246)$$

In order to solve this equation, we set  $q^2/\mu^2 = \exp(2\rho)$ :

$$\bar{C}_i^{(n)}\left(\frac{q^2}{\mu^2}, g\right) = \bar{C}_i^{(n)}(1, \bar{g}(\rho)) \exp\left[-\int_0^\rho d\rho \gamma_n(\bar{g}(\rho))\right]. \quad (20.247)$$

In general,  $\gamma_n$  can be written in the form

$$\gamma_n(g) = \gamma_n g^2 + \mathcal{O}(g^4), \quad (20.248)$$

whence the above exponential function can be determined:

$$\exp\left[-\int_0^\rho d\rho \gamma_n(\bar{g}(\rho))\right] = \left(1 + b g_0^2 \ln \frac{q^2}{\mu^2}\right)^{-\gamma_n/2b}. \quad (20.249)$$

Thus, when  $q^2$  is large,

$$\bar{C}_i^{(n)}\left(\frac{q^2}{\mu^2}, g\right) \approx \text{const.} \left(\ln \frac{q^2}{\mu^2}\right)^{-\gamma_n/2b}, \quad (20.250)$$

where we have set  $\bar{g}(\rho)$  to zero in the first factor on the right-hand side of (20.247). Hence, although the  $q^2$ -dependence remains, and this therefore differs from the original Bjorken scaling law, the dependence is at most logarithmic. This is an interpretation of Bjorken's scaling law based on the asymptotically free QCD. As expected, this result gives support for QCD.

## 20.9 Gauge Dependence of Green's Functions

In gauge theories, there are many ways to choose the gauge-fixing term and the ghost term, and it is important to show that no physical consequence depends on these choices. For instance, a charge  $e$  renormalized in the standard way does not depend on the way we choose the Lagrangian. This is clear from the fact that there are ways to measure  $e$  experimentally. Since QED is an infrared-stable theory, we have the so-called *low-energy limit theorem*. At low energies there are rigorous ways to express the  $S$ -matrix element in terms of  $e$ , and this is why  $e$  can be determined experimentally. In contrast, in non-Abelian gauge theories, there is no preferred renormalization prescription like the one in QED. Therefore, it is unclear whether the coupling constant  $g$  is defined independently of the gauge; and depending on the renormalization prescription, it may depend on how the gauge is chosen.

In this section, we consider ways to distinguish parameters depending on the gauge choice from those independent of it. First, we discuss the situation where renormalization is not involved. We thus consider an unrenormalized Green's function:

$$G_0^{(n)} = \langle \varphi^{(0)}(x_1), \dots, \varphi^{(0)}(x_n) \rangle, \quad (20.251)$$

where we have assumed that  $\varphi$  expresses not only a scalar field but also a general field. Now, considering a gauge theory, we choose the Lagrangian (18.74) in Chap. 18. In this case, what distinguishes the Lagrangians is an unrenormalized gauge parameter  $\alpha_0$ . For the Green's function (20.251), we can write down two equations. The first is an equation which has appeared many times up to now and the second is the equation derived from (18.84) in Chap. 18:

$$\mu \frac{\partial}{\partial \mu} \langle \varphi^{(0)}(x_1) \dots \varphi^{(0)}(x_n) \rangle_{\alpha_0} = 0, \quad (20.252)$$

$$\begin{aligned} \alpha_0 \frac{\partial}{\partial \alpha_0} \langle \varphi^{(0)}(x_1) \dots \varphi^{(0)}(x_n) \rangle_{\alpha_0} \\ = \frac{i\alpha_0}{2} \langle \varphi^{(0)}(x_1) \dots \varphi^{(0)}(x_n) \int d^4x \mathbf{B}^{(0)}(x) \cdot \mathbf{B}^{(0)}(x) \rangle_{\alpha_0}. \end{aligned} \quad (20.253)$$

In the second equation, the term which should be called a gauge insertion term or something similar to the mass-insertion term in the Callan–Symanzik equation appears on the right-hand side. One property of this term is its invariance under the renormalization, i.e., referring to (19.66), we have

$$\alpha_0 = Z_3 \alpha, \quad \mathbf{B}^{(0)}(x) = Z_3^{-1/2} \mathbf{B}(x), \quad (20.254)$$

whence

$$\alpha_0 \mathbf{B}^{(0)}(x) \cdot \mathbf{B}^{(0)}(x) = \alpha \mathbf{B}(x) \cdot \mathbf{B}(x) . \quad (20.255)$$

We shall imitate the Callan–Symanzik equation. Setting

$$\mathbf{F}(x) = \mathbf{B}(x) \cdot \mathbf{B}(x) , \quad (20.256)$$

and referring to (20.125) and (20.126), we carry out the Fourier transforms

$$(\mathcal{D} + \gamma_G)G^{(n)}(p_1, \dots, p_n) = 0 , \quad (20.257)$$

$$(\bar{\mathcal{D}} + \bar{\gamma}_G)G^{(n)}(p_1, \dots, p_n) = -\frac{\alpha}{2}F^{(n)}(0; p_1, \dots, p_n) , \quad (20.258)$$

where  $G^{(n)}$  and  $F^{(n)}$  are the renormalized Green's functions:

$$G_0^{(n)} = Z_G G^{(n)} , \quad \alpha_0 F_0^{(n)} = Z_G \alpha F^{(n)} . \quad (20.259)$$

Furthermore,

$$\mathcal{D} = \mu \frac{\partial}{\partial \mu} + \beta \frac{\partial}{\partial g} + \delta \frac{\partial}{\partial \alpha} , \quad \bar{\mathcal{D}} = \bar{\beta} \frac{\partial}{\partial g} + \bar{\delta} \frac{\partial}{\partial \alpha} , \quad (20.260)$$

$$\beta = \mu \frac{\partial}{\partial \mu} g \Big|_{g_0, \alpha_0} , \quad \bar{\beta} = \alpha_0 \frac{\partial}{\partial \alpha_0} g \Big|_{g_0, \mu} , \quad (20.261)$$

$$\delta = \mu \frac{\partial}{\partial \mu} \alpha \Big|_{g_0, \alpha_0} , \quad \bar{\delta} = \alpha_0 \frac{\partial}{\partial \alpha_0} \alpha \Big|_{g_0, \mu} , \quad (20.262)$$

$$\gamma_G = \mu \frac{\partial}{\partial \mu} \ln Z_G \Big|_{g_0, \alpha_0} , \quad \bar{\gamma}_G = \alpha_0 \frac{\partial}{\partial \alpha_0} \ln Z_G \Big|_{g_0, \mu} . \quad (20.263)$$

We introduce the integrability condition

$$\left( \mu \frac{\partial}{\partial \mu} \right) \left( \alpha_0 \frac{\partial}{\partial \alpha_0} \right) = \left( \alpha_0 \frac{\partial}{\partial \alpha_0} \right) \left( \mu \frac{\partial}{\partial \mu} \right) . \quad (20.264)$$

Expressing this in terms of renormalized parameters,

$$\mathcal{D} \bar{\mathcal{D}} = \bar{\mathcal{D}} \mathcal{D} . \quad (20.265)$$

In this theory, there are three renormalized parameters  $m/\mu$ ,  $g$ , and  $\alpha$ . Introducing three independent functions  $\rho$ ,  $\sigma$ , and  $\tau$ , we now implement a variable transformation, giving the following theorem without proof:

**Theorem 20.1** *If the commutation condition (20.265) holds true, then we can choose functions  $\rho$ ,  $\sigma$ , and  $\tau$  of the three variables  $m/\mu$ ,  $g$ , and  $\alpha$  satisfying the conditions*

$$\mathcal{D}\rho = 1, \quad \bar{\mathcal{D}}\rho = 0, \quad \mathcal{D}\sigma = 0, \quad \bar{\mathcal{D}}\sigma = 1, \quad \mathcal{D}\tau = 0, \quad \bar{\mathcal{D}}\tau = 0. \quad (20.266)$$

Thus, if  $m/\mu$ ,  $g$ , and  $\alpha$  can be rewritten in terms of  $\rho$ ,  $\sigma$  and  $\tau$ , then (20.257) and (20.258) become

$$\left(\frac{\partial}{\partial\rho} + \gamma_G\right)G^{(n)} = 0, \quad \left(\frac{\partial}{\partial\sigma} + \bar{\gamma}_G\right)G^{(n)} = -\frac{\alpha}{2}F^{(n)}. \quad (20.267)$$

The integrability condition (20.264) then implies

$$\frac{\partial}{\partial\sigma}\gamma_G = \frac{\partial}{\partial\rho}\bar{\gamma}_G. \quad (20.268)$$

Therefore, there is a function  $\Phi$  satisfying

$$\gamma_G = \frac{\partial}{\partial\rho}\Phi, \quad \bar{\gamma}_G = \frac{\partial}{\partial\sigma}\Phi. \quad (20.269)$$

Hence, setting

$$G^{(n)} = e^{-\Phi}I^{(n)}, \quad F^{(n)} = e^{-\Phi}J^{(n)}, \quad (20.270)$$

we have

$$\frac{\partial}{\partial\rho}I^{(n)} = 0, \quad \frac{\partial}{\partial\sigma}I^{(n)} = -\frac{\alpha}{2}J^{(n)}. \quad (20.271)$$

We are especially interested in the case where  $\varphi(x_1), \dots, \varphi(x_n)$  are all BRS-invariant. In this case,  $F^{(n)}$  vanishes, and so therefore does  $J^{(n)}$ , whence  $I^n$  becomes a function of  $\tau$  alone.

It is clear from the previous discussion that  $\rho$  is related to a scale transformation, and  $\sigma$  is related to a transformation which changes the gauge. Therefore, *we may say that quantities independent of the unrenormalized gauge parameter  $\alpha_0$  do not depend on  $\sigma$  in a renormalized theory.*

In QED, from the early discussion in this section and (20.40),

$$\bar{\mathcal{D}}e = 0, \quad \mathcal{D}(\alpha e^2) = 0. \quad (20.272)$$

Therefore, in QED,  $\rho$  is expected to become a function of  $e$ , and  $\sigma$  a function of  $\alpha e^2$ . In QCD, things are more complicated. This is related to the fact that, because the  $\beta$ -function generally depends on the renormalization prescription, it may be a function, not only of  $g$ , but also of  $\alpha$ . If there is a renormalization prescription satisfying

$$\bar{\mathcal{D}}g = 0, \quad (20.273)$$

then it may turn out that  $g$  has a significance unrelated to the gauge. One way is to renormalize all quantities using (18.84) in the Landau gauge. Hence, assuming that (18.84) also holds true for renormalized quantities, we define Green's functions in the general gauge. Therefore, both  $\beta$  and  $\gamma_V$  turn out not to depend on  $\alpha$ . However,  $\rho$  can in general be interpreted as playing a role.

# Chapter 21

## Theory of Confinement



So far we have given indirect evidence that all hadrons consist of quarks. Moreover, the long series of successes of this model make other possibilities quite hard to consider. At the same time, searches for fractional charges have been made, but none has ever been detected. This has led to the conclusion that isolated quarks cannot be observed for some reason, an idea known as the *quark confinement hypothesis*. This hypothesis has in fact been generalized to the idea that no coloured particle can ever be observed. This is the *colour confinement hypothesis*.

Many proposals have been put forward to explain this hypothesis. One idea endorsed by many is that the quark–anti-quark potential may be proportional to distance at long ranges. This idea is closely related to the lattice gauge theories introduced by Wilson [185]. He formulated a condition for quark confinement in such a way that a quantity, which we call a *Wilson loop*, satisfies the area law. He has shown that this condition is satisfied in a strong coupling approximation. However, within the framework of lattice gauge theories, it is hard to formulate a confinement condition for gluons. So in order to discuss confinement for general coloured particles, the author has considered a mechanism for confinement based on an analogy with QED [186].

Firstly, if we ask whether there are ever particles which are created as a consequence of quantization but never observed, we immediately recall the photons corresponding to the longitudinal and scalar waves in QED. If, when we write down the condition for unitarity of the  $S$ -matrix, both the initial and final states satisfy the Lorenz condition, the contributions of these photons to intermediate states completely annihilate each other, and so never show up. That is, they cannot be observed. This was discussed in Chap. 5 and Sect. 6.4. This is confinement based on the annihilation mechanism for indefinite metrics, the Lorenz condition being used to eliminate negative norm states.

If we apply this kind of mechanism to confinement for particles with colour, it turns out to be inevitable to introduce indefinite metrics and additional conditions. For non-Abelian gauge fields, all the tools are to hand. That is, in non-Abelian

gauge fields, indefinite metrics are introduced via auxiliary fields, and additional conditions are defined using the BRS charges. Consequently, if both the initial and final states belong to the BRS singlet, then contributions from the BRS doublets completely annihilate each other in the unitarity condition, which was shown in Sect. 19.6. It thus turns out that colour confinement holds true if all particle states with colours belong to BRS doublets. We thus investigate under what kind of conditions such a thing could be realized.

## 21.1 Gauge Independence of the Confinement Condition

In the following discussion, we shall treat the question of confinement in a special gauge, and in particular, in the Landau gauge. We shall thus show that the concept of confinement does not depend on the choice of gauge. So if it holds true in one gauge, it will hold true in other gauges as well. For this purpose, we first recall a generalization of the LSZ reduction formula presented in Sect. 11.7. We write (11.161') in the form

$$\int d^4x \Delta(x-y) K_y ST[\varphi(x_1) \dots \varphi(x_n) \varphi(y)] = [\varphi^{\text{in}}(x), ST[\varphi(x_1) \dots \varphi(x_n)]] . \quad (21.1)$$

For each local operator  $\Phi(x)$ , we assume that the following conditions are satisfied:

- (1)  $\Phi(x)$  is a polynomial of field operators and for an arbitrary elementary particle field  $\varphi(x)$ ,

$$[\Phi(x), \varphi(y)] = 0 , \quad (x-y)^2 > 0 . \quad (21.2)$$

- (2)  $\Phi(x)$  has the same quantum number as  $\varphi(x)$  and for a one-particle state  $|p\rangle$  of the field  $\varphi$ ,

$$\langle \mathbf{0} | \Phi(x) | p \rangle = \langle \mathbf{0} | \varphi(x) | p \rangle . \quad (21.3)$$

In this case, (21.1) extends to

$$\int d^4y \Delta(x-y) K_y ST[\varphi(x_1) \dots \varphi(x_n) \Phi(y)] = [\varphi^{\text{in}}(x), ST[\varphi(x_1) \dots \varphi(x_n)]] . \quad (21.4)$$



Here we have considered scalar fields, but this can be extended to arbitrary fields, noting that, for fields obeying Fermi statistics, we have to replace the commutator in (21.2) by an anti-commutator.

Since we do not have fundamental fields for composite particles, we change the second condition as follows:

(2') For a one-particle state of a composite particle,

$$\langle \mathbf{0} | \Phi(x) | \mathbf{p} \rangle = \frac{1}{\sqrt{2p_0 V}} e^{ip \cdot x} . \quad (21.5)$$

However, if the spin is not 0, we multiply by a suitable spin function.

In this case, introducing asymptotic fields  $\Phi^{\text{in}}$  and  $\Phi^{\text{out}}$  for a composite particle,

$$\int d^4y \Delta(x-y) K_y ST[\dots \Phi(y)] = [\Phi^{\text{in}}(x), ST[\dots]] . \quad (21.6)$$

When we compute the  $S$ -matrix element using the LSZ asymptotic condition, field operators appearing in the Green's function are not unique but arbitrary, if the condition above is satisfied. This formulation derives from treatments of composite particles by Haag [187], Zimmermann [188], and the author [189]. For details, the reader is referred to my book *Fields and Particles* [2]. We now return to the question of confinement.

If all particles with colour are confined, then it turns out that all quanta of fundamental fields are confined. Thus, observable quanta will be particles without colours. In QCD, this concerns only hadrons as composite particles. Such hadrons are described by local composite fields. Moreover, if we consider *unrenormalized composite fields*  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ , then as discussed in Sect. 19.4, such fields should be invariant under the BRS transformation. Hence, (19.58) should hold true. In this case, (19.63) also turns out to hold true. Note that, in order to use the LSZ reduction formula for these Green's functions, we should carry out the renormalization in such a way that a normalization condition like (21.5) is satisfied. Note also that, in this case, from (19.63), we have

$$\langle \mathbf{A}(x) \mathbf{A}^\dagger(y) \rangle_{II} = \langle \mathbf{0} | \mathbf{A}(x) \mathbf{A}^\dagger(y) | \mathbf{p} \rangle_I , \quad (21.7)$$

so that the residues at poles corresponding to one-particle states for these two-point functions are equal. Therefore, for the unrenormalized  $\mathbf{A}$ ,

$$\langle \mathbf{0} | \mathbf{A}(x) | \mathbf{p} \rangle_I = \langle \mathbf{0} | \mathbf{A}(x) | \mathbf{p} \rangle_{II} , \quad (21.8)$$

where  $|\mathbf{p}\rangle$  stands for a one-particle state of a hadron with four-momentum  $p$ . Therefore,

$$\begin{aligned} & \frac{\langle \mathbf{A}(x)\mathbf{B}(y)\dots \rangle_I}{\sqrt{2p_0V}\langle \mathbf{0}|\mathbf{A}(0)|\mathbf{p}\rangle_I\sqrt{2q_0V}\langle \mathbf{0}|\mathbf{B}(0)|\mathbf{q}\rangle_I\dots} \\ &= \frac{\langle \mathbf{A}(x)\mathbf{B}(y)\dots \rangle_{II}}{\sqrt{2p_0V}\langle \mathbf{0}|\mathbf{A}(0)|\mathbf{p}\rangle_{II}\sqrt{2q_0V}\langle \mathbf{0}|\mathbf{B}(0)|\mathbf{q}\rangle_{II}\dots}. \end{aligned} \quad (21.9)$$

This is the renormalized Green's function which does not depend on gauge choices. It turns out that the  $S$ -matrix element obtained by applying the LSZ reduction formula to such a Green's function has meaning independently of how the gauge is chosen. Hence, if the  $S$ -matrix corresponding to a reaction among hadrons is unitary in one gauge, it will be unitary in all other gauges. As a consequence, all particles with colour turn out to be confined in any gauge. Thus, in order to discuss the confinement of particles with colour, we only need to show this in one specific gauge.

## 21.2 Sufficient Condition for Colour Confinement

In Sect. 19.6, we discussed the fact that the unitarity condition for the  $S$ -matrix can be expressed in terms of BRS singlet states alone. We thus assume the asymptotic completeness discussed in Chap. 11 only for BRS singlet states. The reason is that it is unclear, due to so-called infrared divergences, whether we can assume the asymptotic condition for BRS doublet states as well, and for singlet states, the existence of the unitary  $S$ -matrix forms the basis of the theory.

We begin with the Lagrangian density (19.1). Using the field equations and equal-time commutation relations, we can write down the following Ward–Takahashi identities:

$$\begin{aligned} & \partial_\lambda \langle \delta \bar{\delta} A_\lambda^a(x), \psi^\alpha(y), \bar{\psi}^\beta(z) \rangle \\ &= ig t_{\alpha\beta}^a \left[ \delta^4(x-y) S_F(y-z, \beta) - \delta^4(x-z) S_F(y-z, \alpha) \right], \end{aligned} \quad (21.10)$$

$$\begin{aligned} & \partial_\lambda \langle \delta \bar{\delta} A_\lambda^a(x), A_\mu^b(y), A_\nu^c(z) \rangle \\ &= ig M_{bc}^a \left[ \delta^4(x-y) S_{F\mu\nu}(y-z, c) - \delta^4(x-z) D_{F\mu\nu}(y-z, b) \right], \end{aligned} \quad (21.11)$$

where  $M_{bc}^a = -if_{abc}$  corresponds to the adjoint representation, and  $S_F$  and  $D_F$  are propagators of a fermion and a gauge field, respectively. It is only in the presence of

BRS invariance that these identities hold true, even if the symmetry corresponding to the gauge group is spontaneously broken.

So considering (21.10), we begin by defining the vertex function  $V_\lambda^a$  by

$$\langle \delta\bar{\delta}A_\lambda^a(x), \psi^\alpha(y), \bar{\psi}^\beta(z) \rangle = \int d^4y' \int d^4z' S_F(y - y', \alpha) V_\lambda^a(y', z'; x)_{\alpha\beta} S_F(z' - z, \beta) . \quad (21.12)$$

We introduce the Fourier transform

$$S_F(x) = \frac{-i}{(2\pi)^4} \int d^4p e^{ip \cdot x} S_F(p) ,$$

$$V_\lambda^a(yz; x)_{\alpha\beta} = \frac{1}{(2\pi)^8} \int d^4p \int d^4q e^{ip \cdot (y-x) + iq \cdot (x-z)} V_\lambda^a(p, q)_{\alpha\beta} . \quad (21.13)$$

Then (21.10) can be written in the form

$$S_F(p, \alpha)(p - q)_\lambda V_\lambda^a(p, q) S_F(q, \beta) = i g t_{\alpha\beta}^a [S_F(p, \alpha) - S_F(q, \beta)] . \quad (21.14)$$

What is important here is the identity

$$\partial_\lambda (\delta\bar{\delta}A_\lambda) = -\bar{\delta}(\partial_\lambda D_\lambda c) = 0 . \quad (21.15)$$

We now investigate whether  $V_\lambda$  has a pole with respect to the variable  $(p - q)^2$ . First, we assume that  $x$  is connected to  $y'$  and  $z'$  via one-particle lines. This particle can be elementary or composite. In this case, there exists a pole with respect to  $(p - q)^2$ . For instance, if a particle is created by a scalar field  $\phi^b$ , such a pole will appear in the Green's function

$$\langle \delta\bar{\delta}A_\lambda^a(x), \phi^b(w) \rangle . \quad (21.16)$$

Applying  $\partial_\lambda$  to this expression, since from (21.15) an equal-time commutator appears, the result is proportional to  $\delta^4(x - w)$ . Thus, its Fourier transformation becomes a constant, which can be written in the form

$$\text{FT} \langle \delta\bar{\delta}A_\lambda^a(x), \phi^b(w) \rangle = \frac{k_\lambda}{k^2} \delta_{ab} C_0 . \quad (21.17)$$

In contrast, if this one-particle state is created by a vector field  $\phi_\mu^b$ , we have

$$\text{FT} \langle \delta\bar{\delta}A_\lambda^a(x), \phi_\mu^b(w) \rangle = \left( \delta_{\lambda\mu} - \frac{k_\lambda k_\mu}{k^2} \right) \delta_{ab} f(k^2) + \frac{k_\lambda k_\mu}{k^2} \delta_{ab} C_1 . \quad (21.18)$$

However, the first term on the right-hand side does not appear in (21.14). This is because the left-hand side of (21.14) is multiplied by  $k_\lambda = (p - q)_\lambda$ . We can thus replace  $V_\lambda$  in (21.14) by

$$\bar{V}_\lambda(p, q) = \frac{(p - q)_\lambda(p - q)_\mu}{(p - q)^2} V_\mu(p, q), \quad (21.19)$$

which is a projection for a spin 0. With this replacement, the first term on the right-hand side of (21.18) vanishes. We can then rewrite (21.14) in terms of  $\bar{V}_\lambda$ :

$$S_F(p, \alpha)(p - q)_\lambda \bar{V}_\lambda^a(p, q) S_F(q, \beta) = i g t_{\alpha\beta}^a [S_F(p, \alpha) - S_F(q, \beta)]. \quad (21.20)$$

In  $\bar{V}_\lambda^a$ , there can still be poles of spin 0. In fact, it is clear from (19.65) and (19.66) that  $\delta A_\lambda = D_\lambda c$  creates a pole of spin 0. If so, we may ask whether  $\delta\bar{\delta} A_\lambda$  creates a pole of spin 0.

