Graduate Texts in Physics

Edouard B. Manoukian

Quantum Field Theory I

Foundations and Abelian and Non-Abelian Gauge Theories

Springer

Graduate Texts in Physics

Graduate Texts in Physics

Graduate Texts in Physics publishes core learning/teaching material for graduate- and advanced-level undergraduate courses on topics of current and emerging fields within physics, both pure and applied. These textbooks serve students at the MS- or PhD-level and their instructors as comprehensive sources of principles, definitions, derivations, experiments and applications (as relevant) for their mastery and teaching, respectively. International in scope and relevance, the textbooks correspond to course syllabi sufficiently to serve as required reading. Their didactic style, comprehensiveness and coverage of fundamental material also make them suitable as introductions or references for scientists entering, or requiring timely knowledge of, a research field.

More information about this series at http://www.springer.com/series/8431

Edouard B. Manoukian

Quantum Field Theory I

Foundations and Abelian and Non-Abelian Gauge Theories

Springer

Edouard B. Manoukian
The Institute for Fundamental Study
Naresuan University
Phitsanulok, Thailand

ISSN 1868-4513 ISSN 1868-4521 (electronic)
Graduate Texts in Physics
ISBN 978-3-319-30938-5 ISBN 978-3-319-30939-2 (eBook)
DOI 10.1007/978-3-319-30939-2

Library of Congress Control Number: 2016935720

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface to Volume I

This textbook is based on lectures given in quantum field theory (QFT) over the years to graduate students in theoretical and experimental physics. The writing of the book spread over three continents: North America (Canada), Europe (Ireland), and Asia (Thailand). QFT was born about 90 years ago, when quantum mechanics met relativity, and is still going strong. The book covers, pedagogically, the wide spectrum of developments in QFT emphasizing, however, those parts which are reasonably well understood and for which satisfactory theoretical descriptions have been given.

The legendary Richard Feynman *in* his 1958 Cornell, 1959–1960 Cal Tech lectures on QFT of fundamental processes, the first statement he makes, the very first one, is that the *lectures cover all of physics*.[1] One quickly understands what Feynman meant by covering all of physics. The role of fundamental physics is to describe the basic interactions of Nature and *QFT, par excellence, is supposed to do just that*. Feynman's statement is obviously more relevant today than it was then, since the recent common goal is to provide a unified description of *all* the fundamental interactions in nature.

The book requires as background a good knowledge of quantum mechanics, including rudiments of the Dirac equation, as well as elements of the Klein-Gordon equation, and the reader would benefit much by reading relevant sections of my earlier book : *Quantum Theory: A Wide Spectrum (2006), Springer* in this respect.

This book differs from QFT books that have appeared in recent years[2] in several respects and, in particular, it offers something new in its approach to the subject, and the reader has plenty of opportunity to be exposed to many topics not covered, or

[1] R. P. Feynman, The Theory of Fundamental Processes, The Benjamin/Cummings Publishing Co., Menlo Park, California. 6th Printing (1982), page 1.

[2] Some of the fine books that I am familiar with are: L. H. Ryder, Quantum Field Theory; S. Weinberg, The Quantum Theory of Fields I (1995) & II (1996), Cambridge: Cambridge University Press; M. Peskin and D. V. Schroeder, An Introduction to Quantum Field Theory, New York: Westview Press (1995); B. DeWitt, The Global Approach to Quantum Field Theory, Oxford: Oxford University Press (2014).

just touched upon, in standard references. Some notable differences are seen, partly, from unique features in the following material included in ours:

- The very elegant functional *differential* approach of Schwinger, referred to as the quantum dynamical (action) principle, and its underlying theory are used systematically in generating the so-called vacuum-to-vacuum transition amplitude of both abelian and non-abelian gauge theories, in addition to the well-known functional integral approach of Feynman, referred to as the path-integral approach, which are simply related by functional Fourier transforms and delta functionals.
- Transition amplitudes are readily extracted by a direct expansion of the vacuum-to-vacuum transition amplitude in terms of a unitarity sum, which is most closely related to actual experimental setups with particles emitted and detected prior and after a given process and thus represent the underlying physics in the clearest possible way.
- Particular emphasis is put on the concept of a quantum field and its particle content, both physically and technically, as providing an appropriate description of physical processes at sufficiently high energies, for which relativity becomes the indispensable language to do physics and explains the exchange that takes place between energy and matter, allowing the creation of an unlimited number of particles such that the number of particles need not be conserved, and for which a variable number of particles may be created or destroyed. Moreover, quantum mechanics implies that a wavefunction renormalization arises in QFT field independent of any perturbation theory – a point not sufficiently emphasized in the literature.
- The rationale of the stationary action principle and emergence of field equations, via field variations of transformation functions and generators of field variations. The introduction of such generators lead, self consistently, to the field equations. Such questions are addressed as: "Why is the variation of the action, *within* the boundaries of transformation functions, set equal to *zero* which eventually leads to the Euler-Lagrange equations?", "How does the Lagrangian density appear in the formalism?" "What is the significance in commuting/anti-commuting field components within the interaction Lagrangian density in a theory involving field operators?" These are some of the questions many students seem to worry about.
- A panorama of all the fields encountered in present high-energy physics, together with the details of the underlying derivations are given.
- Schwinger's point splitting method of currents is developed systematically in studying abelian and non-abelian gauge theories anomalies. Moreover, an explicit experimental test of the presence of an anomaly is shown by an example.
- Derivation of the Spin & Statistics connection and CPT symmetry, emphasizing for the latter that the invariance of the action under CPT transformation is not sufficient for CPT symmetry, but one has also to consider the roles of incoming and outgoing particles.