According to (19.147),  $D_\lambda c$  includes an asymptotic field  $\partial_\lambda \gamma$  of the spin 0 and massless particle. Hence, the condition that  $V_\lambda$  should not include a pole of spin 0 is

$$\bar{\delta}\gamma^a = 0. \quad (21.21)$$

This condition cannot hold true when a symmetry corresponding to the  $a$ th generator of the gauge group is spontaneously broken. So when  $\bar{V}_\lambda$  does not include a pole of spin 0, we set

$$p - q = \epsilon P, \quad P^2 \neq 0, \quad (21.22)$$

and consider the limit  $\epsilon \rightarrow 0$ . By assumption,  $\bar{V}_\lambda$  does not include poles of either spin 0 or spin 1, and the projection operator is also finite in this limit:

$$\frac{(p - q)_\lambda(p - q)_\mu}{(p - q)^2} = \frac{P_\lambda P_\mu}{P^2}. \quad (21.23)$$

Thus, we have the following finite limit:

$$\lim_{\epsilon \rightarrow 0} \bar{V}_\lambda(p, q) = \bar{V}_\lambda(p, p; P). \quad (21.24)$$

From (21.20), we obtain

$$S_F(p, \alpha) P_\lambda \bar{V}_\lambda(p, p; P) S_F(p, \beta) = i g t_{\alpha\beta}^a \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [S_F(p, \alpha) - S_F(q, \beta)]. \quad (21.25)$$

Although the left-hand side is finite, if the symmetry is broken, we generally have

$$S_F(p, \alpha) - S_F(p, \beta) \neq 0. \quad (21.26)$$

Therefore, the right-hand side diverges and we encounter a contradiction. Thus, it turns out that the condition (21.21) can hold true only if the symmetry is not broken. In this case, if we write

$$S_F(p, \alpha) = S_F(p, \beta) \equiv S_F(p) , \quad (21.27)$$

for components  $\alpha$  and  $\beta$  connected by  $t^a$ , the right-hand side of (21.25) assumes the form

$$igt_{\alpha\beta}^a P_\lambda \frac{\partial}{\partial p_\lambda} S_F(p) . \quad (21.28)$$

If we now assume that quarks belong to BRS singlet states, then (21.25) yields a contradiction. In this case, if we apply the LSZ reduction formula to quarks, the left-hand side implies

$$P_\lambda \langle \mathbf{p}, \alpha | \delta\bar{\delta} \bar{A}_\lambda(0) | \mathbf{p}, \beta \rangle , \quad (21.29)$$

where  $|\mathbf{p}, \alpha\rangle$  and  $|\mathbf{p}, \beta\rangle$  are quark states with four-momentum  $p$ . However,  $\delta\bar{\delta} \bar{A}_\lambda$  is a  $(d, d)$ -type operator. If the quarks belong to BRS singlet states, the matrix element (21.29) should vanish. On the other hand, from the right-hand side, we obtain

$$igt_{\alpha\beta}^a \bar{u}(p) \left[ \frac{\partial}{\partial p_\lambda} S_F^{-1}(p) \right] u(p) \propto igt_{\alpha\beta}^a \bar{u}(p) i\gamma_\lambda u(p) , \quad (21.30)$$

up to a trivial numerical coefficient. This is finite and never vanishes, which is a contradiction.

This means that, if the condition (21.21) is satisfied,  $S_F$  cannot possess a pole corresponding to a BRS singlet state. This in turn means that quarks are confined. Similarly, if we start with (21.11) and assume (21.21), then it can be shown that gluons are also confined. Generalizing this result further, it can be understood that (21.21) is a sufficient condition for all particles with colour to be confined.

Since the condition for confinement has been expressed in terms of the asymptotic field of the auxiliary field, we now investigate the connection with the discussion of the representation in Sect. 19.8. From (19.149),

$$\delta\chi = \gamma . \quad (21.31)$$

Therefore, (21.21) can be expressed in the form

$$\bar{\delta}\delta\chi^a = 0 . \quad (21.32)$$

Hence, using (19.152) which is valid in perturbation theory,

$$\bar{\delta}\delta\chi^a = -i\beta^a, \quad (21.33)$$

so it is clear that confinement does not hold true in perturbation theory. Hence, the quartet representation in Sect. 19.8 does not satisfy the condition for confinement. In contrast, since (21.21) holds in the octet representation given by (19.156), the condition for confinement is satisfied. In this case, as mentioned before,  $\mathbf{c} \times \mathbf{c}$  and  $\bar{\mathbf{c}} \times \bar{\mathbf{c}}$  turn out to have asymptotic fields. Alternatively, since  $\delta\mathbf{c}$  and  $\delta\bar{\mathbf{c}}$  have massless spin-0 asymptotic fields, the following quantity turns out to have a massless spin-0 pole:

$$\langle \delta\mathbf{c}^a(x), \bar{\delta}\bar{\mathbf{c}}^b(y) \rangle = \langle \bar{\delta}\mathbf{c}^a(x), \delta\bar{\mathbf{c}}^b(y) \rangle = -\langle \bar{\mathbf{B}}^a(x), \mathbf{B}^b(y) \rangle. \quad (21.34)$$

Referring to the representation theory for the asymptotic field of the auxiliary field in the Landau gauge, as explained in Sect. 19.9, if we take the representation  $\mathcal{D}_j$ , with  $j$  an integer, the condition for confinement is not satisfied, while if  $j$  is a half integer, the condition is in fact satisfied.

We now apply the above discussion to the grand unified theory (GUT) discussed in Sect. 16.5. The gauge group  $G$  of the grand unified theory, after spontaneous symmetry breaking, is

$$G \rightarrow SU(3)_{\text{colour}} \times U(1)_{\text{em}}. \quad (21.35)$$

As already discussed, confinement is caused by a *strict symmetry*. Since only singlets for  $SU(3)$  can be observed, this means that, among the quanta of gauge fields, only the colourless  $\gamma$ ,  $W$ , and  $Z$  are observed, and if hadrons are colourless bound states of quarks, then they can be observed.

### 21.3 Colour Confinement and Asymptotic Freedom

In the last section, we gave (21.32) as the condition for confinement. This section has been added to describe subsequent progress [190, 191]. To begin with, according to (19.66) and (19.147),  $\mathbf{A}_\mu$  and  $\mathbf{B}$  turn out to have the following asymptotic fields:

$$A_\mu^a(x)^{\text{in}} = \alpha_\mu^a(x) + \partial_\mu\chi^a(x), \quad B^b(y)^{\text{in}} = \beta^b(y), \quad (21.36)$$

where  $\alpha_\mu$  stands for the asymptotic field for a gluon, and  $\chi$  and  $\beta$  are those for massless spin-0 fields, which satisfy (19.151). That is,

$$\langle \chi^a(x), \beta^b(y) \rangle = -\delta_{ab}D_F(x-y). \quad (21.37)$$

Next, we check whether the confinement condition (21.32) is satisfied. Since it is strictly true in QED that

$$i\delta\bar{\delta}\chi = -\beta, \quad (21.38)$$

electrical charge can never be confined.

In QCD, since the computation of  $i\delta\bar{\delta}\chi^a$  is complicated, we consider the lowest order in perturbation theory:

$$i\delta\bar{\delta}\chi^a \approx -\beta^a. \quad (21.39)$$

Higher order corrections alter the coefficient of  $\beta$  on the right-hand side. If this coefficient vanishes in an exact theory, then (21.37) implies that a sufficient condition for confinement is

$$\langle i\delta\bar{\delta}\chi^a(x), \chi^b(y) \rangle = 0. \quad (21.40)$$

We now express the condition including this asymptotic field in terms of the Heisenberg operator. We thus consider a two-point function:

$$\langle i\delta\bar{\delta}A_\mu^a(x), A_\nu^b(y) \rangle. \quad (21.41)$$

The first operator includes  $\partial_\mu(i\delta\bar{\delta}\chi^a)$ , the second includes  $\partial_\nu\chi^b$ , and the left-hand side of (21.40) gives a massless spin-0 pole. We can write the Fourier expression of (21.41), using (21.15):

$$\left( \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2 - i\epsilon} \right) \int dm^2 \frac{\sigma(m^2)}{k^2 + m^2 - i\epsilon} + C \frac{k_\mu k_\nu}{k^2 - i\epsilon}. \quad (21.42)$$

The second term corresponds to (21.40). The coefficient  $C$  appears in

$$\partial_\mu \langle i\delta\bar{\delta}A_\mu^a(x), A_\nu^b(y) \rangle = i\delta_{ab} C \partial_\nu \delta^4(x - y). \quad (21.43)$$

The condition for confinement argued here only concerns the second term in (21.42). Using only the constant  $C$ , the sufficient condition (21.40) for confinement reduces to

$$C = 0. \quad (21.44)$$

Note also that, given (21.15), Eq. (21.43) yields the following equal-time commutation relation:

$$\delta(x_0 - y_0) \langle 0 | i\delta\bar{\delta}A_0^a(x), A_j^b(y) | 0 \rangle = i\delta_{ab} C \partial_j \delta^4(x - y), \quad (21.45)$$

for  $j = 1, 2, 3$ . The left-hand side should be evaluated using renormalized operators. Fortunately, both operators are multiplicatively renormalized:

$$\mathbf{A}_\mu^{(0)} = Z_3^{1/2} \mathbf{A}_\mu, \quad i\delta\bar{\delta}\mathbf{A}_\mu^{(0)} = \tilde{Z}_3^{-1} Z_3^{-1/2} (i\delta\bar{\delta}\bar{\mathbf{A}}_\mu), \quad (21.46)$$

where we have used the notation in Sect. 20.7. Therefore, the two-point function (21.41) and the constant  $C$  satisfy the renormalization group equation in the form

$$(\mathcal{D} - 2\gamma_{\text{FP}})C = 0, \quad (21.47)$$

where  $\gamma_{\text{FP}}$  is the anomalous dimension of the ghost field and  $\mathcal{D}$  is the differential operator appearing in (20.207).

Now recall (19.47). The Ward–Takahashi identities (21.10) and (21.11) used in the last section were in fact derived using this equation. Writing this in the unrenormalized form

$$i\delta\bar{\delta}\mathbf{A}_\nu^{(0)} = \partial_\mu \mathbf{A}_{\mu\nu}^{(0)} + g\partial_\mu (\mathbf{A}_\mu^{(0)} \times \mathbf{A}_\nu^{(0)}) + g\mathbf{J}_\nu^{(0)} \quad (21.48)$$

and inserting in (21.45), we see that the only term contributing to the equal-time commutation relation is the first term involving  $\mathbf{A}_{\mu\nu}^{(0)} = \partial_\mu \mathbf{A}_\nu^{(0)} - \partial_\nu \mathbf{A}_\mu^{(0)}$ . However, if we use renormalized operators, then it turns out that the contribution from the Goto–Imamura–Schwinger term [124, 125] appearing in Sect. 12.8 cannot be neglected. Anyway, inserting the renormalized  $\partial_\mu \mathbf{A}_{\mu\nu}$ , we define a constant  $a$  by

$$\delta(x_0 - y_0) \langle 0 | [\partial_\mu \mathbf{A}_{\mu 0}^a(x), \mathbf{A}_j^b(y)] | 0 \rangle = i\delta_{ab} a \partial_j \delta^4(x - y). \quad (21.49)$$

As is clear from the derivation of (12.223), this  $a$  is in fact equal to  $Z_3^{-1}$ . Additionally, since both sides of (21.49) satisfy the renormalization group equation, which is the same as that of the gluon propagator, we have

$$(\mathcal{D} + 2\gamma_V)a = 0. \quad (21.50)$$

Comparing this with (21.47), we see that  $C$  cannot be equal to  $a$ , whence  $C - a$  survives as the Goto–Imamura–Schwinger term.

In addition, we introduce the gluon propagator in Fourier representation:

$$D_{F_{\mu\nu}}(k) = \left( \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2 - i\epsilon} \right) D(k^2) + \alpha \frac{k_\mu k_\nu}{(k^2 - i\epsilon)^2}, \quad (21.51)$$

$$D(k^2) = \int dm^2 \frac{\rho(m^2)}{k^2 + m^2 - i\epsilon}. \quad (21.52)$$



We thus obtain the following relation corresponding to (12.223):

$$a = Z_3^{-1} = \int dm^2 \rho(m^2) . \quad (21.53)$$

Next, we investigate how we can obtain  $C$  by solving (21.47). In Sect. 20.6, we used the integral method for the homogeneous Callan–Symanzik equation. Here, we can apply exactly the same method to the renormalization group equation above. Corresponding to (20.175), the renormalization group operator is given by

$$R_\rho = \exp(\rho \mathcal{D}) . \quad (21.54)$$

Taking  $Q$  as a function of  $g$ ,  $\alpha$ , and  $\mu$ , we define an application of the renormalization group to  $Q$  by

$$\bar{Q}(\rho) = \exp(\rho \mathcal{D}) Q , \quad \bar{Q}(0) = Q . \quad (21.55)$$

In particular, for an infinitesimal transformation,

$$\delta \bar{Q}(\rho) = \delta \rho \mathcal{D} \bar{Q}(\rho) . \quad (21.56)$$

In particular, for  $g$ ,  $\alpha$ , and  $\mu$ , we obtain

$$\delta \bar{g} = \beta(\bar{g}, \bar{\mu}) \delta \rho , \quad \delta \bar{\alpha} = -2\bar{\alpha} \gamma_V(\bar{g}, \bar{\alpha}, \bar{\mu}) \delta \rho , \quad \delta \bar{\mu} = \bar{\mu} \delta \rho , \quad (21.57)$$

where the variable  $(\rho)$  has been omitted. Their asymptotic values are defined by

$$\bar{g}(\infty) = g_\infty , \quad \bar{\alpha}(\infty) = \alpha_\infty , \quad \bar{\mu}(\infty) = \infty , \quad (21.58)$$

and in particular asymptotic freedom is characterized by  $g_\infty = 0$ .

For a function  $Q(g, \alpha, \mu)$ , we clearly have

$$\bar{Q}(\rho) = \exp(\rho \mathcal{D}) Q(g, \alpha, \mu) = Q(\bar{g}(\rho), \bar{\alpha}(\rho), \bar{\mu}(\rho)) . \quad (21.59)$$

In particular, choosing a Green's function  $G(p_i; g, \alpha, \mu)$  as  $Q$  and taking its anomalous dimension to be  $\gamma$ , we have

$$(\mathcal{D} + \gamma)G = 0 . \quad (21.60)$$

In order to integrate this equation, we consider

$$\bar{G}(\rho) = \exp(\rho \mathcal{D}) G , \quad \bar{G}(0) = G , \quad (21.61)$$

whence

$$\frac{\partial}{\partial \rho} \bar{G}(\rho) = \exp(\rho \mathcal{D}) \mathcal{D} G = -\exp(\rho \mathcal{D}) \gamma G = -\bar{\gamma}(\rho) \bar{G}(\rho) . \quad (21.62)$$

Integrating this,

$$G(p_i; g, \alpha, \mu) = \exp \left[ \int_0^\rho d\rho \bar{\gamma}(\rho) \right] G(p_i; \bar{g}(\rho), \bar{\alpha}(\rho), \bar{\mu}(\rho)) . \quad (21.63)$$

This corresponds to (20.178) in the case of the Callan–Symanzik equation, so if we introduce

$$R(k^2) = k^2 D(k^2) , \quad (21.64)$$

this function satisfies the renormalization condition

$$R(\mu^2) = 1 . \quad (21.65)$$

Since the anomalous dimension of  $R$  is  $2\gamma_V$ , we can use (21.63) to obtain

$$R(k^2; g, \alpha, \mu) = \exp \left[ 2 \int_0^\rho d\rho \bar{\gamma}_V(\rho) \right] R(k^2; \bar{g}(\rho), \bar{\alpha}(\rho), \bar{\mu}(\rho)) , \quad (21.66)$$

and if we take  $k^2 = \bar{\mu}^2(\rho)$  in this equation, then (21.65) implies

$$R(\bar{\mu}^2(\rho); g, \alpha, \mu) = \exp \left[ 2 \int_0^\rho d\rho \bar{\gamma}_V(\rho) \right] . \quad (21.67)$$

If we now take the limits  $\rho \rightarrow \infty$  and  $\bar{\mu}^2(\rho) \rightarrow \infty$ , then (21.52) and (21.64) yield

$$\int dm^2 \rho(m^2) = \exp \left[ 2 \int_0^\infty d\rho \bar{\gamma}_V(\rho) \right] . \quad (21.68)$$

From (21.53), this turns out to be equal to  $a$ . We can thus evaluate  $a$  using this equation. Once  $a$  is known, the next problem is to evaluate  $C$ . Although  $C$  satisfies (21.47), we should give its boundary condition. Note that, in a theory in which a cut-off is introduced, taking a renormalization point at a high energy by considering  $\rho \rightarrow \infty$  and  $\bar{\mu}(\rho) \rightarrow \infty$ , we find that  $\bar{G}(\rho)$  approaches the Green's function for a free field. We may thus consider that the Goto–Imamura–Schwinger term vanishes in this limit. Hence, we introduce the following boundary condition for  $C$ :

$$\lim_{\rho \rightarrow \infty} [\bar{C}(\rho) - \bar{a}(\rho)] = 0 . \quad (21.69)$$

Then, given (21.47) and (21.50) written in the form (21.62), and omitting the variable  $(\rho)$ , we obtain

$$\frac{\partial}{\partial \rho} \bar{C} = 2\bar{\gamma}_{\text{FP}} \bar{C}, \quad \frac{\partial}{\partial \rho} \bar{a} = -2\bar{\gamma}_V \bar{a}. \quad (21.70)$$

These lead immediately to

$$\left( \frac{\partial}{\partial \rho} - 2\bar{\gamma}_{\text{FP}} \right) (\bar{C} - \bar{a}) = 2(\bar{\gamma}_V + \bar{\gamma}_{\text{FP}}) \bar{a}. \quad (21.71)$$

If we solve this equation under the boundary condition (21.69), we find

$$\bar{C}(\rho) = \bar{a}(\rho) - \int_{\rho}^{\infty} d\rho' 2[\bar{\gamma}_V(\rho') + \bar{\gamma}_{\text{FP}}(\rho')] \bar{a}(\rho') \exp \left[ -2 \int_{\rho}^{\rho'} d\rho'' \bar{\gamma}_{\text{FP}}(\rho'') \right]. \quad (21.72)$$

In particular, if we set  $\rho = 0$ , then

$$C = a - \int_0^{\infty} d\rho' 2[\bar{\gamma}_V(\rho') + \bar{\gamma}_{\text{FP}}(\rho')] \bar{a}(\rho') \exp \left[ -2 \int_0^{\rho'} d\rho'' \bar{\gamma}_{\text{FP}}(\rho'') \right]. \quad (21.73)$$

Since the most important assumption required to obtain this result is (21.69), let us see whether this assumption actually holds true in QED.

### Verification in QED

An easy computation in QED yields the value of  $C$  to be exactly given by

$$C = 1. \quad (21.74)$$

Let us therefore check whether this result can be realized using (21.73), obtained by assuming the Goto–Imamura–Schwinger term (21.69). In QED, we can use the simplification

$$\gamma_{\text{FP}} = 0, \quad \beta(e) = e\gamma_V(e). \quad (21.75)$$

Therefore,

$$\int_0^{\infty} d\rho \bar{\gamma}_V(\rho) = \int_e^{e_{\infty}} \frac{d\bar{e}}{\beta(\bar{e})} \frac{\beta(\bar{e})}{e} = \ln \frac{e_{\infty}}{e}. \quad (21.76)$$

Then  $a(e)$  can be obtained immediately from (21.68):

$$a(e) = (e_\infty/e)^2 . \quad (21.77)$$

Inserting this result into (21.73), we see that (21.74) is indeed realized, so from now on we can also apply (21.69) with confidence in QCD. That is,

$$C = \left(\frac{e_\infty}{e}\right)^2 - 2 \int_e^{e_\infty} \frac{d\bar{e}}{\bar{e}} \left(\frac{e_\infty}{\bar{e}}\right)^2 = 1 . \quad (21.78)$$

For two BRS-invariant Lagrangian densities  $\mathcal{L}_1$  and  $\mathcal{L}_2$ , if there is  $\mathcal{M}$  such that

$$\delta\mathcal{L}_1 = \delta\mathcal{L}_2 = 0 , \quad \mathcal{L}_1 - \mathcal{L}_2 = \delta\mathcal{M} , \quad (21.79)$$

then we say that  $\mathcal{L}_1$  and  $\mathcal{L}_2$  belong to the same gauge class. In this case, as discussed in Sect. 19.4, it is obvious from (19.63) that Green's functions including only BRS-invariant operators will be equal for both Lagrangians. For instance, since for two Lagrangians like those given in (19.1) with different values of  $\alpha$ ,

$$\mathcal{L}_2 - \mathcal{L}_1 = \delta\mathcal{M} , \quad \mathcal{M} = -\frac{i}{2}(\alpha_2 - \alpha_1)\bar{c} \cdot B , \quad (21.80)$$

such Lagrangians will belong to the same gauge class. Hence, Green's functions such as those in (19.63) turn out to be independent of  $\alpha$ . If we consider the renormalization group equation (21.60) for such Green's functions, then  $\gamma$  does not of course depend on  $\alpha$ . Thus,  $\beta$  does not depend on  $\alpha$  either, and becomes constant in the same gauge class. This means that the notion of asymptotic freedom becomes a common notion over the same gauge class, i.e., it turns out that there exists a renormalization method such that  $\beta$  does not depend on  $\alpha$ . Now integrating (21.57),

$$\ln \frac{\alpha}{\alpha_\infty} = 2 \int_0^\infty d\rho \bar{\gamma}_V(\rho) . \quad (21.81)$$

Alternatively, from (21.53), (21.31), and so on,

$$a = \int dm^2 \rho(m^2) = \frac{\alpha}{\alpha_\infty} . \quad (21.82)$$

Thus, the problem of obtaining  $\alpha$  reduces to the problem of obtaining  $\alpha_\infty$ . We need therefore to solve (21.57), but  $\gamma_V$  and  $\beta$  are known only in perturbation theory, i.e., as power series expansions of  $g$  :

$$\beta(g) = g^3(\beta_0 + \beta_1 g^2 + \dots) , \quad \gamma_V(g, \alpha) = g^2(\gamma_0 + \gamma_1 g^2 + \dots) , \quad (21.83)$$

where

$$\gamma_0 = \gamma_{00} + \gamma_{01}\alpha, \quad \gamma_1 = \gamma_{10} + \gamma_{11}\alpha + \gamma_{12}\alpha^2, \quad \dots \quad (21.84)$$

In the following, we assume  $\beta_0 < 0$  as the condition for asymptotic freedom, and as a consequence,  $g_\infty = 0$ . Moreover, we assume that  $\gamma_{01} > 0$ , which always holds true for QCD.

If  $g^2 \ll 1$ , then for sufficiently large  $\rho$ ,

$$\bar{g}^2(\rho) = g^2(\mathbf{1} + bg^2\rho)^{-1}, \quad b = -2\beta_0 > 0. \quad (21.85)$$

We now consider the integral (21.81):

1. **Convergent case.** If the integral (21.81) converges, since the left-hand side is finite,  $\alpha_\infty$  is not 0, but finite. In this case, for large  $\rho$ , the asymptotic form of the integrand on the right-hand side is

$$\begin{aligned} \bar{\gamma}_V(\rho) &= \bar{g}^2(\rho)(\gamma_{00} + \gamma_{01}\alpha_\infty) + \mathcal{O}\left(\frac{\mathbf{1}}{\rho^2}\right) \\ &= \frac{1}{b\rho}(\gamma_{00} + \gamma_{01}\alpha_\infty) + \mathcal{O}\left(\frac{\mathbf{1}}{\rho^2}\right). \end{aligned} \quad (21.86)$$

In order for this integral to converge for large  $\rho$ , we must have

$$\gamma_{00} + \gamma_{01}\alpha_\infty = 0, \quad \text{or} \quad \alpha_\infty = -\gamma_{00}/\gamma_{01} \equiv \alpha_0. \quad (21.87)$$

2. **Divergent case.** If the integral (21.81) diverges, then the left-hand side is divergent as well, so the possible values of  $\alpha_\infty$  are

$$\alpha_\infty = 0 \quad \text{or} \quad \infty. \quad (21.88)$$

As a consequence, there are only three possible values of  $\alpha_\infty$ , viz.,  $\alpha_0$ , 0, or  $\infty$ .

We now investigate what happens to  $(\bar{\alpha}(\rho), \bar{g}(\rho))$  for large  $\rho$  if we start out with the initial values  $(\alpha, g)$ . Since  $g_\infty = 0$ , for sufficiently large  $\rho$ , we have  $\bar{g}^2(\rho) \ll 1$ . We thus follow this flow by starting with a sufficiently small  $\bar{g}^2(\rho)$ . Apart from the case where  $\alpha_\infty = \infty$ , the lowest order equation for  $\bar{\alpha}$  can be used, and in this case,

$$\frac{d}{d\rho} \approx -2\bar{\alpha}(\gamma_{00} + \gamma_{01}\bar{\alpha})\bar{g}^2 = -2\gamma_{01}\bar{g}^2\bar{\alpha}(\bar{\alpha} - \alpha_0), \quad \gamma_{01}\bar{g}^2 > 0. \quad (21.89)$$

Considering this equation in the limit  $\rho \rightarrow \infty$ , when  $\bar{\alpha}$  is close to 0 or  $\alpha_0$ , it will turn out that  $\alpha_\infty$  is 0,  $\alpha_0$ , or neither of them. If it is neither of them, then it turns

out that  $\alpha_\infty = \infty$  or  $-\infty$ . Since this problem is straightforward, we give only the results:

1.  $\gamma_{00} < 0$  ( $\alpha_0 > 0$ )

$$\alpha_\infty = \begin{cases} \alpha_0, & \alpha > 0, \\ 0, & \alpha = 0, \\ -\infty, & \alpha < 0. \end{cases} \quad (21.90)$$

2.  $\gamma_{00} > 0$  ( $\alpha_0 < 0$ )

$$\alpha_\infty = \begin{cases} 0, & \alpha > \alpha_0, \\ \alpha_0, & \alpha = \alpha_0, \\ -\infty, & \alpha < \alpha_0. \end{cases} \quad (21.91)$$

Here, we began by assuming that  $g^2$  is sufficiently small at the outset. There exists a region where  $\alpha_\infty = -\infty$  regardless of the sign of  $\gamma_{00}$ . According to (21.82), in this region,  $\alpha = 0$ , whence (21.73) implies  $C = 0$  in the same region. Since  $C = 0$  is a sufficient condition for confinement, if  $g$  and  $\alpha$  belong to this region, then there is confinement. In this case, there are other regions where  $C \neq 0$ , but since  $C = 0$  is not a necessary condition but a sufficient condition, then there is no contradiction if there is confinement. After all, going back to the origin of this proof, it is concluded that colour confinement occurs if there is asymptotic freedom and a spontaneously unbroken non-Abelian gauge symmetry.

## Chapter 22

# Anomalous Terms and Dispersion Theory



Field theory has suffered from the problem of divergences since the day of its birth. At first glance, renormalization theory gives a method for removing these divergences, but in this method we have to compute quantities of the type  $\infty - \infty$ , so some indefiniteness is often left in the result. This is typically exemplified by the self-energy of the photon. Depending on how we compute this quantity, it can be zero or infinity. This is the problem of indefiniteness: depending on how we compute, something that should be gauge invariant turns out not to be invariant. A classic example of this is the problem of vacuum polarization. In Chap. 12, we obtained (12.127), and in order to reach a gauge invariant result, we had to set

$$\int_0^1 dx \int d^4 p \frac{\frac{1}{2} p^2 + m^2 + x(1-x)k^2}{[p^2 + m^2 + x(1-x)k^2 - i\epsilon]^2} = 0. \quad (22.1)$$

Fukuda and Miyamoto discovered similar things when they computed the two-photon decay of the neutral meson including the vacuum polarization [192]. Steinberger also carried out a similar computation in [193].

In fact, it was not noticed that there were in fact two kinds of singularities, which had become a matter of course around that time: one was the indefiniteness mentioned above, and the other was the anomalous term. In 1951, although Schwinger had proposed a computational method that did not spoil the gauge invariance [194], the problem of anomalous terms had gone unnoticed. So what exactly is this problem? Since a product of operators located at the same space-time point is a highly singular quantity in field theory, it is difficult to define it correctly.

The equation for interacting fields is non-linear, so there is no way to get around this issue. In a classical theory, it is easy to define a product of  $c$ -numbers, but in a quantum theory, as a consequence of the quantization, a new kind of singularity is generated. Thus, in a quantum theory, a result derived from a classical field equation does not necessarily hold. Such a term arising from the difference with a classical

equation is called an *anomalous term*. The issue of anomalous terms is even now a central problem in field theory. It is being investigated in detail using modern mathematical methods such as differential geometry and topology. In this chapter, we will treat this issue from a more elementary point of view.

It was around the end of the 1960s that the difference between indefiniteness and anomalous terms was at last clearly acknowledged. This was a consequence of research by Adler et al. [159, 160]. In this chapter, we will discuss this issue from the point of view of dispersion theory.

As discussed in Sect. 1.4, in dispersion theory, the computation of the  $S$ -matrix amplitude is separated into two steps. In the first step, we compute the absorption part of the amplitude from the unitarity condition, and in the second step, we use the scattering formula to compute the dispersion part from the absorption part. In the computations of the first step, the integration is always carried out in a finite phase space, so it converges and there is neither indefiniteness nor divergence. Hence, if divergence arises, it will appear in the second step. In this case, when appropriate, we introduce the dispersion formula with a subtraction. This subtraction corresponds to a renormalization prescription, but since this computation is not of the form  $\infty - \infty$ , hence different from the Feynman–Dyson theory, no indefiniteness is left at all. However, a weak point in the dispersion theory of the  $S$ -matrix is that a complete system for the dispersion formula is not known. However, what we can learn from the above warning is that divergence and indefiniteness in field theory always arise from the dispersion part alone. Computations of the absorption part can be determined uniquely no matter what method we adopt.

In this chapter, we thus introduce the dispersion theory, not for the  $S$ -matrix, but for the Green's functions. The advantage is that a complete system is known for the dispersion formula. Using this method, we investigate the mechanism leading to the appearance of anomalous terms.

## 22.1 Examples of Indefiniteness and Anomalous Terms

As a first example of indefiniteness, we discuss the issue of vacuum polarization.

### 22.1.1 Vacuum Polarization

The integral representation of the propagator in Sect. 12.8 is an example of dispersion relation. If we consider a scalar field, then the absorption part, or the imaginary part, is given by

$$\frac{1}{2} \langle \mathbf{0} | \{ \varphi(x), \varphi(y) \} | \mathbf{0} \rangle = \frac{1}{(2\pi)^4} \int d^4 p e^{ip \cdot (x-y)} \text{Im} \Delta'_F(-p^2) . \quad (22.2)$$



Similarly,

$$\frac{1}{2}\langle\mathbf{0}|\{\mathbf{j}_\mu(x), \mathbf{j}_\nu(y)\}|\mathbf{0}\rangle = \frac{1}{(2\pi)^4} \int d^4k e^{ik\cdot(x-y)} \text{Im} \Pi_{\mu\nu}(k) . \quad (22.3)$$

Inserting an electron–positron pair state into the intermediate state on the left-hand side, we compute to lowest order in perturbation theory:

$$\langle\mathbf{0}|\{\mathbf{j}_\mu(x), \mathbf{j}_\nu(y)\}|\mathbf{0}\rangle = e^2 \text{Tr}[\gamma_\mu S^{(+)}(x-y) \gamma_\nu \tilde{S}^{(+)}(x-y)^T] + (x \rightleftharpoons y, \mu \rightleftharpoons \nu) , \quad (22.4)$$

where

$$S^{(+)}(x) = (\gamma \cdot \partial - m) \Delta^{(+)}(x) , \quad \tilde{S}^{(+)}(x) = (\gamma^T \cdot \partial + m) \Delta^{(+)}(x) . \quad (22.5)$$

Inserting this and taking the Fourier transform,

$$\begin{aligned} \text{Tr}[\dots] &= -\frac{1}{(2\pi)^6} \frac{\pi}{3} \int_{4m^2}^{\infty} d\kappa^2 \sqrt{1 - \frac{4m^2}{\kappa^2}} \left( 2 + \frac{4m^2}{\kappa^2} \right) \\ &\quad \times \int d^4k e^{ik\cdot(x-y)} (\delta_{\mu\nu} k^2 - k_\mu k_\nu) \theta(k_0) \delta(k^2 + \kappa^2) . \end{aligned} \quad (22.6)$$

Therefore,

$$\text{Im} \Pi_{\mu\nu}(k) = \frac{e^2}{3(2\pi)^2} (k_\mu k_\nu - \delta_{\mu\nu} k^2) \int_{4m^2}^{\infty} d\kappa^2 \sqrt{1 - \frac{4m^2}{\kappa^2}} \left( 1 + \frac{2m^2}{\kappa^2} \right) \pi \delta(k^2 + \kappa^2) . \quad (22.7)$$

If we set

$$\Pi_{\mu\nu}(k) = (k_\mu k_\nu - \delta_{\mu\nu} k^2) \Pi(-k^2) , \quad \text{Im} \Pi(-k^2) = \pi \sigma(-k^2) , \quad (22.8)$$

then

$$\sigma(k^2) = \frac{e^2}{12\pi^2} \sqrt{1 - \frac{4m^2}{k^2}} \left( 1 + \frac{2m^2}{k^2} \right) \theta(k^2 - 4m^2) . \quad (22.9)$$

We now compute  $\Pi(-k^2)$  using the dispersion relation. Using the renormalization condition  $\Pi(0) = 0$  and the dispersion relation in which one subtraction has been carried out,

$$\Pi(-k^2) - \Pi(0) = \int d\kappa^2 \left( \frac{1}{k^2 + \kappa^2} - \frac{1}{\kappa^2} \right) \sigma(\kappa^2) , \quad (22.10)$$

or

$$\Pi(-k^2) = -k^2 \int \frac{d\kappa^2}{\kappa^2} \frac{\sigma(\kappa^2)}{k^2 + \kappa^2 - i\epsilon}. \quad (22.11)$$

As a result, the vacuum polarization has been computed without any indefiniteness. Inserting (22.11) into (22.8), one obtains the Fourier transform of the following two-point function:

$$\langle \mathbf{0} | T^* [j_\mu(x), j_\nu(y)] | \mathbf{0} \rangle. \quad (22.12)$$

We can consider the covariant product  $T^*$  to be defined by the dispersion formula. In the following dispersion theory, we use the  $T$ -product in this sense. Then if we compute (22.1) following Feynman and Dyson, we obtain zero for the absorption part. If we differentiate (22.1) twice with respect to  $k^2$ , it vanishes. The expression in (22.1) does not have branch points for  $k^2$ , so its absorption part vanishes. Although the indefiniteness or the divergence appears only in the dispersion part, the dispersion part can be determined uniquely from the absorption part if we compute using the dispersion relation (22.11).

### 22.1.2 Goto–Imamura–Schwinger Term

The Goto–Imamura–Schwinger term [124, 125] discussed in Sect. 12.8 is an example of something that does not arise when we use equal-time commutation relations. It is also an example of a singularity of field operators evaluated at a common space-time point. The Schwinger term also plays an important role in string theory.

We now go back to the two-photon decay of the neutral meson mentioned in the introduction to this chapter.

### 22.1.3 Triangle Anomaly Term

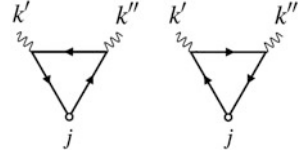
We begin with the Dirac equations in QED:

$$(\gamma \cdot \partial + m)\psi = ie\gamma_\lambda A_\lambda \psi, \quad \bar{\psi}(\gamma \cdot \overleftarrow{\partial} - m) = -ie\bar{\psi}\gamma_\lambda A_\lambda. \quad (22.13)$$

Combining these two equations, we obtain

$$\partial_\lambda j_\lambda^5 = 2m j^5, \quad (22.14)$$

**Fig. 22.1** Feynman diagrams for the calculation of the triangle anomaly term



where

$$j_\lambda^5 = i\bar{\psi}\gamma_\lambda\gamma_5\psi, \quad j^5 = i\bar{\psi}\gamma_5\psi. \tag{22.15}$$

This equation is obtained using the classical equations of motion. In order to find out whether it is also true in quantum theory, we choose a particular matrix element and compare both sides. Hence, sandwiching this equation between the vacuum  $|\mathbf{0}\rangle$  and the two-photon state  $|\mathbf{k}', \mathbf{k}''\rangle$ , we compare both sides. Figure 22.1 shows the two triangle diagrams that arise at the lowest order.

We define the form factors  $F(s)$  and  $G(s)$  by

$$\begin{aligned} \langle \mathbf{k}', \mathbf{k}'' \text{ out} | j_\lambda^5(0) | \mathbf{0} \rangle &= \frac{1}{\sqrt{2k'_0V}\sqrt{2k''_0V}} i\epsilon_{\alpha\beta\gamma\delta} e'_\alpha e''_\beta k'_\gamma k''_\delta F(s), \\ \langle \mathbf{k}', \mathbf{k}'' \text{ out} | j_\lambda^5(0) | \mathbf{0} \rangle &= \frac{1}{\sqrt{2k'_0V}\sqrt{2k''_0V}} i\epsilon_{\alpha\beta\gamma\delta} e'_\alpha e''_\beta k'_\gamma k''_\delta (-i)(k'_\lambda + k''_\lambda) G(s), \end{aligned} \tag{22.16}$$

where  $s$  is the invariant mass squared in the two-photon system, defined by

$$s = -(k' + k'')^2, \tag{22.17}$$

and  $e'$  and  $e''$  are the polarization vectors of each photon.

The absorption or imaginary parts of these form factors can be computed uniquely if we use either the unitarity condition or Feynman and Dyson's method. The controversial relation (22.14) is expressed using the form factors:

$$sG(s) = 2mF(s). \tag{22.18}$$

We first compute the imaginary parts. According to Kikukawa, taking  $\alpha = e^2/4\pi$ ,

$$\text{Im } F(s) = \frac{2\alpha m}{s} \ln \left| \frac{1 + \sqrt{1 - 4m^2/s}}{1 - \sqrt{1 - 4m^2/s}} \right| \theta(s - 4m^2), \tag{22.19}$$

$$\text{Im } G(s) = \frac{4\alpha m}{s} \ln \left| \frac{1 + \sqrt{1 - 4m^2/s}}{1 - \sqrt{1 - 4m^2/s}} \right| \theta(s - 4m^2). \tag{22.20}$$

This implies

$$s \operatorname{Im} G(s) = 2m \operatorname{Im} F(s) , \quad (22.21)$$

and (22.18) certainly holds true for the imaginary parts. Regarding the dispersion parts, i.e., the real parts, we compute using the dispersion relations

$$\operatorname{Re} F(s) = \frac{P}{\pi} \int_{4m^2}^{\infty} ds' \frac{\operatorname{Im} F(s')}{s' - s} , \quad \operatorname{Re} G(s) = \frac{P}{\pi} \int_{4m^2}^{\infty} ds' \frac{\operatorname{Im} G(s')}{s' - s} , \quad (22.22)$$

where P stands for the Cauchy principal value. Going back to (22.18), we have

$$\begin{aligned} s \operatorname{Re} G(s) &= \frac{s}{\pi} \int ds' \frac{\operatorname{Im} G(s')}{s' - s} \\ &= \frac{1}{\pi} \int ds' \frac{s' \operatorname{Im} G(s')}{s' - s} - \frac{1}{\pi} \int ds' \operatorname{Im} G(s') \\ &= 2m \operatorname{Re} F(s) - \frac{1}{\pi} \int ds' \operatorname{Im} G(s') . \end{aligned} \quad (22.23)$$

If (22.18) is to hold true, the integral in the second term on the right-hand side must vanish. In fact, the computation tells us that

$$\int_{4m^2}^{\infty} ds' \operatorname{Im} G(s') = 2\alpha . \quad (22.24)$$

This is the anomalous term. If we compute the real parts from (22.22), we find

$$\operatorname{Re} F(s) = \frac{\alpha}{\pi} \frac{m}{s} \left[ \pi^2 - \left( \ln \left| \frac{1 + \sqrt{1 - 4m^2/s}}{1 - \sqrt{1 - 4m^2/s}} \right| \right)^2 \right] , \quad (22.25)$$

$$\operatorname{Re} G(s) = \frac{2\alpha}{\pi} \frac{m^2}{s} \left\{ -\frac{1}{m^2} + \frac{1}{s} \left[ \pi^2 - \left( \ln \left| \frac{1 + \sqrt{1 - 4m^2/s}}{1 - \sqrt{1 - 4m^2/s}} \right| \right)^2 \right] \right\} . \quad (22.26)$$

Introducing  $B(x) = E(x) \cdot H(x)$ , to lowest order,

$$\langle \mathbf{k}', \mathbf{k}'' \text{ out} | \mathbf{B}(0) | \mathbf{0} \rangle = \frac{1}{\sqrt{2k'_0 V} \sqrt{2k''_0 V}} i \epsilon_{\alpha\beta\gamma\delta} e'_{\alpha} e''_{\beta} k'_{\gamma} k''_{\delta} . \quad (22.27)$$

Then (22.14) is modified as follows:

$$\partial_\lambda j_\lambda^5 = 2m j^5 - \frac{2\alpha}{\pi} \mathbf{E} \cdot \mathbf{H} . \quad (22.28)$$

### 22.1.4 Trace Anomaly Term

As we shall see, the derivation of the Callan–Symanzik equation is deeply related to the trace anomaly term. To investigate this anomalous term, we shall consider a neutral scalar theory with Lagrangian density

$$\mathcal{L} = -\frac{1}{2}[(\partial_\lambda \varphi)^2 + m^2 \varphi] - \frac{g}{4!} \varphi^4 . \quad (22.29)$$

The energy–momentum tensor is

$$T_{\mu\nu} = -\frac{\partial \mathcal{L}}{\partial \varphi_{,\mu}} \varphi_{,\nu} + \delta_{\mu\nu} \mathcal{L} , \quad (22.30)$$

with trace

$$\begin{aligned} T_{\mu\mu} &= -\frac{\partial \mathcal{L}}{\partial \varphi_{,\mu}} \varphi_{,\mu} + 4\mathcal{L} \\ &= -(\partial_\mu \varphi)^2 - 2m^2 \varphi^2 - \frac{g}{6} \varphi^4 \\ &= -\partial_\mu (\varphi \partial_\mu \varphi) - m^2 \varphi^2 , \end{aligned} \quad (22.31)$$

where we have used the field equation. Taking  $T = m^2 \varphi^2$ ,

$$\begin{aligned} T[T_{\mu\mu}(x) \varphi(x_1) \dots \varphi(x_n)] &= -\partial_\mu T[\varphi(x) \partial_\mu \varphi(x), \varphi(x_1) \dots \varphi(x_n)] \\ &\quad - T[T(x) \varphi(x_1) \dots \varphi(x_n)] \\ &\quad + i \sum_{j=1}^n \delta^4(x - x_j) T[\varphi(x_1) \dots \varphi(x_n)] . \end{aligned} \quad (22.32)$$

We now take the Fourier transform:

$$\begin{aligned} \langle \mathbf{0} | T[\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle & \quad (22.33) \\ &= \frac{i^{n-1}}{(2\pi)^{4(n-1)}} \int (d p) e^{i(p_1 \cdot x_1 + \dots + p_n \cdot x_n)} \delta^4(p_1 + \dots + p_n) G^{(n)}(p_1, \dots, p_n), \end{aligned}$$

$$\begin{aligned} \langle \mathbf{0} | T[T_{\mu\mu}(x)\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle & \quad (22.34) \\ &= \frac{i^n}{(2\pi)^{4n}} \int (d p) d^4 q e^{i(q \cdot x + p_1 \cdot x_1 + \dots + p_n \cdot x_n)} \delta^4(q + p_1 + \dots + p_n) T_{\mu\mu}^{(n)}(q; p_1, \dots, p_n), \end{aligned}$$

and so on. Therefore, in the limit  $q \rightarrow \infty$ , Eq. (22.32) becomes

$$T_{\mu\mu}^{(n)}(0; p_1, \dots, p_n) = -T^{(n)}(0; p_1, \dots, p_n) + nG^{(n)}(p_1, \dots, p_n). \quad (22.35)$$

This is called a *trace identity*, but in fact *this equation is not true*. The reason is that it includes a product of operators at the same space-time point.