- The fine-structure effective coupling $\alpha \simeq 1/128$ at high energy corresponding to the mass of the neutral Z^0 vector boson based on all the charged leptons and all those contributing quarks of the three generations.
- Emphasis is put on renormalization theory, including its underlying general subtractions scheme, often neglected in treatments of QFT.
- Elementary derivation of Faddeev-Popov factors directly from the functional differential formalism, with constraints, and their *modifications*, and how they may even arise in some abelian gauge theories.
- A fairly detailed presentation is given of "deep inelastic" experiments as a fundamental application of quantum chromodynamics.
- Schwinger line integrals, origin of Wilson loops, lattices, and quark confinement.
- Neutrino oscillations,[3] neutrino masses, neutrino mass differences, and the "seesaw mechanism."
- QCD jets and parton splitting, including gluon splitting to gluons.
- Equal importance is put on both abelian and non-abelian gauge theories, witnessing the wealth of information also stored in the abelian case.[4]
- A most important, fairly detailed, and semi-technical introductory chapter is given which traces the development of QFT since its birth in 1926 without tears, in abelian and non-abelian gauge theories, including aspects of quantum gravity, as well as examining the impact of supersymmetry, string theory, and the development of the theory of renormalization, as a *pedagogical* strategy for the reader to be able to master the basic ideas of the subject at the outset before they are encountered in glorious technical details later.
- *Solutions* to all the problems are given right at the end of the book.

With the mathematical rigor that renormalization has met over the years and the reasonable agreement between gauge theories and experiments, the underlying theories are in pretty good shape. This volume is organized as follows. The first introductory chapter traces the subject of QFT since its birth, elaborating on many of its important developments which are conveniently described in a fairly simple language and will be quite useful for understanding the underlying technical details of the theory covered in later chapters including those in Volume II. A preliminary chapter follows which includes the study of symmetry transformations in the quantum world, as well as of intricacies of functional differentiation and functional integration which are of great importance in field theory. Chapter 3 deals with quantum field theory methods of spin 1/2 culminating in the study of anomalies in the quantum world. The latter refers to the fact that a conservation law

[3]It is rather interesting to point out that the theory of neutrino oscillations was written up in this book much earlier than the 2015 Nobel Prize in Physics was announced on neutrino oscillations.

[4]With the development of non-abelian gauge theories, unfortunately, it seems that some students are not even exposed to such derivations as of the "Lamb shift" and of the "anomalous magnetic moment of the electron" in QED.

in classical physics does not necessarily hold in the quantum world. Chapter 4, a critical one, deals with the concept of a quantum field, the Poincaré algebra, and particle states. Particular attention is given to the stationary action principle as well as in developing the solutions of QFT via the quantum dynamical principle. This chapter includes the two celebrated theorems dealing with CPT symmetry and of the Spin & Statistics connection. A detailed section is involved with the basic quantum fields one encounters in present day high-energy/elementary-particle physics and should provide a useful reference source for the reader. Chapter 5 treats abelian gauge theories (QED, scalar boson electrodynamics) in quite details and includes, in particular, the derivations of two of the celebrated results of QED which are the anomalous magnetic moment of the electron and the Lamb shift. Chapter 6 is involved with non-abelian gauge theories (electroweak, QCD, Grand unification).[5] Such important topics are included as "asymptotic freedom," "deep inelastic" scattering, QCD jets, parton splittings, neutrino oscillations, the "seesaw mechanism" and neutrino masses, Schwinger-line integrals, Wilson loops, lattices, and quark confinement. Unification of coupling parameters of the electroweak theory and of QCD are also studied, as well as of spontaneous symmetry breaking in both abelian and non-abelian gauge theories, and of renormalizability aspects of both gauge theories, emphasizing the so-called BRS transformations for the latter. We make it a point, pedagogically, to derive things in detail, and some of such details are *relegated* to appendices at the end of the respective chapters with the main results *given* in the sections in question. Five general appendices, at the end of this volume, cover some additional important topics and/or technical details. In particular, I have included an appendix covering some aspects of the general theory of renormalization and its underlying subtractions scheme itself which is often neglected in books on QFT. Fortunately, my earlier book, with *proofs* not just words, devoted completely to renormalization theory – *Renormalization (1983), Academic Press* – may be consulted for more details. The problems given at the end of the chapters form an integral part of the book, and many developments in the text depend on the problems and may include, in turn, additional material. They should be attempted by every serious student. *Solutions* to all the problems are given right at the end of the book for the convenience of the reader. The introductory chapter together with the introductions to each chapter provide the motivation and the *pedagogical* means to handle the technicalities that follow them in the texts.