We now write down the Ward–Takahashi identity for the energy–momentum tensor. This equation is always true:

$$\frac{\partial}{\partial x_\mu} T[T_{\mu\nu}(x)\varphi(x_1) \dots \varphi(x_n)] = i \sum_{j=1}^n \delta^4(x - x_j) \frac{\partial}{\partial x_{j\nu}} T[\varphi(x_1) \dots \varphi(x_n)]. \quad (22.36)$$

Taking the Fourier transform of the vacuum expectation value, we have

$$q_\mu T_{\mu\nu}^{(n)}(q; p_1, \dots, p_n) = \sum_{j=1}^n (p_j + q)_\nu G^{(n)}(p_1, \dots, p_j + q, \dots, p_n). \quad (22.37)$$

Differentiating this equation with respect to  $q_\mu$  and taking the limit  $q \rightarrow 0$ ,

$$T_{\mu\nu}^{(n)}(0; p_1, \dots, p_n) = \left[ (n-1) + \sum_{j=1}^n p_{j\nu} \frac{\partial}{\partial p_{j\mu}} \right] G^{(n)}(p_1, \dots, p_n). \quad (22.38)$$

Taking the trace, we obtain

$$T_{\mu\mu}^{(n)}(0; p_1, \dots, p_n) = \left[ 4(n-1) + \sum_{j=1}^n p_j \cdot \frac{\partial}{\partial p_j} \right] G^{(n)}(p_1, \dots, p_n). \quad (22.39)$$

Since  $G^{(n)}$  has the dimension  $m^{4-3n}$  and is a homogeneous function of  $p_j$  and  $m$ ,

$$\left( \sum_{j=1}^n p_j \cdot \frac{\partial}{\partial p_j} + m \frac{\partial}{\partial m} + 3n - 4 \right) G^{(n)}(p_1, \dots, p_n) = 0. \tag{22.40}$$

Combining (22.39) with (22.40), this yields

$$T_{\mu\mu}^{(n)}(0; p_1, \dots, p_n) = \left( n - m \frac{\partial}{\partial m} \right) G^{(n)}(p_1, \dots, p_n). \tag{22.41}$$

This equation will be used to derive the Callan–Symanzik equation later.

## 22.2 Dispersion Theory for Green's Functions

To explain the dispersion theory for Green's functions, we consider a neutral scalar theory. If we introduce a set of Green's functions  $\{\bar{\tau}\}$  by (11.176), this set turns out to satisfy a generalized unitarity condition like (11.177). We introduce the dispersion relation by picking out the part corresponding to connected diagrams in  $\bar{\tau}$ :

$$\bar{\tau}(x_1, \dots, x_n)_{\text{conn}} = \bar{\rho}(x_1, \dots, x_n). \tag{22.42}$$

Therefore, a recursion equation similar to (11.29) holds between the sets  $\{\bar{\tau}\}$  and  $\{\bar{\rho}\}$ :

$$\bar{\tau}(x, x_1, \dots, x_n) = \bar{\rho}(x, x_1, \dots, x_n) + \sum_{k \neq n} \bar{\rho}(x, x'_1, \dots, x'_k) \bar{\tau}(x'_{k+1}, \dots, x'_n), \tag{22.43}$$

where the sum is taken over all ways to divide  $(x_1, \dots, x_n)$  into two sets,  $(x'_1, \dots, x'_k)$  and  $(x'_{k+1}, \dots, x'_n)$ . We now take the Fourier transform of  $\bar{\rho}$ :

$$\bar{\rho}(x_1, \dots, x_n) = \frac{-i}{(2\pi)^{4(n-1)}} \int (dp) \delta^4(p_1 + \dots + p_n) \mathcal{G}(p_1, \dots, p_n) e^{i(p_1 \cdot x_1 + \dots + p_n \cdot x_n)}. \tag{22.44}$$

Since  $\mathcal{G}$  is a function only of the scalar product  $p_\alpha \cdot p_\beta$ , we write it as  $\mathcal{G}(p_\alpha \cdot p_\beta)$ . Using the generalized unitarity condition, if  $\mathcal{G}$  is known up to order  $g^n$ , then we can compute  $\text{Im } \mathcal{G}$  up to order  $g^{n+1}$ . So if there exists a way to find  $\text{Re } \mathcal{G}$  for a given  $\text{Im } \mathcal{G}$ , then higher orders of  $\mathcal{G}$  can be determined recursively. Moreover, when we compute  $\text{Im } \mathcal{G}$  from the generalized unitarity condition, since the phase space volume is finite, there are no ultraviolet divergences. To compute  $\text{Re } \mathcal{G}$  from  $\text{Im } \mathcal{G}$ ,

we use the following parameter dispersion relation:

$$\operatorname{Re} \mathcal{G}(p_\alpha \cdot p_\beta \xi) = \frac{P}{\pi} \int_{-\infty}^{\infty} \frac{d\xi'}{\xi' - \xi} \epsilon(\xi') \operatorname{Im} \mathcal{G}(p_\alpha \cdot p_\beta \xi'), \quad (22.45)$$

or

$$\mathcal{G}(p_\alpha \cdot p_\beta \xi) = \int_{-\infty}^{\infty} \frac{d\xi'}{\xi'(1 - i\epsilon) - \xi} \operatorname{Im} \mathcal{G}(p_\alpha \cdot p_\beta \xi'), \quad (22.46)$$

where  $p_\alpha \cdot p_\beta \xi$  means that all scalar products between four-momenta should be multiplied by  $\xi$ . Since this dispersion relation holds true for all the  $\mathcal{G}$ , it turns out that we have a complete set for the dispersion relation.

In addition, in order to be able to write the dispersion relation when  $\xi = 0$  for all the  $\mathcal{G}$ , we must have

$$\operatorname{Re} \mathcal{G}(0) = \text{const.}, \quad \operatorname{Im} \mathcal{G}(0) = 0. \quad (22.47)$$

On the other hand, a complete set for the dispersion relation for the  $S$ -matrix element is not known. The proof of the dispersion relation (22.45) is straightforward. Renormalized Green's functions can be written, at all orders of perturbation theory, in the form

$$\mathcal{G}(p_\alpha \cdot p_\beta) = \int \frac{\sigma(z_j) dz_1 \dots dz_k}{[\sum c_{\alpha\beta}(z_j) p_\alpha \cdot p_\beta + M^2(z_j) - i\epsilon]^N}, \quad (22.48)$$

where  $\sigma$ ,  $c$ , and  $M$  are all real numbers and  $z_j$  is a suitable Feynman parameter. Using  $M^2 \geq 0$  and setting  $N = 1$  without loss of generality,

$$\int_0^\infty \frac{d\xi'}{\xi' - \xi} \delta(c\xi' + M^2) = \theta(-c) \frac{1}{c\xi + M^2}, \quad (22.49a)$$

$$\int_{-\infty}^0 \frac{d\xi'}{\xi' - \xi} \delta(c\xi' + M^2) = -\theta(c) \frac{1}{c\xi + M^2}. \quad (22.49b)$$

Taking the difference between them, the following dispersion relation is obtained:

$$\int_{-\infty}^{\infty} \frac{d\xi'}{\xi' - \xi} \epsilon(\xi') \delta(c\xi' + M^2) = \frac{1}{c\xi + M^2}. \quad (22.50)$$

However, this dispersion integral sometimes diverges. This corresponds to divergences in the field theory. In that case, we must implement the subtraction which corresponds to renormalization. Furthermore, through this subtraction, interactions are introduced in perturbation theory. Note also that the dispersion formula pre-



sented in Sect. 22.1 can be proven using a method similar to the one above. We will discuss subtractions in the next section.

## 22.3 Subtractions in Dispersion Relations

There are dispersion relations with no subtractions, with one subtraction, and with two subtractions:

$$\operatorname{Re} f(x) = \frac{1}{\pi} \int \frac{dx'}{x' - x} \operatorname{Im} f(x') \quad (\text{zero}), \quad (22.51)$$

$$\operatorname{Re} f(x) = f(a) + \frac{x - a}{\pi} \int \frac{dx'}{(x' - a)(x' - x)} \operatorname{Im} f(x') \quad (\text{one}), \quad (22.52)$$

$$\operatorname{Re} f(x) = f(a) + (x - a)f'(a) + \frac{(x - a)^2}{\pi} \int \frac{dx'}{(x' - a)^2(x' - x)} \operatorname{Im} f(x') \quad (\text{two}), \quad (22.53)$$

where the integrals are Cauchy principal values and  $a$  is chosen so that  $\operatorname{Im} f(a) = 0$ . If subtractions are necessary, they should be introduced in such a way as to reproduce the results in standard perturbation theory. We should choose subtraction constants so as to determine the normalizations for Green's functions.

In what follows, we start with a two-point function in a neutral scalar theory. Writing  $\mathcal{G}^{(2)}(p) = \mathcal{G}(p, -p)$ , from the integral representation of the two-point function, we have

$$\mathcal{G}^{(2)}(p) = -(p^2 + m^2) \left[ 1 + (p^2 + m^2) \int d\kappa^2 \frac{\sigma(\kappa^2)}{p^2 + \kappa^2 - i\epsilon} \right]. \quad (22.54)$$

If we now replace  $p^2$  by  $p^2\xi$ , the dispersion relation for  $\xi$  requires two subtractions. Note that the subtraction conditions can be obtained immediately from the above integral representation:

$$\mathcal{G}^{(2)} = 0, \quad \frac{\partial \mathcal{G}^{(2)}}{\partial p^2} = -1, \quad \text{where } p^2 + m^2 = 0. \quad (22.55)$$

In fact, these two subtractions correspond to the renormalizations for the mass and the factor  $Z$ .

We now consider a four-point function. At the lowest order, corresponding to (20.107),

$$\mathcal{G}^{(4)} = g. \quad (22.56)$$

Since  $\text{Im } \mathcal{G}^{(4)} = 0$  at this order, we see that  $\mathcal{G}^{(4)}$  requires one subtraction:

$$\text{Re } \mathcal{G}^{(4)}(p_\alpha \cdot p_\beta \xi) = \mathcal{G}^{(4)}(0) + \frac{\xi}{\pi} \int \frac{d\xi'}{\xi'(\xi' - \xi)} \epsilon(\xi') \text{Im } \mathcal{G}^{(4)}(p_\alpha \cdot p_\beta \xi'). \quad (22.57)$$

As the subtraction condition, we equate  $\mathcal{G}^{(4)}$  at a subtraction point (s.p.) with the renormalized coupling constant  $g$  :

$$\mathcal{G}^{(4)}(\text{s.p.}) = g, \quad (22.58)$$

using (20.112) to choose the s.p. such that

$$p_\alpha \cdot p_\beta = \frac{m^2}{3} (4\delta_{\alpha\beta} - 1). \quad (22.59)$$

This is the same as the subtraction condition in the  $S$ -matrix theory.

We now see that a two-point function needs two subtractions regardless of the details of the interactions. If no further subtractions are required for the other Green's functions, we can show that the fields become free, at least in perturbation theory. To do this, we write the part of  $\mathcal{G}^{(n)}$  linear in the coupling constant as  $\mathcal{G}^{(1)}$ . From the unitarity condition, we can show that

$$\text{Im } \mathcal{G}_1^{(n)} = 0. \quad (22.60)$$

According to the unitarity condition,  $\text{Im } \mathcal{G}_1^{(n)}$  is higher than second order with respect to  $\mathcal{G}$  and  $\mathcal{G}^\dagger$ , so at least one Green's function becomes a free Green's function  $\mathcal{G}_0$ . The only Green's function of free fields that does not vanish is a two-point function. However, the two-point function only appears in the unitarity condition in the form

$$\int d^4v \Delta^{(+)}(u-v) \bar{\rho}(v, x) = - \int d^4v \Delta^{(+)}(u-v) K_v K_x \Delta'_F(v-x). \quad (22.61)$$

The Fourier transform of this expression is

$$\delta(p^2 + m^2)(p^2 + m^2) \left[ 1 + (p^2 + m^2) \int d\kappa^2 \frac{\sigma(\kappa^2)}{p^2 + \kappa^2 - \epsilon} \right] = 0. \quad (22.62)$$

We thus see that a two-point function never appears in the non-linear contributions to the unitarity condition. Hence, if (22.60) holds true and there is no subtraction, we have

$$\text{Re } \mathcal{G}_1^{(n)} = 0, \quad n > 2, \quad (22.63)$$

and this reduces to a free field.

In the Feynman–Dyson theory, propagators play a fundamental role, but in the dispersion theory, the two-point function never appears in the middle of computations. Other Green’s functions are determined independently of the two-point function. However, if other Green’s functions are determined, then from the unitarity condition, a two-point function turns out to be determined through the spectral function  $\sigma(\kappa^2)$ .

## 22.4 Heisenberg Operators

Here we consider a Heisenberg operator  $A(x)$  and, in particular, the problems that occur when  $A(x)$  is a product of operators at a common space-time point. We start with the Green’s function

$$\bar{\tau}_A(x; x_1, \dots, x_n) = (-i)^{n+1} K_{x_1} \dots K_{x_n} \langle \mathbf{0} | T[A(x)\varphi(x_1) \dots \varphi(x_n)] | \mathbf{0} \rangle. \quad (22.64)$$

An arbitrary matrix element of  $A(x)$  is obtained using the LSZ reduction formula. We set the  $n$  four-momenta on the mass shell and take the Fourier transform of  $\bar{\tau}_A$ , which yields a matrix element of the form

$$\langle \boldsymbol{\beta}, \text{out} | A(x) | \boldsymbol{\alpha}, \text{in} \rangle. \quad (22.65)$$

If the set  $\{\bar{\tau}_A\}$  is known, it turns out that the field operator  $A(x)$  is determined.

The unitarity condition for  $\bar{\tau}_A$  is obtained along the same lines as the  $\bar{\tau}$  case:

$$\begin{aligned} 0 &= \bar{\tau}_A(x; x_1, \dots, x_n) + \bar{\tau}_A^*(x; x_1, \dots, x_n) \\ &+ \sum_{\text{comb}} \sum_{l=0}^{\infty} \frac{i^l}{l!} \int (du)(dv) \bar{\tau}_A(x; x_1, \dots, x_n) \\ &\quad \times \Delta^{(+)}(u_1 - v_1) \dots \Delta^{(+)}(u_l - v_l) \bar{\tau}^*(x'_{k+1}, \dots, x'_n) \\ &+ (\bar{\tau}_A \rightarrow \bar{\tau}, \bar{\tau}^* \rightarrow \bar{\tau}_A^*). \end{aligned} \quad (22.66)$$

Since this unitarity condition is linear in  $\bar{\tau}_A$  and  $\bar{\tau}_A^*$ , we call it a *linear unitarity condition*. We write the connected part of  $\bar{\tau}_A$  as  $\bar{\rho}_A$ , as in (22.43). This yields the recursion formula

$$\bar{\tau}_A(x; x_1, \dots, x_n) = \bar{\rho}_A(x; x_1, \dots, x_n) + \sum_{\substack{\text{comb} \\ k \neq n}} \bar{\rho}_A(x; x'_1, \dots, x'_k) \bar{\tau}(x'_{k+1}, \dots, x'_n). \quad (22.67)$$

Taking the Fourier transform of  $\bar{\rho}_A$ ,

$$\bar{\rho}_A(x; x_1, \dots, x_n) = \frac{-i}{(2\pi)^4} \int (dp) d^4q \delta(p_1 + \dots + p_n + q) \quad (22.68)$$

$$\times e^{i(q \cdot x + p_1 \cdot x_1 + \dots + p_n \cdot x_n)} \mathcal{A}(q; p_1, \dots, p_n).$$

The functional system  $\{\mathcal{A}\}$  satisfies a dispersion relation similar to  $\{\mathcal{G}\}$ :

$$\text{Re } \mathcal{A}(p_\alpha \cdot p_\beta \xi) = \frac{P}{\pi} \int_{-\infty}^{\infty} \frac{d\xi'}{\xi' - \xi} \epsilon(\xi') \text{Im } \mathcal{A}(p_\alpha \cdot p_\beta \xi). \quad (22.69)$$

It is obvious that, in perturbation theory, all the  $\mathcal{A}$  will vanish unless we introduce a subtraction. In this case, if  $\{\mathcal{G}\}$  is known, combining the linear unitarity condition with the dispersion relation,  $\{\mathcal{A}\}$  can be determined recursively. In the above discussion, we started with the operator  $A$  and set out to determine the set  $\{\mathcal{A}\}$ , but it sometimes happens that, if we start with two conditions and determine the functional system, that does not necessarily have a corresponding operator realization. Such a functional system turns up when we consider the various Ward–Takahashi identities.

## 22.5 Subtraction Condition

When we evaluate  $\{\mathcal{G}\}$  or  $\{\mathcal{A}\}$ , we find that the absorption part, i.e., the imaginary part, does not include divergences, due to the unitarity condition. Thus, if divergences appear in field theory, this happens in the dispersion part, i.e., the real part. We compute higher orders recursively by carrying out subtractions for an  $n$ -point function  $\mathcal{G}^{(n)}$ , the number of these subtractions being determined in advance. If a dispersion integral always converges at any higher order due to a previously determined number of subtractions, then that theory is renormalizable. Otherwise, if the higher the order, the more subtractions must be made, then it is a non-renormalizable theory.

In the following, we determine the number of subtractions ensuring renormalizability within the framework of perturbation theory. For this purpose, we make the following assumption:

*Assumption* For a large value of  $|\xi|$  and for almost all  $\{p_j\}$ , we have

$$\mathcal{G}^{(n)}(p_\alpha \cdot p_\beta \xi) \sim \xi^{c(n)/2}, \quad (22.70)$$

where  $c(n)$  is an index independent of  $\{p_j\}$ .

This is true in the Feynman–Dyson theory within the perturbative approach. The meaning of the above equation for an arbitrary positive  $\epsilon$  is

$$\lim_{\xi \rightarrow \infty} \frac{\mathcal{G}^{(n)}(p_\alpha \cdot p_\beta \xi)}{\xi^{c(n)/2+\epsilon}} = 0, \quad \lim_{t \rightarrow \infty} \frac{\mathcal{G}^{(n)}(p_\alpha \cdot p_\beta \xi)}{\xi^{c(n)/2-\epsilon}}. \quad (22.71)$$

Furthermore, (22.70) is sometimes written formally as

$$\mathcal{G}^{(n)}(p_\alpha \cdot p_\beta) \sim p^{c(n)}. \quad (22.72)$$

To determine  $c(n)$ , we use the unitarity condition:

$$\begin{aligned} 0 &= \bar{\rho}(x_1, \dots, x_n) + \bar{\rho}^*(x_1, \dots, x_n) \\ &+ \sum_{\text{comb}}' \sum_{l=1}^{\infty} \frac{i^l}{l!} \int (du)(dv) \bar{\rho}(x'_1, \dots, x'_k, u_1, \dots, u_l) \\ &\quad \times \Delta^{(+)}(u_1 - v_1) \dots \Delta^{(+)}(u_l - v_l) \bar{\rho}^*(x'_{k+1}, \dots, x'_k, v_1, \dots, v_l) \\ &+ \text{terms} \geq \text{third order in } \rho \text{ and } \rho^*. \end{aligned} \quad (22.73)$$

An exponent of  $\text{Im } \mathcal{G}^{(n)}$  never exceeds the order itself. Exponents of the first and second terms are given by the exponent of  $\text{Im } \mathcal{G}^{(n)}$ , while exponents of other terms are expressed in terms of those of  $\mathcal{G}$  and  $\mathcal{G}^*$ . Therefore,

$$c(n) + 4(n - 1) \geq \max[c(k + l) + c(n - k + l) + 4(n + 2l - 2) - 6l]. \quad (22.74)$$

Inequalities generated by terms higher than the third order hold true automatically if the above inequality does. Furthermore, terms other than  $c(n)$  are generated from kinematic factors, and since there is no two-point function,

$$k + l > 2, \quad n - k + l > 2. \quad (22.75)$$

We introduce the notation

$$d(n) = c(n) + n - 4. \quad (22.76)$$

Then, the above inequality can be written in the form

$$d(n) \geq \max[d(k + l) + d(n - k + l)]. \quad (22.77)$$

This immediately implies the necessary condition

$$d(n) \leq 0. \quad (22.78)$$

Checking the case with non-negative  $c(n)$ , in the  $\varphi^4$ -theory, from the lowest order result,

$$c(2) \geq 2, \quad c(4) \geq 0, \quad (22.79)$$

$$d(2) \geq 0, \quad d(4) \geq 0. \quad (22.80)$$

Combining this with the necessary condition (22.78), for even  $n$ ,

$$d(n) = 0, \quad \text{or} \quad c(n) = 4 - n. \quad (22.81)$$

This corresponds to a renormalizable theory. In contrast, in the case of non-renormalizable theories such as the  $\varphi^6$ -theory, even at the lowest order, it turns out that

$$c(2) \geq 2, \quad c(6) \geq 0, \quad \text{or} \quad d(2) \geq 0, \quad d(6) \geq 2, \quad (22.82)$$

so there is no solution consistent with (22.78).