I hope this book will be useful for a wide range of readers. In particular, I hope that physics graduate students, not only in quantum field theory and high-energy physics, but also in other areas of specializations will also benefit from it as, according to my experience, they seem to have been left out of this fundamental area of physics, as well as instructors and researchers in theoretical physics. The content of this volume may be covered in one-year (two semesters) quantum field theory courses.

[5] QED and QCD stand, respectively, for quantum electrodynamics and quantum chromodynamics.

In Volume II, the reader is introduced to quantum gravity, supersymmetry, and string theory,[6] which although may, to some extent, be independently read by a reader with a good background in field theory, the present volume sets up the language, the notation, provides additional background for introducing these topics, and will certainly make it much easier for the reader to follow. In this two-volume set, aiming for completeness in covering the basics of the subject, I have included topics from the so-called conventional field theory (the classics) to ones from the modern or the new physics which I believe that every serious graduate student studying quantum field theory should be exposed to.

Without further ado, and with all due respect to the legendary song writer Cole Porter, let us find out "what is this thing called QFT?"

Edouard B. Manoukian

---

[6]Entitled: Quantum Field Theory II: *Introductions to Quantum Gravity, Supersymmetry, and String Theory"* (2016), Springer.

Acknowledgements

In the beginning of it all, I was introduced to the theoretical aspects of quantum field theory by Theodore Morris and Harry C. S. Lam, both from McGill and to its mathematical intricacies by Eduard Prugovečki from the University of Toronto. I am eternally grateful to them. Over the years, I was fortunate enough to attend a few lectures by Julian Schwinger and benefited much from his writings as well. Attending a lecture by Schwinger was quite an event. His unique elegant, incisive, physically clear approach and, to top it off, short derivations were impressive. When I was a graduate student, I would constantly hear that Schwinger "does no mistakes." It took me years and years to understand what that meant. My understanding of this is because he had developed such a powerful formalism to do field theory that, unlike some other approaches, everything in the theory came out automatically and readily without the need to worry about multiplicative factors in computations, such as 2π's and other numerical factors, and, on top of this, is relatively easy to apply. Needless to say this has much influenced my own approach to the subject. He had one of the greatest minds in theoretical physics of our time.

I want to take this opportunity as well to thank Steven Weinberg, the late Abdus Salam, Raymond Streater, and Eberhard Zeidler for the keen interest they have shown in my work on renormalization theory.

I acknowledge with thanks the support I received from several colleagues, while visiting their research establishments for extended periods, for doing my own thing and writing up the initial notes on this project. These include Yasushi Takahashi and Anton Z. Capri from the University of Alberta, Lochlainn O'Raifeartaigh and John Lewis from the Dublin Institute for Advanced Studies, and Jiri Patera and Pavel Winternitz from the University of Montreal. For the final developments of the project, I would like to thank Sujin Jinahyon, the President of Naresuan University, Burin Gumjudpai, Seckson Sukhasena, Suchittra Sa-Nguansin, and Jiraphorn Chomdaeng of the university's Institute for Fundamental Study for encouragement, as well as Ahpisit Ungkitchanukit and Chai-Hok Eab from Chulalongkorn University.

I am also indebted to many of my former graduate students, who are now established physicists in their own rights, particularly to Chaiyapoj Muthaporn, Nattapong Yongram, Siri Sirininlakul, Tukkamon Vijaktanawudhi (aka Kanchana Limboonsong), Prasopchai Viriyasrisuwattana, and Seckson Sukhasena, who through their many questions, several discussions, and collaborations have been very helpful in my way of analyzing this subject.

Although I have typed the entire manuscripts myself, and drew the figures as well, Chaiyapoj Muthaporn prepared the LaTeX input files. Without his constant help in LaTeX, this work would never have been completed. I applaud him, thank him and will always remember how helpful he was. My special thanks also go to Nattapong Yongram for downloading the endless number of papers I needed to complete the project.

I was fortunate and proud to have been associated with the wonderful editorial team of Maria Bellantone and Mieke van der Fluit of Springer. I would like to express my deepest gratitude to them for their excellent guidance, caring, patience, and hard work in making this project possible and move forward toward its completion. I have exchanged more emails with Mieke than with anybody else on the globe. This has led to such an enjoyable association that I will always cherish.

This project would not have been possible without the patience, encouragement, and understanding of my wife Tuenjai. To my parents, who are both gone, this work is affectionately dedicated.