On the other hand, for super-renormalizable theories such as the  $\varphi^3$ -theory,

$$d(2) \geq 0, \quad d(3) \geq -1. \quad (22.83)$$

Looking more closely, we find that  $d(n) = 0$  does not hold true, but

$$d(n) = 2 - n. \quad (22.84)$$

In the discussion above, we have considered scalar theories, but extensions are easy. For instance, in the pion–nucleon system, the function  $\bar{\tau}$  is given by

$$\begin{aligned} & (-i)^n K_{x_1} \dots K_{x_n} D_{y_1} \dots D_{y_l} \tilde{D}_{z_1} \dots \tilde{D}_{z_l} \\ & \times \langle \mathbf{0} | T[\varphi(x_1) \dots \varphi(x_n) \psi(y_1) \dots \psi(y_l) \bar{\psi}(z_1) \dots \bar{\psi}(z_l)] | \mathbf{0} \rangle, \end{aligned} \quad (22.85)$$

where

$$D = \gamma \cdot \partial + M, \quad \tilde{D} = \gamma^T \cdot \partial - M. \quad (22.86)$$

The contraction functions for the nucleon are

$$\langle 0 | \psi^{\text{in}}(x) \bar{\psi}^{\text{in}}(y) | 0 \rangle = -iS^{(+)}(x - y) = -i(\gamma \cdot \partial - M)\Delta^{(+)}(x - y), \quad (22.87)$$

$$\langle 0 | \bar{\psi}^{\text{in}}(x) \psi^{\text{in}}(y) | 0 \rangle = -i\tilde{S}^{(+)}(x - y) = -i(\gamma^T \cdot \partial + M)\Delta^{(+)}(x - y). \quad (22.88)$$

Setting  $m = 2l$ , we introduce the exponent of  $\mathcal{G}^{(n,m)}$ :

$$\mathcal{G}^{(n,m)} \sim p^{c(n,m)} . \quad (22.89)$$

Using the so-called *regular dimensions* 1 and  $3/2$  of  $\varphi$  and  $\bar{\psi}$  or  $\bar{\psi}$ , we define

$$d(n, m) = c(n, m) + n + \frac{3}{2}m - 4 . \quad (22.90)$$

The inequality corresponding to (22.77) is then

$$d(n, m) \geq \max[d(k + l, k' + l') + d(n - k + l, m - k' + l')] . \quad (22.91)$$

From this inequality, we obtain

$$d(n, m) \leq 0 , \quad \text{or} \quad c(n, m) \leq 4 - n - \frac{3}{2}m . \quad (22.92)$$

We consider the following two kinds of renormalizable interaction known in the pion–nucleon system:

$$i\bar{\psi}\gamma_5\psi \cdot \varphi , \quad \varphi^4 , \quad (22.93)$$

where the isospin has been omitted. In addition to the conditions arising from the two-point function, viz.,

$$c(2, 0) \geq 2 , \quad c(0, 2) \geq 1 , \quad (22.94)$$

we obtain the conditions

$$c(1, 2) \geq 0 , \quad c(4, 0) \geq 0 . \quad (22.95)$$

Again, the unique solution is given by

$$d(n, m) = 0 . \quad (22.96)$$

As we see above, it can be understood that in typical renormalizable theories, we always obtain  $d = 0$ . For a proof of renormalizability, in addition to this, we have to prove that the exponents of the real and imaginary parts of  $\mathcal{G}^{(n)}$  are equal. This requires detailed analyses which will be omitted here.

In the above, we considered the set  $\{\mathcal{G}\}$  of typical Green's functions. We now consider  $\{\mathcal{A}\}$ :

$$\mathcal{A}^{(n,m)}(p_\alpha \cdot p_\beta \xi) \sim \xi^{a(n,m)/2} . \quad (22.97)$$

Using the linear unitarity condition (22.66) corresponding to (22.91), we obtain

$$b(n, m) \geq \max[b(k + l, k' + l') + d(n - k + l, m - k' + l')], \quad (22.98)$$

where

$$b(n, m) = a(n, m) + n + \frac{3}{2}m - 4. \quad (22.99)$$

Therefore, in a renormalizable theory characterized by  $d = 0$ , the solution (22.98) is

$$b(n, m) = b. \quad (22.100)$$

This  $b$  does not depend on the values of  $n$  and  $m$ . It is called an *index of the set*  $\{\mathcal{A}\}$ , and is determined by perturbation theory.

For instance, in a scalar theory, if we consider

$$A(x) = \frac{1}{2}[\varphi(x)]^2, \quad (22.101)$$

then at the lowest order,

$$\mathcal{A}^{(2)} = 1. \quad (22.102)$$

Hence, it turns out that  $a(2) = 0$  and  $b = -2$ . For a polynomial of field operators, it is given by

$$b = (\text{dimension of } A) - 4. \quad (22.103)$$

Although the dispersion theory is formally elegant, the Feynman–Dyson theory is more useful for practical purposes. However, since there exist neither divergences nor indefiniteness, in a delicate issue such as derivation of the anomalous term, it gives a unique solution, and that is eminently useful. We thus give a theorem which provides a basis for this idea in the  $\varphi^4$ -theory:

**Theorem 22.1** *We assume that four sets of functions  $\{\mathcal{A}\}$ ,  $\{\mathcal{B}\}$ ,  $\{\mathcal{C}\}$ , and  $\{\mathcal{D}\}$  in  $\varphi^4$ -theory all satisfy the linear unitarity condition, that the index is 0 or  $-2$ , and that they have a scalar transformation property. Then there exists a linear dependence relation among these four sets of functions.*

To prove this theorem, we introduce the following  $\mathcal{S}^{(n)}$ :

$$\mathcal{S}^{(n)} = a\mathcal{A}^{(n)} + b\mathcal{B}^{(n)} + c\mathcal{C}^{(n)} + d\mathcal{D}^{(n)}. \quad (22.104)$$



In general, the index of  $\{\mathcal{A}\}$  is 0. It satisfies the linear unitarity relation. If we add the following three subtraction conditions, then this set is uniquely determined:

$$\begin{aligned} \mathcal{S}^{(2)}(0; p, -p), \quad \frac{\partial}{\partial p^2} \mathcal{S}^{(2)}(0; p, -p) \quad \text{for } p^2 + m^2 = 0, \\ \mathcal{S}^{(4)}(0; p_1, \dots, p_4) \quad \text{for } p_\alpha \cdot p_\beta = \frac{m^2}{3}(1 - 4\delta_{\alpha\beta}). \end{aligned} \quad (22.105)$$

Since these subtraction constants are linear in  $a$ ,  $b$ , and  $c$ , if we choose their ratio properly, all three of the above constants can be set to zero. However, the values of  $\mathcal{A}, \dots$  at the above subtraction point are assumed to be already known. If all the subtraction constants are zero, from the linear unitarity condition and the dispersion formula, all members of  $\{\mathcal{S}\}$  must vanish, i.e., if we determine the ratios of  $a$ ,  $b$ ,  $c$ , and  $d$ , according to the above argument, we must have

$$\mathcal{S}^{(n)} = a\mathcal{A}^{(n)} + b\mathcal{B}^{(n)} + c\mathcal{C}^{(n)} + d\mathcal{D}^{(n)} = 0, \quad (22.106)$$

which proves the existence of the linear dependence relation.

## 22.6 Anomalous Trace Identity

We have already presented the Ward–Takahashi identity for the energy–momentum tensor (22.37) in Sect. 22.1. This can be considered as a defining identity for  $T_{\mu\nu}$ . We now rewrite it for Green’s functions in which the legs of propagators are removed:

$$\begin{aligned} q_\mu \mathcal{T}_{\mu\nu}^{(n)}(q; p_1, \dots, p_n) \\ = \sum_{j=1}^n \frac{p_j^2 + m^2}{(p_j + q)^2 + m^2 - i\epsilon} (p_j + q)_\nu \mathcal{G}^{(n)}(p_1, \dots, p_j + q, \dots, p_n), \end{aligned} \quad (22.107)$$

where  $q + p_1 + \dots + p_n = 0$ . We may take this to define  $\{\mathcal{T}_{\mu\nu}\}$ . If all the  $p_j$  are on the mass shell, then for almost all values of  $q$ , i.e., when the denominator on the right-hand side does not vanish, the right-hand side does vanish:

$$\langle \boldsymbol{\beta}, \text{out} | \partial_\mu \mathbf{T}_{\mu\nu}(x) | \boldsymbol{\alpha}, \text{in} \rangle = 0. \quad (22.108)$$

We thus obtain the conservation law

$$\partial_\mu \mathbf{T}_{\mu\nu} = 0. \quad (22.109)$$

Moreover, Eq. (22.107) implies that the index of  $\{\mathcal{T}_{\mu\nu}\}$  is zero, so

$$a(n) = c(n) = 4 - n , \quad (22.110)$$

and for  $n = 2$ , the quantity  $\mathcal{T}_{\mu\nu}^{(2)}$  can be determined by solving (22.107):

$$\begin{aligned} \mathcal{T}_{\mu\nu}^{(2)}(q; p_1, p_2) &= \delta_{\mu\nu}(p_1 \cdot p_2 - m^2) - (p_{1\mu} \cdot p_{1\nu} + p_{2\mu} p_{1\nu}) + \frac{1}{2}(p_1^2 + m^2)(p_2^2 + m^2) \\ &\quad \times \left[ \delta_{\mu\nu} \int d\kappa^2 \sigma(\kappa^2) \left( \frac{1}{p_1^2 + \kappa^2 - i\epsilon} + \frac{1}{p_2^2 + \kappa^2 - i\epsilon} \right) \right. \\ &\quad \left. + k_\mu k_\nu \int d\kappa^2 \frac{\sigma(\kappa^2)}{(p_1^2 + \kappa^2 - i\epsilon)(p_2^2 + \kappa^2 - i\epsilon)} \right] + S_{\mu\nu}^{(1)} F + S_{\mu\nu}^{(2)} G , \end{aligned} \quad (22.111)$$

where  $q = -(p_1 + p_2)$  and  $k = p_1 - p_2$ , while  $S_{\mu\nu}^{(1)}$  and  $S_{\mu\nu}^{(2)}$  are defined by

$$\begin{aligned} S_{\mu\nu}^{(1)} &= q^2 \delta_{\mu\nu} - q_\mu q_\nu , \\ S_{\mu\nu}^{(2)} &= (q \cdot k)^2 q_\mu q_\nu - (q^2)^2 k_\mu k_\nu - q^2 (q \cdot k) (q_\mu k_\nu + q_\nu k_\mu) , \end{aligned} \quad (22.112)$$

and satisfy  $q_\mu S_{\mu\nu}^{(i)} = 0$ . Therefore, for a large value of  $\xi$ ,

$$F(\xi) \sim \xi^0 , \quad G(\xi) \sim \xi^{-2} . \quad (22.113)$$

Thus, in order to determine  $\{\mathcal{T}_{\mu\nu}\}$  uniquely, we must give a subtraction condition for  $F$ . However, this arbitrariness does not remain for  $q = 0$ .

Since we have  $\mathcal{T}_{\mu\nu}^{(4)} \sim \xi^0$  for  $n = 4$ , even if a tensor such as  $S_\nu$  appears, its coefficient does not require subtraction. Now, taking  $F \neq 0$ , this arbitrariness is transmitted to a multi-point function via the linear unitarity condition. So in order to investigate the properties of this term, we consider the operator  $\mathbf{T} = m^2 \boldsymbol{\varphi}^2$ . Considering the corresponding set of functions

$$\mathcal{T}^{(n)}(q; p_1, \dots, p_n) , \quad (22.114)$$

its index is  $-2$ , so the unique subtraction is completely determined by

$$\mathcal{T}^{(2)}(0; p_1, \dots, p_n) = 2m^2 , \quad \text{where } p^2 + m^2 = 0 . \quad (22.115)$$

Thus, the set starting with  $\mathcal{T}^{(2)}$  should be proportional to the set starting with  $F$ :

$$F^{(n)}(q; p_1, \dots, p_n) = a \mathcal{T}^{(n)}(q; p_1, \dots, p_n) . \quad (22.116)$$

Therefore, assuming that  $\{\mathcal{T}_{\mu\nu}\}$  is one of the solutions satisfying the Ward–Takahashi identity and the linear unitarity condition, we obtain the following solution with the same property:

$$\mathcal{T}_{\mu\nu}^{(n)}(q; p_1, \dots, p_n) + a(q^2\delta_{\mu\nu} - q_\mu q_\nu)\mathcal{T}^{(n)}(q; p_1, \dots, p_n). \quad (22.117)$$

If we express this in terms of operators, it yields the generalization

$$\mathbf{T}_{\mu\nu}(x) \rightarrow \mathbf{T}_{\mu\nu}(x) + \lambda(\delta_{\mu\nu}\square - \partial_\mu\partial_\nu)\boldsymbol{\varphi}^2(x). \quad (22.118)$$

This arbitrariness has been used by Callan et al. to introduce the improved energy–momentum tensor [195].

However, this arbitrariness disappears for  $q = 0$ , and (22.38) holds. We have already introduced two sets of Green’s function systems, viz.,  $\{\mathcal{T}_{\mu\nu}\}$  and  $\mathcal{T}$ . We now introduce a third set:

$$\mathcal{W}^{(n)}(q; p_1, \dots, p_n) = \sum_f \frac{p_j^2 + m^2}{(p_j + q)^2 + m^2 - i\epsilon} \mathcal{G}^{(n)}(p_1, \dots, p_j + q, \dots, p_n). \quad (22.119)$$

This set is the Fourier transform of the Green’s function

$$(-i)^n K_{x_1} \dots K_{x_n} \sum_j \delta^4(x - x_j) \langle \mathbf{0} | T[\boldsymbol{\varphi}(x_1) \dots \boldsymbol{\varphi}(x_n)] | \mathbf{0} \rangle. \quad (22.120)$$

It is straightforward to check that this set satisfies the linear unitarity condition. However, assuming that all the  $q_j$  are on the mass shell for  $q \neq 0$ ,  $\mathcal{W}^{(n)}$  vanishes and hence does not correspond to any operator. This was discussed at the end of Sect. 22.4. This is a set of functions which appear through the connection with the Ward–Takahashi identity.

We now investigate what happens if  $q \rightarrow 0$ . Equation (22.119) is a highly singular function, since the denominator and the numerator cancel. In this case,

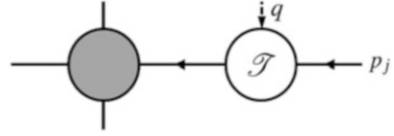
$$\mathcal{W}^{(n)}(0; p_1, \dots, p_n) = n\mathcal{G}^{(n)}(p_1, \dots, p_n). \quad (22.121)$$

This function neither vanishes on the mass shell nor satisfies the linear unitarity condition. Moreover, in this case, similar difficulties arise for other sets. This is caused by the pole

$$[(p_j + q)^2 + m^2 - i\epsilon]^{-1}. \quad (22.122)$$

In the case of  $\{\mathcal{T}_{\mu\nu}\}$  and  $\{\mathcal{T}\}$ , this pole appears in the diagram when a two-point function is connected to another part by a propagator, as shown in Fig. 22.2. If  $p_j^2 + m^2 = 0$  for  $q = 0$ , then (22.122) diverges. This is the reason why in field theory a

**Fig. 22.2** Diagram involving the pole (22.122), when a two-point function is connected to another part by a propagator



one-particle irreducible function is often introduced instead of a connected Green's function. To ensure that there is no pole of the form (22.122) for  $q = 0$  even on the mass shell, it turns out that we must take  $\mathcal{S}^{(n)}$  to be a suitable linear combination of  $\mathcal{T}_{\mu\nu}^{(n)}$ ,  $\mathcal{T}^{(n)}$ , and  $\mathcal{W}^{(n)}$ , and in the vicinity of  $p^2 + m^2 = 0$ , impose the relation

$$\mathcal{S}^{(2)}(0; p, -p) = \mathcal{O}((p^2 + m^2)^2). \quad (22.123)$$

We thus take  $\mathcal{S}^{(n)}$  to have the form

$$\mathcal{S}^{(n)} = \mathcal{T}_{\mu\mu}^{(n)} + \mathcal{T}^{(n)} - d_\varphi(g)\mathcal{W}^{(n)}. \quad (22.124)$$

We now determine the coefficient  $d_\varphi(g)$  which ensures that the condition (22.123) is satisfied. In addition, we assume that the above  $\mathcal{S}^{(n)}$  is defined only for  $q = 0$ . We thus expand  $\mathcal{T}_{\mu\mu}^{(2)}$ ,  $\mathcal{T}^{(2)}$ , and  $\mathcal{W}^{(2)}$  as power series in  $(p^2 + m^2)$  when  $q = 0$ . If we use (22.111) for  $\mathcal{T}_{\mu\mu}^{(2)}$ , then

$$\begin{aligned} \mathcal{T}_{\mu\mu}^{(2)}(0; p, -p) &= -2m^2 - 2(p^2 + m^2) + \mathcal{O}((p^2 + m^2)^2), \\ \mathcal{G}^{(2)}(p, -p) &= -(p^2 + m^2) + \mathcal{O}((p^2 + m^2)^2), \\ \mathcal{T}^{(2)}(0; p, -p) &= 2m^2 - 2\gamma_\varphi(g)(p^2 + m^2) + \mathcal{O}((p^2 + m^2)^2), \end{aligned} \quad (22.125)$$

where  $\gamma_\varphi(g)$  must be determined by computing higher order terms in perturbation theory. Therefore, in order to satisfy (22.123), we should choose the coefficient  $d_\varphi(g)$  as follows:

$$d_\varphi(g) = 1 + \gamma_\varphi(g). \quad (22.126)$$

Here,  $\gamma_\varphi(g)$  and 1 are called the *anomalous dimension of  $\varphi$*  and the *holomorphic dimension*, respectively. For a general  $n$ ,

$$\mathcal{S}^{(n)} = \mathcal{T}_{\mu\mu}^{(n)}(0; p_1, \dots, p_n) + \mathcal{T}^{(n)}(0; p_1, \dots, p_n) - nd_\varphi(g)\mathcal{G}^{(n)}(p_1, \dots, p_n). \quad (22.127)$$

It turns out that this set satisfies the linear unitarity condition. Moreover, for  $n = 2$ , the subtraction condition is

$$\mathcal{S}^{(2)}(p, -p) = 0, \quad \frac{\partial}{\partial p^2}\mathcal{S}^{(2)}(p, -p) = 0, \quad \text{for } p^2 + m^2 = 0. \quad (22.128)$$

For  $n = 4$ , using the subtraction point (22.59), we have to use the result of the computation

$$\mathcal{T}_{\mu\mu}^{(4)}(0; \text{s.p.}) + \mathcal{T}^{(4)}(0; \text{s.p.}) - 4d_\varphi(g)\mathcal{G}^{(4)}(\text{s.p.}) \equiv \beta(g). \quad (22.129)$$

This determines the subtraction condition.

To obtain the linear dependence relation, we introduce another set of functions:

$$\left\{ \frac{\partial}{\partial g} \mathcal{G}^{(n)}(p_1, \dots, p_n) \right\}. \quad (22.130)$$

Differentiating the non-linear unitarity condition for  $\mathcal{G}$  with respect to  $g$ , it can be shown that the above set satisfies the linear unitarity condition as well. Furthermore, for  $p^2 + m^2$ , the subtraction condition is

$$\frac{\partial}{\partial g} \mathcal{G}^{(2)}(p, -p) = 0, \quad \frac{\partial}{\partial p^2} \left[ \frac{\partial}{\partial g} \mathcal{G}^{(2)}(p, -p) \right] = 0, \quad (22.131)$$

$$\frac{\partial}{\partial g} \mathcal{G}^{(4)}(\text{s.p.}) = 1. \quad (22.132)$$

Comparing  $\{\mathcal{S}\}$  with  $\{\partial\mathcal{G}/\partial g\}$ , we see that they are proportional to each other:

$$\mathcal{S}^{(2)}(p_1, \dots, p_n) = \beta(g) \frac{\partial}{\partial g} \mathcal{G}^{(n)}(p_1, \dots, p_n). \quad (22.133)$$

This is an example of the linear dependence relation mentioned in the previous section. The four sets and their indices are

$$\begin{array}{cccc} \{\mathcal{T}_{\mu\nu}\}, & \{\mathcal{T}\}, & \{\mathcal{W}\}, & \left\{ \frac{\partial}{\partial g} \mathcal{G} \right\} \\ 0 & -2 & 0 & 0 \end{array} \quad (22.134)$$

If we express the relation obtained above by a Green's function without removing the legs of the propagators, we obtain

$$\begin{aligned} T_{\mu\mu}^{(n)}(0; p_1, \dots, p_n) + T^{(n)}(0; p_1, \dots, p_n) - nd_\varphi(g)G^{(n)}(p_1, \dots, p_n) \\ = \beta(g) \frac{\partial}{\partial g} G^{(n)}(p_1, \dots, p_n). \end{aligned} \quad (22.135)$$

This is called an *anomalous trace identity*. Comparing this with the incorrect (22.35), it should be clear what kinds of anomalous terms have been generated.

Furthermore, combining with (22.41),

$$\left[ m \frac{\partial}{\partial m} + \beta(g) \frac{\partial}{\partial g} + n \gamma_\varphi(g) \right] G^{(n)}(p_1, \dots, p_n) = T^{(n)}(0; p_1, \dots, p_n). \quad (22.136)$$

This is nothing but the Callan–Symanzik equation given by (20.136). We thus see that, although this equation has been derived from unrenormalized divergent quantities, it can be derived solely from finite fixed quantities. Moreover, in this method, anomalous terms are generated without any ambiguities, so it is a good approach when discussing more delicate issues.

## 22.7 Triangle Anomaly Terms

In the last section, we investigated the anomalous trace identity. Another typical anomalous term is the triangle anomaly term. We shall now investigate this anomalous term.

We first introduce the *unrenormalized operators*

$$\begin{aligned} A_\lambda^{(0)} &= i\bar{\psi}^{(0)} \gamma_\lambda \gamma_5 \psi^{(0)}, & P^{(0)} &= i\bar{\psi}^{(0)} \gamma_5 \psi^{(0)}, & S^{(0)} &= \bar{\psi}^{(0)} \psi^{(0)}, \\ C_\lambda^{(0)} &= \frac{i}{4} \epsilon_{\lambda\alpha\beta\gamma} \phi_\alpha^{(0)} F_{\beta\gamma}^{(0)}, & B^{(0)} &= \partial_\lambda C_\lambda^{(0)} = \mathbf{E}^{(0)} \cdot \mathbf{H}^{(0)}, \end{aligned} \quad (22.137)$$

where we have used  $\phi_\lambda^{(0)}$  for the potential of the electromagnetic field since  $A_\lambda^{(0)}$  stands for an axial vector current. The triangle anomaly terms appear in the *renormalized equation*

$$\partial_\lambda (A_\lambda - a C_\lambda) = 2m \mathbf{P}, \quad (22.138)$$

$$a = -\frac{e^2}{2\pi^2}, \quad (22.139)$$

where (22.139) is the lowest order result. However, assuming that these operators are renormalized properly, this result becomes strict. This is the *Adler–Bardeen theorem* [196]. We shall also prove this theorem.