Contents

Notation and Data

- Latin indices $i, j, k, \ldots$ are generally taken to run over 1,2,3, while the Greek indices $\mu, \nu, \ldots$ over $0, 1, 2, 3$ in 4D. Variations do occur when there are many different types of indices to be used, and the meanings should be evident from the presentations.
- The Minkowski metric $\eta_{\mu\nu}$ is defined by $[\eta_{\mu\nu}] = \mathrm{diag}[-1, 1, 1, 1] = [\eta^{\mu\nu}]$ in 4D.
- Unless otherwise stated, the fundamental constants $\hbar$, c are set equal to one.
- The gamma matrices satisfy the anti-commutation relations $\{\gamma^{\mu}, \gamma^{\nu}\} = -2\,\eta^{\mu\nu}$.
- The Dirac, the Majorana, and the chiral representations of the γ^{μ} matrices are defined in Appendix I at the end of the book.
- The charge conjugation matrix is defined by $\mathscr{C} = \mathrm{i}\gamma^{2}\gamma^{0}$.
- $\overline{\psi} = \psi^{\dagger}\gamma^{0}$, $\overline{u} = u^{\dagger}\gamma^{0}$, $\overline{v} = v^{\dagger}\gamma^{0}$. A Hermitian conjugate of a matrix M is denoted by $M^{\dagger}$, while its complex conjugate is denoted by M^{*}.
- The step function is denoted by $\theta(x)$ which is equal to 1 for $x > 0$, and 0 for $x < 0$.
- The symbol ε is used in dimensional regularization (see Appendix III). ϵ is used in defining the boundary condition in the denominator of a propagator $(Q^2 + m^2 - \mathrm{i}\,\epsilon)$ and should not be confused with ε used in dimensional regularization. We may also use either one when dealing with an infinitesimal quantity, in general, with ϵ more frequently, and this should be self-evident from the underlying context.
- For units and experimental data, see the compilation of the "Particle Data Group": Beringer et al. [1] and Olive et al. [2]. The following (some obviously approximate) numerical values should, however, be noted:

$$\begin{aligned} 1\ \mathrm{MeV} &= 10^{6}\ \mathrm{eV} \\ 1\ \mathrm{GeV} &= 10^{3}\ \mathrm{MeV} \\ 10^{3}\ \mathrm{GeV} &= 1\ \mathrm{TeV} \\ 1\ \mathrm{erg} &= 10^{-7}\ \mathrm{J} \end{aligned}$$

$$\begin{aligned} 1\ \mathrm{J} &= 6.242 \times 10^{9}\ \mathrm{GeV} \\ \mathrm{c} &= 2.99792458 \times 10^{10}\ \mathrm{cm/s\ (exact)} \\ \hbar &= 1.055 \times 10^{-34}\ \mathrm{J\ s} \\ \hbar\,\mathrm{c} &= 197.33\ \mathrm{MeV\ fm} \\ 1\ \mathrm{fm} &= 10^{-13}\ \mathrm{cm} \end{aligned}$$

(Masses) $M_p = 938.3\,\mathrm{MeV/c^2}$, $M_n = 939.6\,\mathrm{MeV/c^2}$,
$M_W = 80.4\ \mathrm{GeV/c^2}$, $M_Z = 91.2\,\mathrm{GeV/c^2}$,
$m_e = 0.511\,\mathrm{MeV/c^2}$, $m_\mu = 105.66\ \mathrm{MeV/c^2}$, $m_\tau = 1777\,\mathrm{MeV/c^2}$.
Mass of $\nu_e < 2\,\mathrm{eV/c^2}$, Mass of $\nu_\mu < 0.19\,\mathrm{MeV/c^2}$, Mass of $\nu_\tau < 18.2\,\mathrm{MeV/c^2}$,
Mass of the neutral Higgs $H^0 \approx 125.5\,\mathrm{GeV/c^2}$.
For approximate mass values of some of the quarks taken, see Table 5.1 in Sect. 5.19.2. For more precise range of values, see Olive et al. [2].
(Newton's gravitational constant) $\mathrm{G_N} = 6.709 \times 10^{-39}\,\hbar\,\mathrm{c^5/GeV^2}$.
(Fermi weak interaction constant) $\mathrm{G_F} = 1.666 \times 10^{-5}\,\hbar^3\,\mathrm{c^3/GeV^2}$.
Planck mass $\sqrt{\hbar\,\mathrm{c/G_N}} \approx 1.221 \times 10^{19}\,\mathrm{GeV/c^2}$,
Planck length $\sqrt{\hbar\,\mathrm{G_N/c^3}} \approx 1.616 \times 10^{-33}$ cm.
Fine structure constant $\alpha = 1/137.04$ at $Q^2 = 0$, and $\approx 1/128$ at $Q^2 \approx M_Z^2$.
For the weak-mixing angle θ_W, $\sin^2\theta_W \approx 0.232$, at $Q^2 \approx M_Z^2$.
$\alpha/\sin^2\theta_W \approx 0.034$, at $Q^2 \approx M_Z^2$.
Strong coupling constant $\alpha_s \approx 0.119$, at $Q^2 \approx M_Z^2$.

References

1. Beringer, J., et al. (2012). Particle data group. *Physical Review D, 86*, 010001.
2. Olive, K. A., et al. (2014). Particle data group. *Chinese Physics C, 38*, 090001.

Chapter 1
Introduction

Donkey Electron, Bare Electron, Electroweak Frog, God Particle, "Colored" Quarks and Gluons, Asymptotic Freedom, Beyond Resonances into the Deep Inelastic Region, Partons, QCD Jets, Confined Quarks, Bekenstein – Hawking Entropy of a Black Hole, Sparticles, Strings, Branes, Various Dimensions and even Quanta of Geometry, AdS/CFT Correspondence and Holographic Principle, CPT, and Spin & Statistics

The major theme of quantum field theory is the development of a unified theory that may be used to describe nature from microscopic to cosmological distances. Quantum field theory was born 90 years ago, when quantum theory met relativity, and has captured the hearts of the brightest theoretical physicists in the world. It is still going strong. It has gone through various stages, met various obstacles on the way, and has been struggling to provide us with a coherent description of nature in spite of the "patchwork" of seemingly different approaches that have appeared during the last 40 years or so, but still all, with the common goal of unification.

As mentioned in our Preface, Feynman, in his 1958 Cornell, 1959–1960 Cal Tech, lectures on the quantum field theory of fundamental processes, the first statement he makes, the very first one, is that the lectures will cover *all* of physics [76, p. 1]. One quickly understands what Feynman meant by covering all of physics. After all, the role of fundamental physics is to describe the basic interactions we have in nature and quantum field theory is supposed to do just that. Feynman's statement is obviously more relevant today than it was then, since the recent common goal is to provide a unified description of *all* the fundamental forces in nature. With this in mind, let us trace the development of this very rich subject from the past to the present, and see what the theory has been telling us all these years.

When the energy and momentum of a quantum particle are large enough, one is confronted with the requirement of developing a formalism, as imposed by nature, which extends quantum theory to the relativistic regime. A relativistic theory, as a result of the exchange that takes place between energy and matter, allows the creation of an unlimited number of particles and the number of particles in a given physical process need not be conserved. An appropriate description of such physical processes for which a variable number of particles may be created or destroyed, in the quantum world, is provided by the very rich concept of a quantum field. For example, photon emissions and absorptions, in a given process, are explained by the introduction of the electromagnetic quantum field. The theory which emerges from extending quantum physics to the relativistic regime is called "Relativistic Quantum

E.B. Manoukian, *Quantum Field Theory I*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-30939-2_1

Field Theory" or just *"Quantum Field Theory"*. Quantum Electrodynamics is an example of a quantum field theory and is the most precise theory devised by man when confronted with experiments. The essence of special relativity is that all inertial frames are completely equivalent in explaining a physical theory as one inertial frame cannot be distinguished from another. This invariance property of physical theories in all inertial frames, as required by special relativity, as well as by the many symmetries one may impose on such theories, are readily implemented in the theory of quantum fields. The implementation of symmetries and describing their roles in the explanation of observed phenomena has played a key role in elementary particle physics.

Of course it took years before the appropriate language of quantum field theory, described concisely above, by marrying quantum theory and relativity, was spelled out and applied consistently to physical processes in the quantum world in the relativistic regime. An appropriate place to start in history is when Dirac [47–49] developed his relativistic equation of spin $1/2$, from which one learns quite a bit about the subsequent development of the subject as a multi-particle theory. We will then step back a year or two, and then move again forward in time to connect the dots between the various stages of the underlying exciting developments. His relativistic equation, which incorporated the spin of the electron, predicted the existence of negative energy states with negative mass, with energies going down to $-\infty$, implying the instability of the corresponding systems. For example, an electron in the ground-state energy of an atom would spontaneously decay to such lower and lower negative energy states emitting radiation of arbitrary large energies leading eventually to the collapse of the atom with the release of an infinite amount of energy. Historically, a relativistic equation for spin 0, was developed earlier by Klein and Gordon in 1926,[1] referred to as the Klein-Gordon equation, which also shared this problem, but unlike Dirac's theory it led to negative probabilities as well. Dirac being aware of the negative probabilities encountered in the theory of the latter authors, was able to remedy this problem in his equation. To resolve the dilemma of negative energies, Dirac, in 1930,[2] assumed that a priori all the negative energy states are filled with electrons in accord to the Pauli exclusion principle, giving rise to the so-called Dirac sea or the Dirac vacuum, so that no transitions to such states are possible, thus ensuring the stability of the atom.

The consequences of the assumption made by Dirac above were many. A negative energy electron in the Dirac sea, may absorb radiation of sufficient energy so as to overcome an energy gap arising from the level $-mc^2$ to $+mc^2$, where m is the mass of an electron, thus making such a negative energy electron jump to a positive energy state, leaving behind a surplus of positive energy and a surplus of

[1] Klein [128] and Gordon [101].