We first express (22.138) using Green's functions:

$$\partial_\lambda [A_\lambda^{(n,m)} - a C_\lambda^{(n,m)}] = 2m P^{(n,m)} + W^{(n,m)}, \quad (22.140)$$

where these Green's functions are defined by

$$A_\lambda^{(n,m)}(w; x \dots y \dots z \dots) = \langle \mathbf{0} | T[A_\lambda(w)\phi(x) \dots \psi(y) \dots \bar{\psi}(z) \dots] | \mathbf{0} \rangle. \quad (22.141)$$

Here  $n$  is the number of fields  $\phi$  and  $m$  is the number of fields  $\psi$  and  $\bar{\psi}$ . Additionally,

$$\begin{aligned} W^{(n,m)}(w; x \dots y \dots z \dots) &= - \sum_j \delta^4(w - y_j)(\gamma_5)_j G^{(n,m)}(x \dots y \dots z \dots) \\ &\quad - \sum_j \delta^4(w - z_j) G^{(n,m)}(x \dots y \dots z \dots)(\gamma_5)_j, \end{aligned} \quad (22.142)$$

with

$$G^{(n,m)}(x \dots y \dots z \dots) = \langle \mathbf{0} | T[\phi(x) \dots \psi(y) \dots \bar{\psi}(z) \dots] | \mathbf{0} \rangle. \quad (22.143)$$

Moreover, in order to chop the legs off the propagators, we multiply  $K$ ,  $D$ ,  $\tilde{D}$ , and so on. Here,  $D$  and  $\tilde{D}$  are given by (22.86) and  $K$  is given by

$$K_{\mu\nu}(\partial) = (\delta_{\mu\nu}\square - \partial_\mu\partial_\nu) + \frac{1}{\alpha}\partial_\mu\partial_\nu. \quad (22.144)$$

This defines Green's functions like the ones in (22.68). For instance,

$$\begin{aligned} (-i)^n K_x \dots D_y \dots \tilde{D}_z \dots \langle \mathbf{0} | T[\mathbf{P}(w)\phi(x) \dots \psi(y) \dots \bar{\psi}(z) \dots] | \mathbf{0} \rangle &\quad (22.145) \\ &= \frac{1}{(2\pi)^{4(n+m)}} \int (dk)(dp)(d\bar{p}) \delta^4(q + k + \dots + p + \dots + \bar{p} + \dots) \\ &\quad \times \exp[i(q \cdot w + k \cdot x + \dots + p \cdot y + \dots + \bar{p} \cdot z + \dots)] \\ &\quad \times \mathcal{P}^{(n,m)}(q; k \dots p \dots \bar{p} \dots). \end{aligned}$$

Equation (22.140) then becomes

$$iq_\lambda [\mathcal{A}_\lambda^{(n,m)} - a\mathcal{E}_\lambda^{(n,m)}] = 2m\mathcal{P}^{(n,m)} + \mathcal{W}^{(n,m)}. \quad (22.146)$$

### 22.7.1 Renormalization Condition

We now investigate several kinds of renormalization condition, i.e., normalization, for Green's functions.

### The Set $\{\mathcal{P}\}$

In this case, it can be shown using perturbation theory that the order of divergence is given by

$$d(\mathcal{P}) = 3 - n - \frac{3}{2}m. \quad (22.147)$$

From the invariance under charge conjugation, we see that  $d(\mathcal{P})$  is non-negative only for the channels  $(0, 2)$  and  $(2, 0)$ . We thus expand  $\mathcal{P}^{(0,2)}$  as follows:

$$\begin{aligned} \mathcal{P}^{(0,2)}(q; p, \bar{p}) &= (i\gamma_5)\mathcal{P}_1^{(0,2)} \\ &+ (ip \cdot \gamma + m)(i\gamma_5)\mathcal{P}_2^{(0,2)} + (i\gamma_5)(-i\bar{p} \cdot \gamma + m)\tilde{\mathcal{P}}_2^{(0,2)} \\ &+ i(p \cdot \gamma + m)(i\gamma_5)(-i\bar{p} \cdot \gamma + m)\mathcal{P}_3^{(0,2)}. \end{aligned} \quad (22.148)$$

Then  $\mathcal{P}_1$  and  $\mathcal{P}_3$  are invariant under charge conjugation  $p \rightleftharpoons \bar{p}$ , while  $\mathcal{P}_2$  and  $\tilde{\mathcal{P}}_2$  are interchanged by this operation. Since  $\mathcal{P}^{(0,2)}$  diverges logarithmically overall, only  $\mathcal{P}_1$  requires renormalization. We choose the subtraction point such that

$$q = 0, \quad p^2 + m^2 = \bar{p}^2 + m^2 = 0, \quad (22.149)$$

and we choose the subtraction condition

$$\mathcal{P}_1^{(0,2)}(\text{s.p.}) + 2m\mathcal{P}_2^{(0,2)}(\text{s.p.}) = 1. \quad (22.150)$$

Usually, we choose  $\mathcal{P}_1^{(0,2)}(\text{s.p.}) = 1$ , but in that case the Adler–Bardeen theorem does not hold. Renormalizing as in (22.150), in the vicinity of the mass shell,

$$\mathcal{P}^{(0,2)}(0; p, -p) = i\gamma_5[1 + \mathcal{O}(p^2 + m^2)]. \quad (22.151)$$

This coincides with the renormalization condition given by Adler and Bardeen.

In the channel  $(2, 0)$ , a linear divergence is expected. We expand as

$$\mathcal{P}_{\rho\sigma}^{(2,0)}(q; k_1, k_2) = i\epsilon_{\rho\sigma\alpha\beta}(k_1)_\alpha(k_2)_\beta\mathcal{P}_1^{(2,0)}. \quad (22.152)$$

Here  $\rho$  and  $\sigma$  are vector indices. Written like this, since  $\mathcal{P}_1^{(2,0)}$  goes as  $p^{-1}$ , there is no divergence and no renormalization is necessary.

Therefore, only the channel  $(0, 2)_1$  requires subtraction for  $\mathcal{P}$ , and since this subtraction condition determines its normalization in this case, it becomes a multiplicative renormalization.



### The Set $\{\mathcal{W}\}$

This set is typical of the Ward–Takahashi identity and has the property that it vanishes if all the  $p_j$  except  $q$  are on the mass shell. In this case, for other Green's functions, for instance for  $\mathcal{P}$ , this becomes a matrix element of the operator  $\mathbf{P}$ , but vanishes for  $\mathcal{W}$ . The definition immediately implies the following:

$$\mathcal{W}_1^{(0,2)}(\text{s.p.}) = 0, \quad (22.153)$$

$$\mathcal{W}_2^{(0,2)}(\text{s.p.}) = \tilde{W}_2^{(0,2)}(\text{s.p.}) = -1. \quad (22.154)$$

These give the subtraction condition. In the channel (2, 0),

$$\mathcal{W}^{(2,0)} = 0. \quad (22.155)$$

### The Set $\{\mathcal{A}_\lambda\}$

$\mathcal{A}_\lambda$  also has the same index as  $\mathcal{P}$ . In the channel (0, 2),

$$\begin{aligned} \mathcal{A}_\lambda^{(0,2)}(q; p, \bar{p}) &= (i\gamma_\lambda \gamma_5) \mathcal{A}_1^{(0,2)} \\ &+ (i\mathbf{p} \cdot \boldsymbol{\gamma} + m)(i\gamma_\lambda \gamma_5) \mathcal{A}_2^{(0,2)} + (i\gamma_\lambda \gamma_5)(-\bar{\mathbf{p}} \cdot \boldsymbol{\gamma} + m) \tilde{\mathcal{A}}_2^{(0,2)} \\ &+ (i\mathbf{p} \cdot \boldsymbol{\gamma} + m)(i\gamma_\lambda \gamma_5)(-i\bar{\mathbf{p}} \cdot \boldsymbol{\gamma} + m) \mathcal{A}_3^{(0,2)} + q_\lambda \gamma_5 \mathcal{A}_4^{(0,2)} \\ &+ (i\mathbf{p} \cdot \boldsymbol{\gamma} + m)(q_\lambda \gamma_5) \mathcal{A}_5^{(0,2)} + (q_\lambda \gamma_5)(-i\bar{\mathbf{p}} \cdot \boldsymbol{\gamma} + m) \tilde{A}_5^{(0,2)} \\ &+ (i\mathbf{p} \cdot \boldsymbol{\gamma} + m)(q_\lambda \gamma_5)(-i\bar{\mathbf{p}} \cdot \boldsymbol{\gamma} + m) \mathcal{A}_6^{(0,2)}. \end{aligned} \quad (22.156)$$

Only  $\{\mathcal{A}_1^{(0,2)}\}$  requires subtractions. The subtraction condition is

$$\mathcal{A}_1^{(0,2)}(\text{s.p.}) = \mathcal{P}_1^{(0,2)}(\text{s.p.}). \quad (22.157)$$

Taking into account the bosonic symmetry, the expansion in the channel (2, 0) is

$$\begin{aligned} \mathcal{A}_{\lambda\rho\sigma}^{(2,0)}(q; k_1, k_2) &= \epsilon_{\rho\sigma\alpha\beta}(k_1)_\alpha(k_2)_\beta [(k_1)_\lambda \mathcal{A}_1^{(2,0)} + (k_2)_\lambda \tilde{\mathcal{A}}_1^{(2,0)}] \\ &+ [\epsilon_{\lambda\rho\alpha\beta}(k_1)_\sigma(k_1)_\alpha(k_2)_\beta + (k_1 \cdot k_2)\epsilon_{\lambda\rho\sigma\alpha}(k_1)_\alpha] \mathcal{A}_2^{(2,0)} \\ &- [\epsilon_{\lambda\sigma\alpha\beta}(k_2)_\rho(k_1)_\alpha(k_2)_\beta + (k_1 \cdot k_2)\epsilon_{\lambda\rho\sigma\alpha}(k_2)_\alpha] \tilde{\mathcal{A}}_2^{(2,0)} \\ &+ [\epsilon_{\lambda\rho\alpha\beta}(k_2)_\sigma(k_1)_\alpha(k_2)_\beta + k_2^2 \epsilon_{\lambda\rho\sigma\alpha}(k_1)_\alpha] \mathcal{A}_3^{(2,0)} \\ &- [\epsilon_{\lambda\sigma\alpha\beta}(k_1)_\rho(k_1)_\alpha(k_2)_\beta + k_1^2 \epsilon_{\lambda\rho\sigma\alpha}(k_2)_\alpha] \tilde{\mathcal{A}}_3^{(2,0)} \\ &+ (k_1^2 - k_2^2)\epsilon_{\lambda\rho\sigma\alpha}(k_1 + k_2)_\alpha \mathcal{A}_4^{(2,0)} \\ &+ \epsilon_{\lambda\rho\sigma\alpha}(k_1 - k_2)_\alpha \mathcal{A}_5^{(2,0)}. \end{aligned} \quad (22.158)$$

Note that  $\mathcal{A}_j$  and  $\tilde{\mathcal{A}}_j$  are mapped to one another for  $j = 1, 2, 3$ , under the exchange  $k_1 \rightleftharpoons k_2$ . Here,  $\mathcal{A}^{(2,0)}$  is expected to diverge linearly. Looking at the total coefficients, we see that only  $\mathcal{A}_5$  requires subtractions. For the gauge invariant  $A_\lambda$ , we have

$$(k_1)_\rho \mathcal{A}_{\lambda\rho\sigma}^{(2,0)} = (k_2)_\sigma \mathcal{A}_{\lambda\rho\sigma}^{(2,0)} = 0 . \quad (22.159)$$

From this, we obtain

$$\mathcal{A}_4^{(2,0)} = \mathcal{A}_5^{(2,0)} = 0 . \quad (22.160)$$

Hence, in this channel, subtractions are not required. The only subtraction condition appears in  $(2, 0)_1$ , whence it finally becomes a multiplicative renormalization.

### The Set $\{\mathcal{C}_\lambda\}$

Since  $\{\mathcal{C}_\lambda\}$  is a completely new operator, to determine the renormalization condition, we choose a perturbation theory at the lowest order as a guide. Because the index is the same as that of  $\{\mathcal{A}_\lambda\}$ , if we first consider the channel  $(0, 2)$ , it vanishes at the lowest order, so we choose the renormalization condition

$$\mathcal{C}_1^{(2,0)}(\text{s.p.}) = 0 . \quad (22.161)$$

Next, in the channel  $(2, 0)$ , the lowest order computation is

$$\mathcal{C}_{\lambda\rho\sigma}^{(2,0)} = \frac{1}{2} \epsilon_{\lambda\rho\sigma\alpha} (k_1 - k_2)_\alpha . \quad (22.162)$$

We thus choose (22.149) as the subtraction point and the renormalization condition

$$\mathcal{C}_5^{(0,2)}(\text{s.p.}) = \frac{1}{2} . \quad (22.163)$$

The component above does not disappear, because  $C_\lambda$  is not gauge invariant. After all, there are two renormalization conditions for  $C_\lambda$ , viz., (22.161) and (22.163), and the renormalization is not multiplicative. Instead, renormalizations by linear transformations among operators are required. As given in (20.156), they have the form

$$A_\lambda^{(0)} = Z_A A_\lambda , \quad C_\lambda^{(0)} = Z_C C_\lambda + z A_\lambda . \quad (22.164)$$

**The Set  $\{\mathcal{D}\}$** 

In the Ward–Takahashi identity (22.146) for triangle anomaly terms,  $\mathcal{A}_\lambda$  always appears in the form

$$\mathcal{D}^{(n,m)} = iq_\lambda \mathcal{A}_\lambda^{(n,m)}. \quad (22.165)$$

Its power is

$$d(\mathcal{D}) = 4 - n - \frac{3}{2}m. \quad (22.166)$$

Thus, (2, 0) yields linear divergences, while (0, 2)<sub>1</sub> and (0, 2)<sub>2</sub> require subtractions. From the subtraction condition (22.157) for  $\{\mathcal{A}_\lambda\}$ ,

$$\mathcal{D}_1^{(0,2)}(\text{s.p.}) = 2m \mathcal{P}_1^{(0,2)}(\text{s.p.}), \quad (22.167)$$

$$\mathcal{D}_2^{(0,2)}(\text{s.p.}) = \tilde{D}_2^{(0,2)}(\text{s.p.}) = -\mathcal{P}_1^{(0,2)}(\text{s.p.}). \quad (22.168)$$

Similarly,

$$\mathcal{D}_1^{(2,0)}(\text{s.p.}) = 0. \quad (22.169)$$

**The Set  $\{\mathcal{B}\}$** 

In this case,  $\mathcal{C}_\lambda$  also appears in (22.146) in the form

$$\mathcal{B}^{(n,m)} = iq_\lambda \mathcal{C}_\lambda^{(n,m)}. \quad (22.170)$$

The subtraction conditions are obtained from (22.161) and (22.163) as

$$\mathcal{B}_1^{(0,2)}(\text{s.p.}) = 0, \quad \mathcal{B}_2^{(0,2)}(\text{s.p.}) = \tilde{\mathcal{B}}_2^{(0,2)}(\text{s.p.}) = 0, \quad (22.171)$$

$$\mathcal{B}_1^{(2,0)}(\text{s.p.}) = 0. \quad (22.172)$$

**The Set  $\{\mathcal{S}\}$** 

Although the scalar  $\mathcal{S}$  does not appear in the Ward–Takahashi identity, it appears in the Callan–Symanzik equation, so we give its subtraction conditions. In the channel (0, 2),

$$\begin{aligned} \mathcal{S}^{(0,2)}(q; p, \bar{p}) &= \mathcal{S}_1^{(0,2)} + (ip \cdot \gamma + m)\mathcal{S}_2^{(0,2)} + (-i\bar{p} \cdot \gamma + m)\tilde{\mathcal{S}}_2^{(0,2)} \\ &\quad + (ip \cdot \gamma + m)(-i\bar{p} \cdot \gamma + m)\mathcal{S}_3^{(0,2)}. \end{aligned} \quad (22.173)$$

Thus, taking (22.149) as a subtraction point, the subtraction condition is

$$\mathcal{S}_1^{(0,2)}(\text{s.p.}) = 1 . \quad (22.174)$$

In the channel (2, 0),

$$\mathcal{S}_{\rho\sigma}^{(2,0)}(q; k_1, k_2) = [(k_1 \cdot k_2)\delta_{\rho\sigma} - (k_1)_\rho(k_2)_\sigma] \mathcal{S}_1^{(2,0)} , \quad (22.175)$$

and  $\mathcal{S}_1^{(2,0)}$  does not require subtractions.

Since the operators appearing in (22.146) have dimension at most four, their indices are at most zero. The channels which are expected to require subtractions are

$$(0, 2)_1 , \quad (0, 2)_2 , \quad (2, 0)_1 , \quad (4, 0) , \quad (1, 2) .$$

The first three channels have already been analyzed. Moreover, we see that subtractions are not in fact necessary in (4, 0) because of the gauge invariance. In the channel (1, 2), if we introduce the set

$$\mathcal{V}^{(n,m)} = [2m \mathcal{P}^{(n,m)} + \mathcal{W}^{(n,m)}]_{q=0} , \quad (22.176)$$

although we have not discussed it here, the *Kroll–Ruderman theorem* [197] implies that

$$\bar{u}(p)\mathcal{V}_\lambda^{(1,2)}(k, p, \bar{p})u(p) = 0 , \quad k = 0 , \quad p^2 + m^2 = 0 , \quad (22.177)$$

where  $\lambda$  is an index of the electromagnetic potential. From this,  $\mathcal{V}^{(n,m)}$  can be expanded in terms of quantities which vanish for  $q = 0$ . The right-hand side of (22.146) is of course proportional to  $q_\lambda$ . If we examine its coefficient, we see that subtractions are not necessary.

We can therefore apply the theorem given at the end of Sect. 22.5 in this case as well. We investigate the value at each subtraction point of the following linear combination:

$$\mathcal{T}^{(n,m)} = \mathcal{D}^{(n,m)} - a\mathcal{B}^{(n,m)} - 2m\mathcal{P}^{(n,m)} - \mathcal{W}^{(n,m)} . \quad (22.178)$$

Hence,

$$\mathcal{T}_1^{(0,2)}(\text{s.p.}) = 2m\mathcal{P}_1^{(0,2)}(\text{s.p.}) - 0 - 2m\mathcal{P}_1^{(0,2)}(\text{s.p.}) - 0 = 0 , \quad (22.179)$$

$$\mathcal{T}_2^{(0,2)}(\text{s.p.}) = -\mathcal{P}_1^{(0,2)}(\text{s.p.}) - 0 - 2m\mathcal{P}_2^{(0,2)}(\text{s.p.}) - (-1) = 0 , \quad (22.180)$$

$$\mathcal{T}_1^{(2,0)}(\text{s.p.}) = 0 - a - 2m\mathcal{P}_1^{(2,0)}(\text{s.p.}) - 0 = 0 , \quad (22.181)$$

where (22.181) effectively defines the coefficient  $a$  :

$$a = -2m \mathcal{P}_1^{(2,0)}(\text{s.p.}) . \quad (22.182)$$

In this way, all the subtraction constants vanish, and from the theorem in Sect. 22.5,

$$\mathcal{F}^{(n,m)} = 0 . \quad (22.183)$$

This is the proof of (22.146). Thus, it turns out that (22.138) also holds.

### 22.7.2 Ward–Takahashi Identity for $C_\lambda$

From (22.164), we obtain

$$C_\lambda = -z Z_C^{-1} A_\lambda + \frac{i}{4} Z_C^{-1} Z_3 \epsilon_{\lambda\alpha\beta\gamma} \phi_\alpha F_{\beta\gamma} . \quad (22.184)$$

Combining the Ward–Takahashi identity (12.190) in QED with the equation above, we find

$$\begin{aligned} \partial_\mu K_{\mu\rho}(\partial_x) \langle \mathbf{0} | T[\phi_\rho(x) \phi_\sigma(y) C_\lambda(z)] | \mathbf{0} \rangle \\ = \frac{1}{2} Z_C^{-1} Z_3 \epsilon_{\lambda\alpha\beta\gamma} \frac{\partial}{\partial x_\alpha} \delta^4(x-z) \frac{\partial}{\partial y_\beta} G_{\sigma\gamma}^{(2,0)}(y-z) , \end{aligned} \quad (22.185)$$

where  $G^{(2,0)}$  is the Green's function for the electromagnetic field. We note that, using the renormalization condition (22.163),

$$Z_C^{-1} Z_3 = 1 , \quad \text{or} \quad Z_C = Z_3 , \quad (22.186)$$

and the right-hand side of (22.185) is also simplified using this equation.