[2] Dirac [50, 51].

positive charge $+|e|$ *relative* to the Dirac sea. This has led Dirac eventually,[3] in 1931 [52], to interpret the "hole" left behind by the transition of the negative energy electron to a positive energy state, as a particle that has the same mass as the electron but of opposite charge. It is interesting to note that George Gamow referred[4] to Dirac's predicted particle as a *"donkey electron"*, because it would move in the opposite direction of an appropriate applied force. The physics community found it difficult to accept Dirac's prediction until Anderson[5] discovered this particle (the positron e^+), who apparently was not aware of Dirac's prediction at the time of the discovery.[6] With the positron now identified, the above argument just given has provided an explanation of the so-called *pair production* $\gamma \to e^+e^-$ by a photon (in the vicinity of a nucleus).[7] Conversely, if a "hole" is created in the vacuum, then an electron may make a transition to such a state releasing radiation giving rise to the phenomenon of *pair annihilation*. A Pair created, as described above, in the vicinity of a positively charged nucleus, would lead to a partial screening of the charge of the nucleus as the electron within the pair would be attracted by the nucleus and the positively charged one would be repelled. Accordingly, an electron, in the atom, at sufficiently large distances from the nucleus would then see a smaller charge on the nucleus than an electron nearby (such as one in an *s*-state). This leads to the concept of *vacuum polarization*, and also to the concept of *charge renormalization* as a result of the partial charge screening mentioned above.

The Dirac equation is Lorentz covariant, that is, it has the same form in every inertial frame with its variables being simply relabeled reflecting the variables used in the new inertial frame. It predicted, approximately, the gyromagnetic ratio $g = 2$ of the electron, the fine-structure of the atom, and eventually anti-matter was discovered such as antiprotons.[8] It was thus tremendously successful. Apparently,[9] Dirac himself remarked in one of his talks that *his equation was more intelligent than its author*.[10]

Thus the synthesis of relativity and quantum physics, led to the discovery of the antiparticle. The Dirac equation which was initially considered to describe a single particle necessarily led to a *multi*-particle theory, and a single particle description in the relativistic regime turned out to be not complete. A formalism which would naturally describe creation and annihilation of particles and take into account this

---

[3]Dirac [50, 51] assumed that the particle is the proton as the positron was not discovered yet at that time. Apart from the large mass difference between the proton and the electron, there were other inconsistencies with such an assumption.

[4]Weisskopf [242].

[5]Anderson [5, 6].

[6]Weisskopf [242].

[7]The presence of the nucleus is to conserve energy and momentum.

[8]Chamberlain et al. [30].

[9]Weisskopf [242].

[10]For a systematic treatment of the intricacies of Dirac's theory and of the quantum description of relativistic particles, in general, see Manoukian [151], Chapter 16.

multi-particle aspect became necessary. The so-called "hole" theory although it gave insight into the nature of fundamental processes involving quantum particles in the relativistic regime, and concepts such as vacuum polarization, turned out to be also not complete. For example, in the "hole" theory, the number of electrons minus the number of positrons, created is conserved by the simultaneous creation of a "hole" for every electron ejected from the Dirac Sea. In nature, there are processes, where just an electron or just a positron is created while conserving charge of course. Examples of such processes are β^- decay: $n \to p + e^- + \tilde{\nu}_e$, muon decay: $\mu^- \to e^- + \tilde{\nu}_e + \nu_\mu$, and β^+ decay: $p \to n + e^+ + \nu_e$, for a bound proton in a nucleus for the latter process. Finally, Dirac's argument of a sea of negatively charged bosons did not work with the Klein-Gordon equation because of the very nature of the Bose statistics of the particles. A new description to meet all of the above challenges including the creation and annihilation of particles, mentioned above, was necessary.

After the conceptual framework of quantum mechanics was developed, Born, Heisenberg, and Jordan in 1926 [26], applied quantum mechanical methods to the electromagnetic field, now, giving rise to a system with an infinite degrees of freedom, and described as a set of independent harmonic oscillators of various frequencies. Then Dirac in 1927 [46], prior to the development of his relativistic spin 1/2 equation, also extended quantum mechanical methods to the electromagnetic field now with the latter field treated as an operator, and provided a theoretical description of how photons emerge in the quantization of the electromagnetic field. This paper is considered to mark the birthdate of "Quantum Electrodynamics", a name coined by Dirac himself, and provided a prototype for the introduction of field operators for other particles with spin, such as for spin 1/2, where in the latter case commutators in the theory are replaced by anti-commutators [125, 126] for the fermion field.