### 22.7.3 Proof of the Adler–Bardeen Theorem Using the Callan–Symanzik Equation

Recall (20.166) in which the mass-insertion terms are introduced for Green's functions. We now do the same thing for QED. For example,

$$\begin{aligned} A^{(n,m)}(w; x \dots y \dots z \dots; K) \\ = \langle \mathbf{0} | T \left[ A(w) \phi(x) \dots \psi(y) \dots \bar{\psi}(z) \dots \exp \left[ -i K m \int d^4 u S(u) \right] \right] | \mathbf{0} \rangle . \end{aligned} \quad (22.187)$$

The Callan–Symanzik equation is then

$$[\tilde{D} + \gamma^{(n,m)} + \gamma_A]A^{(n,m)}(\dots : K) = 0 . \quad (22.188)$$

Here,  $\gamma_A$  is the anomalous dimension of  $A$  :

$$\tilde{D} = m \frac{\partial}{\partial m} + \beta(e) \frac{\partial}{\partial e} - 2\alpha \gamma_{\text{ph}}(e) \frac{\partial}{\partial \alpha} - \left\{ 1 + [1 - \gamma_S(e)]K \right\} \frac{\partial}{\partial K} . \quad (22.189)$$

Here, for  $A$ , we use  $S$ ,  $P$ ,  $A_\lambda$ ,  $W$ , and so on. From the definition,

$$\gamma_W = 0 . \quad (22.190)$$

Moreover, for  $C_\lambda$ , Eq. (20.158) implies

$$[\tilde{D} + \gamma^{(n,m)} + \gamma_C]C_\lambda^{(n,m)}(\dots : K) + \gamma_M A_\lambda^{(n,m)}(\dots : K) = 0 . \quad (22.191)$$

The quantity  $W^{(n,m)}$  was originally generated by the equal-time commutation relations between  $(A_0 - aC_0)$  and  $\psi$ ,  $\bar{\psi}$ . If there is a mass-insertion term, the commutator of  $(A_0 - aC_0)$  and  $S$  is generated. This is proportional to  $P$ , and if we take its proportionality coefficient as  $b$ , then (22.140) generalizes to

$$\begin{aligned} \partial_\lambda [A_\lambda^{(n,m)}(\dots : K) - aC_\lambda^{(n,m)}(\dots : K)] \\ = 2m(1 + bK)P^{(n,m)}(\dots : K) + W^{(n,m)}(\dots : K) . \end{aligned} \quad (22.192)$$

Similarly, the generalization of (22.185) is

$$\partial_\mu K_{\mu\rho}(\partial_x)C_{\lambda\rho\sigma}^{(2,0)}(x, y, z : K) = \frac{1}{2}\epsilon_{\lambda\alpha\beta\gamma} \frac{\partial}{\partial x_\alpha} \delta^4(x - z) \frac{\partial}{\partial y_\beta} G_{\sigma\gamma}^{(2,0)}(y - z : K) . \quad (22.193)$$

Multiplying (22.192) by  $(\tilde{\mathcal{D}} + \gamma^{(n,m)})$ , a linear combination of  $\{\partial_\lambda A_\lambda\}$ ,  $\{\partial_\lambda C_\lambda\}$ , and  $\{P\}$  is obtained once again, and  $\{W\}$  disappears. Without  $\{W\}$ , the remaining three sets are linearly independent, so all coefficients turn out to disappear. From this, we obtain

$$\left( \beta \frac{d}{de} - \gamma_C \right) a = 0 , \quad (22.194)$$

$$\gamma_A = a\gamma_M , \quad (22.195)$$

$$b = 1 - \gamma_P , \quad (22.196)$$

$$\left( \beta \frac{d}{de} + \gamma_S - \gamma_P \right) b = 0 . \quad (22.197)$$

In particular, combining the last two equations, (20.164) is reproduced. Alternatively,

$$\beta \frac{d}{de} \gamma_P = (1 - \gamma_P)(\gamma_S - \gamma_P). \quad (22.198)$$

We combine (22.193) with the two equations

$$(\tilde{\mathcal{G}} + 2\gamma_{\text{ph}})G_{\sigma\gamma}^{(2,0)}(y - z : K) = 0 \quad (22.199)$$

and

$$(\tilde{\mathcal{G}} + 2\gamma_{\text{ph}} + \gamma_C)C_{\lambda\rho\sigma}^{(2,0)}(x, y, z : K) + \gamma_M A_{\lambda\rho\sigma}^{(2,0)}(x, y, z : K) = 0. \quad (22.200)$$

We thus multiply (22.200) by the differential operator

$$\partial_\mu K_{\mu\rho}(\partial) = \frac{1}{\alpha} \square \partial_\rho. \quad (22.201)$$

Hence,

$$(\tilde{\mathcal{G}} + \gamma_C)\partial_\mu K_{\mu\rho}(\partial_x)C_{\lambda\rho\sigma}^{(2,0)}(x, y, z : K) = 0. \quad (22.202)$$

According to (22.193), since this should be equal to (22.199), we have

$$\gamma_C = 2\gamma_{\text{ph}}. \quad (22.203)$$

Hence, by (20.152), we have  $\beta = e\gamma_{\text{ph}}$ , and (22.194) simplifies to

$$\left( e \frac{d}{de} - 2 \right) a = 0. \quad (22.204)$$

Therefore, the value of  $a$  is strictly given by the lowest order in perturbation theory:

$$a = -\frac{e^2}{2\pi^2}, \quad (22.205)$$

and all higher order corrections disappear. This is the Adler–Bardeen theorem. This proof was given by Higashijima et al. [198].

As we saw above, divergences and indefiniteness do not appear explicitly in the dispersion theory, so it is suitable for treating the delicate issue of anomalous terms.

Regarding these terms, an extremely powerful method using the path integral method was proposed by Fujikawa [199]. This method determines transformations of the path integral measure for fermions under various types of transformations using suitable regularization procedures. The reader is referred to the original work in [199].

## Postface

In this book, we have treated various aspects, focusing on well-established methods in field theory. However, many issues remain unaddressed. These include the Nambu–Bethe–Salpeter equation [200, 201] and the dispersion theory for the  $S$ -matrix and form factors, which are rarely used nowadays. For these subjects, the reader is referred to my book, “Fields and Particles” [2]. Regarding the renormalisation theory for higher order corrections in the Feynman–Dyson theory and the issue of infrared divergences, the reader is referred to the book “Quantum Field Theory” by Noboru Nakanishi [83].



# References

1. K. Nishijima, *Fundamental Particles* (Benjamin, New York, 1963)
2. K. Nishijima, *Fields and Particles* (Benjamin, New York, 1969)
3. P.A.M. Dirac, Quantum theory of emission and absorption of radiation, Proc. R. Soc. Lond. A **114**, 243 (1927)
4. P.A.M. Dirac, The quantum theory of the electron. Part 2, Proc. R. Soc. Lond. A **118**, 351 (1928)
5. W. Heisenberg, W. Pauli, On quantum field theory (In German). Z. Phys. **56**, 1 (1929)
6. W. Heisenberg, W. Pauli, On quantum field theory. 2 (In German), Z. Phys. **59**, 168 (1930)
7. C.D. Anderson, The positive electron. Phys. Rev. **43**, 491 (1933)
8. O. Chamberlain, E. Segrè, C. Wiegand, T. Ypsilantis, Observation of antiprotons, UCRL-3172 (1995), ACC0411, W-7405-eng-48, printed for the US Atomic Energy Commission
9. J. Chadwick, Possible existence of a neutron. Nature **129**, 312 (1932)
10. J. Chadwick, The existence of a neutron. Proc. R. Soc. A **136**, 692 (1932)
11. W. Pauli, Dear radioactive ladies and gentlemen, Phys. Today **31N9**, 27 (1978)
12. W. Pauli, Discussions, in Septième Conseil de Physique Solvay, Noyaux Atomiques, Bruxelles (1933), p. 324
13. E. Fermi, Attempts to produce a theory of beta radiation (In Italian). Nuovo Cim. **11**, 1 (1934)
14. D. Iwanenko, Interaction of neutrons and protons. Nature **133**, 981 (1934)
15. I.G. Tamm, Exchange forces between neutrons and protons, and Fermi's theory. Nature **133**, 981 (1934)
16. I. Tamm, Interaction of neutrons and protons. Nature **134**, 1010 (1934)
17. H. Yukawa, On the interaction of elementary particles, Proc. Phys. Math. Soc. Jap. **17**, 48 (1935)
18. O. Klein, International Institute of Intellectual Cooperation, Paris (1938), in *Proceedings of the Conference New Theories in Physics in Warsaw* (1938)
19. J.C. Street, E.C. Stevenson, New evidence for the existence of a particle of mass intermediate between the proton and electron. Phys. Rev. **52**, 1003 (1937)
20. M. Conversi, E. Pancini, O. Piccioni, On the decay process of positive and negative mesons. Phys. Rev. **68**, 232 (1945)
21. C.M.G. Lattes, H. Muirhead, G.P.S. Occhialini, C.F. Powell, Processes involving charged mesons. Nature **159**, 694 (1947)
22. S. Sakata, T. Inoue, On the correlations between mesons and Yukawa particles. Prog. Theo. Phys. **1**, 143 (1946)
23. Y. Tanikawa, On the cosmic-ray meson and the nuclear meson. Prog. Theor. Phys. **2**, 220 (1947)

24. R.E. Marshak, H.A. Bethe, On the two-meson hypothesis. *Phys. Rev.* **72**, 506 (1947)
25. E. Gardner, C.M.G. Lattes, Production of mesons by the 184-inch Berkeley cyclotron, in *Topics on Cosmic Rays*, ed. by J. Bellandi, vol. 2 (1987), pp. 110–111. [*Science* **107** (1948) 270–271]
26. H.L. Anderson, E. Fermi, E.A. Long, D.E. Nagle, Total cross-sections of positive pions in hydrogen. *Phys. Rev.* **85**, 936 (1952)
27. N. Kemmer, The charge-dependence of nuclear forces. *Proc. Camb. Philos. Soc.* **34**, 354 (1938)
28. G.F. Chew, F.E. Low, Effective range approach to the low-energy p wave pion–nucleon interaction. *Phys. Rev.* **101**, 1570 (1956)
29. S. Mandelstam, Determination of the pion–nucleon scattering amplitude from dispersion relations and unitarity. General theory. *Phys. Rev.* **112**, 1344 (1958)
30. T. Regge, Introduction to complex orbital momenta. *Nuovo Cim.* **14**, 951 (1959)
31. G.F. Chew, S.C. Frautschi, Regge trajectories and the principle of maximum strength for strong interactions. *Phys. Rev. Lett.* **8**, 41 (1962)
32. G.D. Rochester, C.C. Butler, Evidence for the existence of new unstable elementary particles. *Nature* **160**, 855 (1947)
33. A.J. Seriff, R.B. Leighton, C. Hsiao, E.W. Cowan, C.D. Anderson, Cloud-chamber observations of the new unstable cosmic ray. *Phys. Rev.* **78**, 209 (1950)
34. R. Armenteros, K.H. Barker, C.C. Butler, A. Cachon, A.H. Chapman, Decay of V-particles. *Nature* **167**, 501 (1951)
35. Y. Nambu, K. Nishijima, Y. Yamaguchi, On the nature of V-particles. I. *Prog. Theor. Phys.* **6**, 615–619 (1951). In Eguchi, T. and Nishijima, K. (eds.): *Broken symmetry* 47–51
36. Y. Nambu, K. Nishijima, Y. Yamaguchi, On the nature of V-particles. II. *Prog. Theor. Phys.* **6**, 619–622 (1951). In Eguchi, T. and Nishijima, K. (eds.): *Broken symmetry* 51–54
37. S. Oneda, About V-particles and  $\tau$ -mesons, in Japanese. *Soryushiron Kenkyu* **3**(4), 271 (1951)
38. H. Miyazawa, A model of V-particles. *Prog. Theor. Phys.* **6**, 631 (1951)
39. A. Pais, Some remarks on the V-particles. *Phys. Rev.* **86**, 663 (1952)
40. W.B. Fowler, R.P. Shutt, A.M. Thorndike, W.L. Whittemore, Production of  $V_1^0$  particles by negative pions in hydrogen. *Phys. Rev.* **91**, 1287 (1953)
41. T. Nakano, K. Nishijima, Charge independence for V-particles. *Prog. Theor. Phys.* **10**, 581 (1953)
42. M. Gell-Mann, Isotopic spin and new unstable particles. *Phys. Rev.* **92**, 833 (1953)
43. R.H. Dalitz, On the analysis of  $\tau$ -meson data and the nature of the  $\tau$ -meson. *Philos. Mag.* **44**, 1068 (1953)
44. E. Fabri, A study of tau-meson decay. *Nuovo Cim.* **11**, 479 (1954)
45. T.D. Lee, C.N. Yang, Question of parity conservation in weak interactions. *Phys. Rev.* **104**, 254 (1956)
46. C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes, R.P. Hudson, Experimental test of parity conservation in beta decay. *Phys. Rev.* **105**, 1413 (1957)
47. T.D. Lee, C.N. Yang, Parity nonconservation and a two-component theory of the neutrino. *Phys. Rev.* **105**, 1671 (1957)
48. E.C.G. Sudarshan, R.E. Marshak, The nature of the four-fermion interaction, in *Proc. of the Conference on Mesons and Newly-Discovered Particles, Padua-Venice, September* (1957)
49. R.P. Feynman, M. Gell-Mann, Theory of the Fermi interaction. *Phys. Rev.* **109**, 193 (1958)
50. J.J. Sakurai, Mass reversal and weak interactions. *Nuovo Cim.* **7**, 649 (1958)
51. S.S. Gershtein, Y.B. Zeldovich, Meson corrections in the theory of beta decay. *Zh. Eksp. Teor. Fiz.* **29**, 698 (1955)
52. M. Gell-Mann, Symmetries of baryons and mesons. *Phys. Rev.* **125**, 1067 (1962)
53. M. Gell-Mann, The symmetry group of vector and axial vector currents. *Physics* **1**, 63 (1964)
54. S.L. Adler, Calculation of the axial vector coupling constant renormalization in beta decay. *Phys. Rev. Lett.* **14**, 1051 (1965)
55. W.I. Weisberger, Renormalization of the weak axial vector coupling constant. *Phys. Rev. Lett.* **14**, 1047 (1965)

56. W.I. Weisberger, Unsubtracted dispersion relations and the renormalization of the weak axial vector coupling constants. *Phys. Rev.* **143**, 1302 (1966)
57. N. Cabibbo, Unitary symmetry and leptonic decays. *Phys. Rev. Lett.* **10**, 531 (1963)
58. C.L. Cowan, F. Reines, F.B. Harrison, H.W. Kruse, A.D. McGuire, Detection of the free neutrino: A confirmation. *Science* **124**, 103 (1956)
59. E.J. Konopinski, H.M. Mahmoud, The universal Fermi interaction. *Phys. Rev.* **92**, 1045 (1953)
60. K. Nishijima, Vanishing of the neutrino rest mass. *Phys. Rev.* **108**, 907 (1957)
61. J.S. Schwinger, A theory of the fundamental interactions. *Ann. Phys.* **2**, 407 (1957)
62. G. Danby, J.M. Gaillard, K.A. Goulianos, L.M. Lederman, N.B. Mistry, M. Schwartz, J. Steinberger, Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Phys. Rev. Lett.* **9**, 36 (1962)
63. E. Fermi, C.N. Yang, Are mesons elementary particles? *Phys. Rev.* **76**, 1739 (1949)
64. S. Sakata, On a composite model for the new particles, *Prog. Theor. Phys.* **16**, 686 (1956)
65. M. Ikeda, S. Ogawa, Y. Ohnuki, A possible symmetry in Sakata's model for bosons-baryons system. *Prog. Theor. Phys.* **22**, 715 (1959)
66. A. Pevsner et al., Evidence for a three pion resonance near 550 MeV. *Phys. Rev. Lett.* **7**, 421 (1961)
67. M. Gell-Mann, The Eightfold Way: A theory of strong interaction symmetry, CTSL-20, TID-12608
68. Y. Ne'eman, Derivation of strong interactions from a gauge invariance. *Nucl. Phys.* **26**, 222 (1961)
69. M. Gell-Mann, A theory of strong interaction symmetry, Caltech Report CTLS-20 (1961)
70. S. Okubo, Note on unitary symmetry in strong interactions. *Prog. Theor. Phys.* **27**, 949 (1962)
71. V.E. Barnes et al., Observation of a hyperon with strangeness minus three. *Phys. Rev. Lett.* **12**, 204 (1964)
72. G.F. Chew, S.C. Frautschi, Principle of equivalence for all strongly interacting particles within the S matrix framework. *Phys. Rev. Lett.* **7**, 394 (1961)
73. G.F. Chew, Nuclear democracy and bootstrap dynamics, UCRL-11163
74. M. Gell-Mann, A schematic model of baryons and mesons. *Phys. Lett.* **8**, 214 (1964)
75. G. Zweig, An  $SU(3)$  model for strong interaction symmetry and its breaking. Version 1, CERN-TH-401
76. M.Y. Han, Y. Nambu, Three triplet model with double  $SU(3)$  symmetry. *Phys. Rev.* **139**, B1006 (1965)
77. Y. Nambu, A systematics of hadrons in subnuclear physics, in *Preludes in Theoretical Physics*, ed. by A. de Shalit, H. Feshbach, L. Van Hove (North Holland, Amsterdam, 1966)
78. K. Nishijima, *Relativistic Quantum Mechanics* (New Physics Series 13) (Baifukan, 1973) (in Japanese)
79. M. Goldhaber, L. Grodzins, A.W. Sunyar, Helicity of neutrinos. *Phys. Rev.* **109**, 1015 (1958)
80. W. Pauli, P. Jordan, Zur Quantenelektrodynamik ladungsfreier Felder. *Zeits. Phys.* **47**, 151 (1928)
81. M. Fierz, Force-free particles with any spin. *Helv. Phys. Acta* **12**, 3 (1939)
82. W. Pauli, The connection between spin and statistics. *Phys. Rev.* **58**, 716 (1940)
83. N. Nakanishi, *Quantum Field Theory* (New Physics Series 19) (Baifukan, 1975) (in Japanese)
84. P.A.M. Dirac, Bakerian lecture. The physical interpretation of quantum mechanics. *Proc. R. Soc. A* **180**, 1 (1942)
85. S.N. Gupta, Theory of longitudinal photons in quantum electrodynamics. *Proc. Phys. Soc. A* **63**, 681 (1950)
86. K. Bleuler, A new method of treatment of the longitudinal and scalar photons. *Helv. Phys. Acta* **23**, 567 (1950)
87. S. Tomonaga, On a relativistically invariant formulation of the quantum theory of wave fields. *Prog. Theor. Phys.* **1**, 27 (1946)
88. J.S. Schwinger, Quantum electrodynamics. I. A covariant formulation. *Phys. Rev.* **74**, 1439 (1948)
89. P.A.M. Dirac, Relativistic quantum mechanics. *Proc. R. Soc. A* **136**, 453 (1932)

90. C.N. Yang, D. Feldman, The S matrix in the Heisenberg representation. *Phys. Rev.* **79**, 972 (1950)
91. Y. Takahashi, *Introduction to Analytical Mechanics for Learning Quantum Mechanics*, Enlarged second edition, (Kodansha Scientific, 2000) (in Japanese)
92. F.J. Belinfante, On the current and the density of the electric charge, the energy, the linear momentum and the angular momentum of arbitrary fields. *Physica* **7**, 449 (1940)
93. L. Rosenfeld, Sur le tenseur d'impulsion-énergie. *Acad. R. Belg. Mémoires de classes de Science* **18** (1940)
94. W. Pauli, On the conservation of the lepton charge. *Nuovo Cimento* **6**, 204 (1957)
95. F. Gürsey, Relation of charge independence and baryon conservation to Pauli's transformation. *Nuovo Cimento* **7**, 411 (1958)
96. R.P. Feynman, Space-time approach to nonrelativistic quantum mechanics. *Rev. Mod. Phys.* **20**, 367 (1948)
97. F.J. Dyson, The Radiation theories of Tomonaga, Schwinger, and Feynman. *Phys. Rev.* **75**, 486 (1949)
98. F.J. Dyson, The S matrix in quantum electrodynamics. *Phys. Rev.* **75**, 1736 (1949)
99. W.H. Furry, A symmetry theorem in the positron theory. *Phys. Rev.* **51**, 125 (1937)
100. R.P. Feynman, The theory of positrons. *Phys. Rev.* **76**, 749 (1949)
101. L.D. Landau, On the angular momentum of a two-photon system. *Dokl. Akad. Nauk Ser. Fiz.* **60**, 207 (1948)
102. C.N. Yang, Possible experimental determination of whether the neutral meson is scalar or pseudoscalar. *Phys. Rev.* **77**, 722 (1950)
103. C. Møller, Zur Theorie des Durchgangs schneller Elektronen durch Materie. *Ann. Physik* **14**, 531 (1932)
104. O. Klein, Y. Nishina, Über die Streuung von Strahlung durch freie Elektronen nach der neuen relativistischen Quantendynamik von Dirac. *Zeitsch. Physik* **52**, 853 (1929)
105. J.D. Bjorken, Asymptotic sum rules at infinite momentum. *Phys. Rev.* **179**, 1547 (1969)
106. L. Michel, Interaction between four half-spin particles and the decay of the  $\mu$ -meson. *Proc. Phys. Soc. A* **63**, 514 (1950)
107. R.H. Dalitz, Decay of  $\tau$  mesons of known charge. *Phys. Rev.* **94**, 1046 (1954)
108. J.H. Christenson, J.W. Cronin, V.L. Fitch, R. Turlay, Evidence for the  $2\pi$  decay of the  $K_2^0$  meson. *Phys. Rev. Lett.* **13**, 138 (1964)
109. G. Lüders, On the equivalence of invariance under time reversal and under particle-antiparticle conjugation for relativistic field theories. *Kong. Dan. Vid. Sel. Mat. Fys. Med.* **28N5**, 1 (1954)
110. W. Pauli, *Niels Bohr and the Development of Physics*. Essays Dedicated to Niels Bohr on the Occasion of his Seventieth birthday (Pergamon Press, London, 1955)
111. M. Gell-Mann, F. Low, Bound states in quantum field theory. *Phys. Rev.* **84**, 350 (1951)
112. P.T. Matthews, The application of Dyson's methods to meson interactions. *Phys. Rev.* **76**, 684 (1949)
113. H. Lehmann, K. Symanzik, W. Zimmermann, On the formulation of quantized field theories. *Nuovo Cim.* **1**, 205 (1955)
114. B.A. Lippmann, J. Schwinger, Variational principles for scattering processes. I. *Phys. Rev.* **79**, 469 (1950)
115. G.F. Chew, M.L. Goldberger, The scattering of elementary particles by complex nuclei. A generalization of the impulse approximation. *Phys. Rev.* **87**, 778 (1952)
116. J.S. Schwinger, On quantum electrodynamics and the magnetic moment of the electron. *Phys. Rev.* **73**, 416 (1948)
117. H. Fukuda, Y. Miyamoto, S. Tomonaga, A self-consistent subtraction method in the quantum field theory. II. *Prog. Theor. Phys.* **2**, 121 (1949)
118. J.C. Ward, An identity in quantum electrodynamics. *Phys. Rev.* **78**, 182 (1950)
119. Y. Takahashi, On the generalized Ward identity. *Nuovo Cim.* **6**, 371 (1957)
120. K. Nishijima, Asymptotic conditions and perturbation theory. *Phys. Rev.* **119**, 485 (1960)