The first comprehensive treatment of a general quantum field theory, involving Lagrangians, as in modern treatments, was given by Heisenberg and Pauli in 1929, 1930 [116, 117], where canonical quantization procedures were applied directly to the fields themselves. A classic review of the state of affairs of quantum electrodynamics in 1932 [68] was given by Fermi. The problem of negative energy solutions was resolved and its equivalence to the Dirac "hole" theory was demonstrated by Fock in 1933 [83], and Furry and Oppenheimer in 1934 [90], where the (Dirac) field operator and its adjoint were expanded in terms of appropriate creation and annihilation operators for the electron and positron, thus providing a unified description for the particle and its antiparticle. The method had a direct generalization to bosons. The old "hole" theory became unnecessary and obsolete.[11] The problem of negative energy solutions was also resolved for spin 0 bosons by

[11] As a young post-doctoral fellow, I remember attending Schwinger's lecture tracing the Development of Quantum Electrodynamics in "The Physicist's Conception of Nature" [202], making the statement, regarding the "hole" theory, that it is now best regarded as an historical curiosity, and forgotten.

similar methods by Pauli and Weisskopf in 1934 [170]. The fields thus introduced from these endeavors have become operators for creation and annihilation of particles and antiparticles, rather than probability amplitudes.[12]

The explanation that interactions are generated by the exchange of quanta was clear in the classic work of Bethe and Fermi in 1932 [18]. For example, charged particles, as sources of the electromagnetic field, influence other charged particles via these electromagnetic fields. Fields as operators of creation and destruction of particles, and the association of particles with forces is a natural consequence of field theory. The same idea was used by Yukawa in 1935 [249], to infer that a massive scalar particle is exchanged in describing the strong interaction (as understood in those days), with the particle necessarily being massive to account for the short range nature of the strong force unlike the electromagnetic one which is involved with the massless photon describing an interaction of infinite range. The mass μ of the particle may be estimated from the expression $\mu \approx \hbar/R\mathrm{c}$, obtained formally from the uncertainty principle, where R denotes the size of the proton, i.e., $R = 1\,\mathrm{fm} = 10^{-13}\,\mathrm{cm}$. In natural units, i.e., for $\hbar = 1, \mathrm{c} = 1$, $1\,\mathrm{fm} \approx 1/(200\,\mathrm{MeV})$. This gives $\mu \approx 200\,\mathrm{MeV}$. Such a particle (the pion) was subsequently discovered by the C. F. Powell group in 1947 [136].

As early as 1930s, infinities appeared in explicit computations in quantum electrodynamics by Oppenheimer [168], working within an atom, by Waller [233, 234], and by Weisskopf [239]. The nature of these divergences, arising in these computations, came from integrations that one had to carry out over energies of photons exchanged in describing the interaction of the combined system of electrons and the electromagnetic field to arbitrary high-energies. By formally restricting the energies of photons exchanged, as just described, to be less than, say, κ, Weisskopf, in his calculations, has shown [239, 240], within the full quantum electrodynamics, that the divergences encountered in the self-energy acquired by the electron from its interaction with the electromagnetic field is of the logarithmic[13] type $\sim \ln(\kappa/m\mathrm{c}^2)$, improving the preliminary calculations done earlier, particularly, by Waller, mentioned above. That such divergences, referred to as "ultraviolet divergence",[14] are encountered in quantum field theory should

[12]It is important to note, however, that the matrix elements of these field operators between particle states and the vacuum naturally lead to amplitudes of particles creation by the fields and to the concept of *wavefunction* renormalization (see Sect. 4.1) independently of any perturbation theories.

[13]The corresponding expression occurs with higher powers of the logarithm for higher orders in the fine-structure constant $\mathrm{e}^2/4\pi\hbar c$.

[14]That is, divergences arising from the high-energy behavior of a theory. Another type of divergence, of different nature occurring in the low energy region, referred to as the "infrared catastrophe", was encountered in the evaluation of the probability that a photon be emitted in a collision of a charged particle. In computations of the scattering of charged particles, due to the zero mass nature of photons, their simultaneous emissions in arbitrary, actually infinite, in number must necessary be taken into account for a complete treatment. By doing so finite expressions for the probabilities in question were obtained [22].

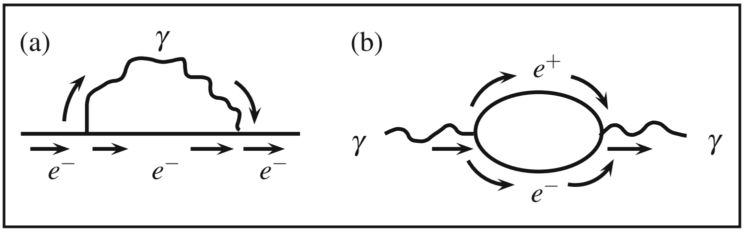

Fig. 1.1 Processes leading to an electron self-energy correction, and vacuum polarization, respectively

be of no surprise as one is assuming that our theories are valid up to infinite energies![15]

The $2S_{1/2}$, $2P_{1/2}$ states of the Hydrogen atom are degenerate in Dirac's theory. In 1947 [134], Lamb and Retherford, however, were able to measure the energy difference between these states, referred to as the "Lamb Shift", using then newly developed microwave methods with great accuracy. Bethe [17] then made a successful attempt to compute this energy difference by setting an upper limit for the energy of photon exchanged in describing the electromagnetic interaction of the order of the rest energy of the electron $m\mathrm{c}^2$, above which relativistic effects take place, relying on the assumption that the electron in the atom is non-relativistic, and, in the process, took into consideration of the mass shift[16] of the electron. He obtained a shift of the order of 1000 megacycles which was in pretty good agreement with the Lamb-Retherford experiment.