121. H. Umezawa, S. Kamefuchi, The vacuum in quantum electrodynamics. *Prog. Theor. Phys.* **6**, 543 (1951)
122. G. Källén, On the definition of the renormalization constants in quantum electrodynamics. *Helv. Phys. Acta* **25**(4), 417 (1952)
123. H. Lehmann, Über Eigenschaften von Ausbreitungsfunktionen und Renormierungskonsten quantisierter Felder. *Nuovo Cim.* **11**, 342 (1954)
124. T. Goto, T. Imamura, Note on the non-perturbation-approach to quantum field theory. *Prog. Theor. Phys.* **14**, 396 (1955)
125. J.S. Schwinger, Field theory commutators. *Phys. Rev. Lett.* **3**, 296 (1959)
126. M.L. Perl et al., Evidence for anomalous lepton production in  $e^+e^-$  annihilation. *Phys. Rev. Lett.* **35**, 1489 (1975)
127. H. Frauenfelder et al., Parity and the polarization of electrons from  $^{60}\text{Co}$ . *Phys. Rev.* **106**, 386 (1957)
128. H. Frauenfelder et al., Parity and electron polarization: Møller scattering. *Phys. Rev.* **107**, 643 (1957)
129. M. Gell-Mann, M. Levy, The axial vector current in beta decay. *Nuovo Cim.* **16**, 705 (1960)
130. S.L. Glashow, J. Iliopoulos, L. Maiani, Weak interactions with lepton-hadron symmetry. *Phys. Rev. D* **2**, 1285 (1970)
131. M. Kobayashi, T. Maskawa, CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.* **49**, 652 (1973)
132. J.J. Aubert et al. [E598 Collaboration], Experimental observation of a heavy particle. *J. Phys. Rev. Lett.* **33**, 1404 (1974)
133. J.E. Augustin et al. [SLAC-SP-017 Collaboration], Discovery of a narrow resonance in  $e^+e^-$  annihilation. *Phys. Rev. Lett.* **33**, 1406 (1974). [*Adv. Exp. Phys.* **5** (1976) 141]
134. Z. Maki, The fourth baryon, Sakata model and modified  $B - L$  symmetry. I. *Prog. Theor. Phys.* **31**, 331 (1964)
135. Y. Hara, Unitary triplets and the eightfold way. *Phys. Rev.* **134**, B701 (1964)
136. S.W. Herb et al., Observation of a dimuon resonance at 9.5 GeV in 400 GeV proton-nucleus collisions. *Phys. Rev. Lett.* **39**, 252 (1977)
137. R.P. Feynman, The behavior of hadron collisions at extreme energies. *Conf. Proc. C* **690905**, 237 (1969)
138. C.G. Callan Jr., D.J. Gross, High-energy electroproduction and the constitution of the electric current. *Phys. Rev. Lett.* **22**, 156 (1969)
139. C.N. Yang, R.L. Mills, Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.* **96**, 191 (1954)
140. R. Utiyama, Invariant theoretical interpretation of interaction. *Phys. Rev.* **101**, 1597 (1956)
141. Y. Nambu, G. Jona-Lasinio, Dynamical model of elementary particles based on an analogy with superconductivity. I. *Phys. Rev.* **122**, 345 (1961)
142. Y. Nambu, G. Jona-Lasinio, Dynamical model of elementary particles based on an analogy with superconductivity. II. *Phys. Rev.* **124**, 246 (1961)
143. J. Goldstone, Field theories with superconductor solutions. *Nuovo Cim.* **19**, 154 (1961)
144. F. Englert, R. Brout, Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.* **13**, 321 (1964)
145. P.W. Higgs, Broken symmetries, massless particles and gauge fields. *Phys. Lett.* **12**, 132 (1964)
146. P.W. Higgs, Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.* **13**, 508 (1964)
147. P.W. Higgs, Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.* **145**, 1156 (1966)
148. N. Nakanishi, Indefinite-metric quantum theory of genuine and Higgs-type massive vector fields. *Prog. Theor. Phys.* **49**, 640 (1973)
149. P.W. Anderson, Plasmons, gauge invariance, and mass. *Phys. Rev.* **130**, 439 (1963)
150. T.W.B. Kibble, Symmetry breaking in non-Abelian gauge theories. *Phys. Rev.* **155**, 1554 (1967)

151. S. Weinberg, A model of leptons. *Phys. Rev. Lett.* **19**, 1264 (1967)
152. A. Salam, Weak and electromagnetic interactions. *Conf. Proc. C* **680519**, 367 (1968)
153. S. Glashow, Partial-symmetries of weak interactions. *Nucl. Phys.* **22**(4), 579–588 (1961)
154. G. Arnison et al. [UA1 Collaboration], Experimental observation of isolated large transverse energy electrons with associated missing energy at  $s^{1/2} = 540$  GeV. *Phys. Lett. B* **122**, 103 (1983)
155. M. Banner et al. [UA2 Collaboration], Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN anti-p p collider. *Phys. Lett. B* **122**, 476 (1983)
156. G. Arnison et al. [UA1 Collaboration], Experimental observation of lepton pairs of invariant mass around 95 GeV/c<sup>2</sup> at the CERN SPS collider. *Phys. Lett. B* **126**, 398 (1983)
157. N.S. Craigie, K. Hidaka, P. Ratcliffe, The role helicity asymmetries could play in the search for supersymmetric interactions. *Phys. Lett. B* **129**, 310 (1983)
158. F.J. Hasert et al. [Gargamelle Neutrino Collaboration], Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment. *Phys. Lett. B* **46**, 138 (1973)
159. S.L. Adler, Axial vector vertex in spinor electrodynamics. *Phys. Rev.* **177**, 2426 (1969)
160. J.S. Bell, R. Jackiw, A PCAC puzzle:  $\pi^0 \rightarrow \gamma\gamma$  in the sigma model. *Nuovo Cim. A* **60**, 47 (1969)
161. W.A. Bardeen, Anomalous Ward identities in spinor field theories. *Phys. Rev.* **184**, 1848 (1969)
162. H. Georgi, S.L. Glashow, Gauge theories without anomalies. *Phys. Rev. D* **6**, 429 (1972)
163. H. Georgi, S.L. Glashow, Unity of all elementary particle forces. *Phys. Rev. Lett.* **32**, 438 (1974)
164. L.D. Faddeev, V.N. Popov, Feynman diagrams for the Yang–Mills field. *Phys. Lett. B* **25**, 29 (1967)
165. G. 't Hooft, Renormalization of massless Yang–Mills fields. *Nucl. Phys. B* **33**, 173 (1971)
166. T. Kugo, I. Ojima, Manifestly covariant canonical formulation of Yang–Mills field theories. 1. The case of Yang–Mills fields of Higgs–Kibble type in the Landau gauge. *Prog. Theor. Phys.* **60**, 1869 (1978)
167. C. Becchi, A. Rouet, R. Stora, Renormalization of gauge theories. *Ann. Phys.* **98**, 287 (1976)
168. G. Curci, R. Ferrari, Slavnov transformations and supersymmetry. *Phys. Lett. B* **63**, 91 (1976)
169. I. Ojima, Another BRS transformation. *Prog. Theor. Phys.* **64**, 625 (1980)
170. L. Bonora, M. Tonin, Superfield formulation of extended BRS symmetry. *Phys. Lett. B* **98**, 48 (1981)
171. A.A. Slavnov, Invariant regularization of gauge theories. *Teor. Mat. Fiz.* **13**, 174 (1972)
172. J.C. Taylor, Ward identities and charge renormalization of the Yang–Mills field. *Nucl. Phys. B* **33**, 436 (1971)
173. T. Kugo, I. Ojima, Local covariant operator formalism of nonabelian gauge theories and quark confinement problem. *Prog. Theor. Phys. Suppl.* **66**, 1 (1979)
174. T. Kugo, I. Ojima, Manifestly covariant canonical formulation of Yang–Mills field theories: Physical state subsidiary conditions and physical S matrix unitarity. *Phys. Lett. B* **73**, 459 (1978)
175. N. Nakanishi, I. Ojima, Superalgebras of nonabelian gauge theories in the manifestly covariant canonical formalism. *Z. Phys. C* **6**, 155 (1980). [*Z. Phys. C* **8** (1981) 94]
176. E.C.G. Stueckelberg, A. Petermann, La normalisation des constantes dans la théorie des quanta (Normalization of constants in the quantum theory). *Helv. Phys. Acta* **26**, 499 (1953)
177. M. Gell-Mann, F.E. Low, Quantum electrodynamics at small distances. *Phys. Rev.* **95**, 1300 (1954)
178. N.N. Bogolyubov, D.V. Shirkov, Introduction to the theory of quantized fields. *Intersci. Monogr. Phys. Astron.* **3**, 1 (1959)
179. C.G. Callan Jr., Broken scale invariance in scalar field theory. *Phys. Rev. D* **2**, 1541 (1970)
180. K. Symanzik, Small distance behavior in field theory and power counting. *Commun. Math. Phys.* **18**, 227 (1970)

181. K. Symanzik, Small distance behavior analysis and Wilson expansion. *Commun. Math. Phys.* **23**, 49 (1971)
182. K. Higashijima, K. Nishijima, Renormalization groups of Gell-Mann–Low and of Callan–Symanzik. *Prog. Theor. Phys.* **64**, 2179 (1980)
183. D.J. Gross, F. Wilczek, Ultraviolet behavior of nonabelian gauge theories. *Phys. Rev. Lett.* **30**, 1343 (1973)
184. H.D. Politzer, Reliable perturbative results for strong interactions? *Phys. Rev. Lett.* **30**, 1346 (1973)
185. K.G. Wilson, Confinement of quarks. *Phys. Rev. D* **10**, 2445 (1974)
186. K. Nishijima, BRS transformation and color confinement. *Phys. Lett. B* **116**, 295 (1982)
187. R. Haag, Quantum field theories with composite particles and asymptotic conditions. *Phys. Rev.* **112**, 669 (1958)
188. W. Zimmermann, On the bound state problem in quantum field theory. *Nuovo Cim.* **10**, 597 (1958)
189. K. Nishijima, Formulation of field theories of composite particles. *Phys. Rev.* **111**, 995 (1958)
190. M. Chaichian, K. Nishijima, Renormalization constant of the color gauge field as a probe of confinement. *Eur. Phys. J. C* **22**, 463 (2001)
191. K. Nishijima, A. Tureanu, Gauge-dependence of Green’s functions in QCD and QED. *Eur. Phys. J. C* **53**, 649 (2008)
192. H. Fukuda, Y. Miyamoto, On the  $\gamma$ -decay of neutral meson. *Prog. Theor. Phys.* **4**, 374 (1949)
193. J. Steinberger, On the use of subtraction fields and the lifetimes of some types of meson decay. *Phys. Rev.* **76**, 1180 (1949)
194. J.S. Schwinger, On gauge invariance and vacuum polarization. *Phys. Rev.* **82**, 664 (1951)
195. C.G. Callan Jr., S.R. Coleman, R. Jackiw, A new improved energy–momentum tensor. *Ann. Phys.* **59**, 42 (1970)
196. S.L. Adler, W.A. Bardeen, Absence of higher order corrections in the anomalous axial vector divergence equation. *Phys. Rev.* **182**, 1517 (1969)
197. N.M. Kroll, M.A. Ruderman, A theorem on photomeson production near threshold and the suppression of pairs in pseudoscalar meson theory. *Phys. Rev.* **93**, 233 (1954)
198. K. Higashijima, K. Nishijima, M. Okawa, The Adler–Bardeen theorem in quantum electrodynamics. *Prog. Theor. Phys.* **67**, 668 (1982)
199. K. Fujikawa, Path integral measure for gauge invariant fermionic theories. *Phys. Rev. Lett.* **42**, 1195 (1979)
200. E.E. Salpeter, H.A. Bethe, A relativistic equation for bound-state problems. *Phys. Rev.* **84**, 1232 (1951)
201. Y. Nambu, Force potentials in quantum field theory. *Prog. Theor. Phys.* **5**, 614–633 (1950). In Eguchi, T. and Nishijima, K. (eds.): Broken symmetry 27–46; Path integral for gauge theories with fermions, *Phys. Rev. D* **21** (1980) 2848 [*Phys. Rev. D* **22** (1980) 1499]

# Index

## A

Action integral, 411  
Additional condition, 93  
Additional term method, 423  
Adjoint representation, 380  
Adler–Bardeen theorem, 548  
Adler–Weisberger formula, 10, 340  
Advanced  $\Delta$ -function, 114  
Annihilation operator, 32  
Anomalous magnetic moment, 294  
Anomalous term, 397  
    trace anomaly term, 531  
    triangle anomaly term, 528, 548  
Anomalous trace identity, 543  
Anti-commutation relation, 51  
Antiparticle, 34, 212  
Antiunitary transformation, 217  
Asymptotic condition, 248  
Asymptotic field, 248  
Asymptotic freedom, 499  
Axial gauge, 434

## B

Baryon number, 314  
Belinfante–Rosenfeld tensor, 136  
Beta-decay, 2, 332  
Bianchi identity, 356, 358, 359  
Bjorken scaling law, 196, 350, 502  
Bra, 81  
Branching ratio, 191  
BRS algebra, 450  
    for auxiliary fields, 459  
    extended, 457  
    representation of, 450

BRS charge, 442  
BRS doublet, 451  
BRS identity, 446  
BRSNO algebra, 462  
BRS singlet, 451  
BRS transformation, 439  
    another, 444  
Bubble graph, 165

## C

Cabibbo's theory, 341, 342  
Callan–Symanzik equation, 477, 485, 555  
    homogeneous, 491  
Canonical commutation relation, 16, 22  
Canonical number, 383  
Casimir operator, 54, 138, 329  
Cauchy's theorem, 266  
Cayley–Hamilton theorem, 84  
Characteristic polynomial, 84  
Charge conjugation, 34, 55, 174, 209  
Charge independence, 5, 117  
Chew–Frautschi plot, 7  
Chew–Goldberger formal solution, 267  
Chiral transformation, 148  
Christoffel symbol, 358  
Classical electron radius, 190  
Colour, 343, 400  
Complex scalar field, 32  
Complex vector field, 59  
Compton scattering, 166  
Confinement, 509  
    colour, 512  
Conservation of current density, 116  
Conserved vector current (CVC), 9



Continuity equation, 34, 36  
 Contraction, 160  
 Covariant derivative, 355, 357, 358  
 CPT theorem, 221  
 Creation operator, 32  
 Cross-section, 182  
 Current algebra, 9, 339, 366  
 CVC hypothesis, 338

**D**

Dalitz plot, 202, 205  
 Decomposition theorem, 86  
 Degeneracy, 83  
 $\Delta$  (resonance state), 5  
 Deuteron, 120  
 Dipole ghost, 101  
 Dirac's equation, 34  
 Direct product representation, 311  
 Discrete symmetry, 211  
 Discrete unitary transformation, 209  
 Dispersion relation, 5, 527, 530  
 Dispersion theory, 533  
 Dyson's formula, 152

**E**

Effective potential, 369  
 Eightfold way, 11, 324  
 Einstein relation, 28  
 Electron–positron collision, 499  
 Energy–momentum tensor, 131  
   canonical, 131  
   improved, 545  
   symmetric, 136  
 Energy spectrum, 202  
 Euler–Lagrange equation, 17, 24  
 External lines, 162

**F**

Faddeev–Popov ghost field, 79, 430  
 Fermi gauge, 91, 161  
 Fermi interaction, 3, 199  
 Fermi's method, 92  
 Fermi–Yang model, 11  
 Feynman diagrams, 161  
 Feynman–Dyson theory, 172  
 Feynman propagator, 160  
 Feynman's formula, 172  
 Flavours, 343  
   number of, 405  
 Formula for decay width, 182  
 Functional derivative, 108

Furry's theorem, 174, 215

**G**

Gamma-matrices, 39  
 Gauge condition, 422, 424  
 Gauge fields, 13, 353  
 Gauge transformation, 115  
   of electromagnetic field, 354  
   of Green's function, 432  
   local, 355  
 G-conjugation, 216  
 Gell-Mann–Low relation, 229  
 Generating functional, 229, 368  
 Generation, 4, 7, 10, 393  
 Generator of the Lorentz transformation, 135  
 Generic transform function, 109  
 Georgi–Glashow  $SU(5)$  model, 402  
 Ghost number, 444  
 GIM mechanism, 394  
 Gluon, 13  
 Goto–Imamura–Schwinger term, 306, 341, 518, 528  
 Grand unified theory (GUT), 402, 403  
 Gravitational field, 358  
 Green's function, 229  
   connected, 232, 368  
   gauge dependence of, 504  
   one-particle irreducible, 368  
 Group of phase transformations, 145

**H**

Hadron, 6, 309  
 Hamilton's principle, 17  
 Heisenberg picture, 15  
 Helicity, 46  
   of  $v$ , 49  
   of  $e^-$ , 335  
 Higgs mechanism, 371  
 Higgs particle, 375, 388  
 Hypercharge, 314

**I**

Inclusive reaction, 192  
 Indefinite metric, 81  
 Indefiniteness, 525  
 Inner product, 81  
 Integrability condition, 109, 154  
 Integral representation of propagator, 299  
   of electromagnetic field, 306  
 Intermediate bosons, W, Z, 391  
 Invariant functions, 68

Invariant subspace, 84  
 Irreducibility, 59  
 Irreducible representation, 311  
 Isospin, 5, 117, 134, 216  
 Isospin parity, 215  
 Isospin space, 5, 119

**J**

Jacobi identity, 19, 356  
 Jordan–Pauli invariant  $\Delta$ -function, 66

**K**

Ket, 81  
 Kibble’s theorem, 383  
 Klein–Gordon equation, 25  
 Klein–Nishina formula, 189  
 Kobayashi–Maskawa matrix, 344  
 Kroll–Foldy theorem, 215

**L**

Lagrange equation, 17  
 Landau gauge, 426  
 Legendre transformation, 233  
 Lepton, 7  
 Lepton number conservation, 10  
 Lie algebra, 313  
 Lie group, 311  
 Linear operator, 82  
 Lippmann–Schwinger equation, 267  
 Local commutativity, 68  
 Lorentz transformation, 38, 134  
 Lorenz condition, 97

**M**

Magnetic moment, 117, 326, 347  
 Mass formula, 325, 345  
 Mass insertion operator, 487  
 Matthew’s theorem, 236, 281  
 Maxwell equations, 143  
 Meson field, 3  
 Metric matrix, 83  
 Michel parameter, 200, 202  
 Microscopic causality, 68, 109  
 Minkowski metric, 96  
 Møller’s formula, 183  
 Multi-time formulation, 107  
 Multipole state, 88  
 Muon, 4

**N**

Nambu–Goldstone particle, 362, 367, 370, 386  
 Nambu–Goldstone theorem, 362  
 Nambu–Jona-Lasinio model, 362  
 Natural units, 15  
 Negative energy solutions, 48  
 Neutrino, 2
 

- helicity of, 49
- two-component theory of, 9, 45
- two-types of, 10

Nilpotency, 443  
 Nilpotent operator, 84  
 Noether’s theorem, 127, 130, 132  
 Non-Abelian gauge theory, 356  
 Non-conservation of CP, 214  
 Non-conservation of parity, 8, 207  
 Norm, 87  
 Normal product, 155  
 Nuclear force, 2, 117  
 Nucleon decay, 404

**O**

Observable, 90, 98  
 Optical theorem, 196, 271  
 Orthogonal group  $O(n)$ , 311  
 Ovsianikov’s equation, 479

**P**

Pair-production theory, 8  
 Parameter dispersion relation, 534  
 Parity
 

- intrinsic, 211
- non-conservation, 8, 207

Partial wave expansion, 196  
 Particle number operator, 32  
 Parton model, 348, 499  
 Path integral, 409  
 Pauli–Gürsey group, 144  
 Pauli–Liubanski operator, 137  
 Pauli metric, 25  
 Pauli principle, 117  
 Pauli representation, 37  
 Pauli term, 294  
 PCAC hypothesis, 340, 367  
 Pion ( $\pi$ -meson), 4, 119
 

- decay, 167
- parity of, 122
- spin of, 120

Planck mass, 403  
 Poincaré transformation, 134  
 Polarization, 120  
 Polarization vector, 94

Positronium, 214  
 Principle of detailed balance, 120, 220  
 Principle of least action, 21  
 Probability amplitude, 15  
 Projection operator, 84, 86  
 Proper Lorentz transformation, 27, 140

## Q

Quantum chromodynamics (QCD), 6, 402  
 Quantum electrodynamics (QED), 4, 114, 470  
 Quark model, 12, 314, 341, 344

## R

Rank, 313  
 Real scalar field, 27, 65  
 Recombination reaction, 272  
 Reduction formula, 161, 247  
   LSZ, 251  
 Regge trajectory, 6  
 Renormalizability, 274  
 Renormalization, 4  
   of charge, 276  
   condition, 274  
   of mass, 275, 276  
   multiplicative, 280  
   of vertex function, 288  
 Renormalization constant, 295  
   of field operator, 279  
 Renormalization group, 469  
   for non-Abelian gauge theory, 495  
   linear equation for, 483  
 Renormalized interaction picture, 281  
 Renormalized propagator, 282  
 Representation of group, 310  
 Representations of the Lorentz group, 139  
 Retarded  $\Delta$ -function, 75  
 Retarded product, 112

## S

Sakata's model, 11, 211, 314  
 Schrödinger picture, 15  
 Self-energy, 125, 302  
 Sigma model, 364  
 Slavnov–Taylor Identity, 446  
 $S$ -matrix, 5, 110, 151  
 Space inversion, 43, 210  
 Space-like hypersurface, 107  
 Special unitary group  $SU(2)$ , 313  
 Special unitary group  $SU(3)$ , 11, 314  
 Spectral function, 301  
 Spin–statistics theorem, 78  
 Spinor, 46

Static approximation, 122  
 Strange particle, 7  
 Strangeness, 7, 314  
 Structure function, 195, 350  
 Stueckelberg causality, 263  
 Subtraction condition, 538  
 Supermultiple-time formulation, 108  
 Symmetry, 361  
   hidden, 361  
   manifest, 361

## T

$\theta$ – $\tau$  puzzle, 9, 207  
 Thomson's formula, 189  
 Time-ordering operator, 153, 234  
 Tomonaga–Schwinger equation, 105  
 Trace, 170  
 Two-meson theory, 4  
 Two-photon decay of  $\pi^0$ -meson, 168, 400  
 Two-photon decay of neutral mesons, 175

## U

Unimodular, 44  
 Unitarity, 151, 255  
 Unitarity gauge, 377  
 Unitary group  $U(n)$ , 312  
 Unitary transformation, 123, 209  
 Universality, 330, 336

## V

$V - A$  theory, 9, 393  
 Variation, 24, 127  
 Vector space, 81

## W

Ward identity, 295  
 Ward–Takahashi identity, 297, 426, 543, 545, 555  
 Weak isospin, 390  
 Weinberg angle, 392  
 Weinberg–Salam model, 389  
 Weyl equation, 37, 142  
 Wick's theorem, 154

## Y

Yang–Feldman equation, 114  
 Yang–Mills field, 354  
 Yukawa interaction, 122  
 Yukawa potential, 125