Very accurate computations were then made, within the full relativistic quantum electrodynamics, and positron theory. Notably, Schwinger[17] in 1948 [192], computed the magnetic moment of the electron modifying the gyromagnetic ratio, $g = 2$ in the Dirac theory, to $2\,(1 + \alpha/2\pi)$, to lowest order in the fine-structure constant. The computation of the Lamb-Shift was also carried out in a precise manner by Kroll and Lamb [133], and, for example, by French and Weisskopf [85], and Fukuda et al. [88, 89].

State of affairs changed quite a bit. It became clear that an electron is accompanied by an electromagnetic field which in turn tends to alter the nature of the electron that one was initially aiming to describe. The electron e^-, being a charged particle, produces an electromagnetic field (γ). This field, in turn, interacts back with the electron as shown below in Fig. 1.1a. Similarly, the electromagnetic field (γ) may lead to the creation of an electron-positron pair $e^+\,e^-$, which in turn annihilate each other re-producing an electromagnetic field, a process referred to as vacuum-polarization, shown in part (b). Because of these processes, the parameters initially

[15]See also the discussion in Sect. 5.19.

[16]See also the important contribution to this by Kramers [130]. This reference also includes contributions of his earlier work.

[17]See also Appendix B of Schwinger [193].

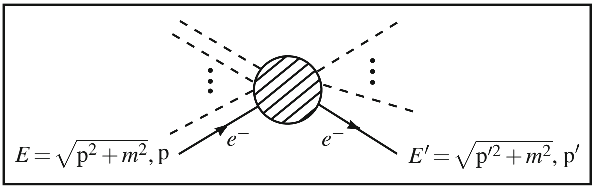

Fig. 1.2 As a result of the self-energy correction in Fig. 1.1a, where an electron emits and re-absorbs a photon, the mass parameter m_0, one initially starts with, does not represent the physical mass of the electron determined in the lab. Here this is emphasized by the energy dependence on the physical mass m of an electron in a scattering process. The *dashed lines* represent additional particles participating in the process

appearing in the theory, such as mass, say, m_0, vis-à-vis Fig. 1.1a, and the electron charge, say, e_0, vis-à-vis Fig. 1.1b, that were associated with the electron one starts with, are not the parameters actually measured in the lab. For example, the energy of a scattered electron of momentum $\mathbf{p}$, in a collision process, turned up to be not equal to $\sqrt{\mathbf{p}^2+m_0^2}$ but rather to $\sqrt{\mathbf{p}^2+m^2}$, self-consistently,[18] with m identified with the actual, i.e., tabulated, mass of the electron, and $m \neq m_0$, with a scattering process shown in Fig. 1.2, where the dashed lines represent other particles (such as γ, e^-, e^+), where the total charge as well as the total energy and momentum are conserved in the scattering process.

Similarly, the potential energy between two widely separated electrons, by a distance r, turned up to be not $\mathrm{e}_0^2/4\pi r$ but rather $\mathrm{e}^2/4\pi r$, with $\mathrm{e}^2 \neq \mathrm{e}_0^2$, where e is identified with the charge, i.e., the tabulated charge, of the electron. As we will see later, the physical parameters are related to the initial ones by scaling factors, referred to as mass and charge renormalization constants, respectively. An electron parametrized by the couple (m_0, e_0), is referred to as a *bare electron* as it corresponds to measurements of its properties by going down to "zero" distances all the way into the "core" of the electron – a process that is unattainable experimentally. On the other hand, the physical parameters (m, e), correspond to measurements made on the electron from sufficiently large distances.

One thus, in turn, may generate parameters, corresponding to a wide spectrum of scales running from the very small to the very large. Here one already notices that in quantum field theory, one encounters so-called effective parameters which are functions of different scales (or energies). Functions of these effective parameters turn out to satisfy invariance properties under scale transformations, thus introducing a concept referred to as the renormalization group. Clearly, due to the screening effect via vacuum polarization of e^+e^- pairs creation, as discussed earlier, the magnitude of the physical charge is smaller than the magnitude of the bare charge.

[18]An arbitrary number of photons of vanishingly small energies are understood to be attached to the external electron lines, as discussed in Footnote 14 when dealing with infrared divergence problems.

A process was, in turn, then carried out, referred to as "renormalization", to eliminate the initial parameters in the theory in favor of *physically* observed ones. This procedure related the theory at very small distances to the theory at sufficiently large distances at which particles emerge on their way to detectors as it happens in actual experiments. All the difficulties associated with the ultraviolet divergences in quantum electrodynamics were isolated in renormalization constants, such as the ones discussed above, and one was then able to eliminate them in carrying out physical applications giving rise to completely finite results. This basic idea of the renormalization procedure was clearly spelled out in the work of Schwinger, Feynman, and Tomonaga.[19] The renormalization group,[20] mentioned above, describes the connection of renormalization to scale transformations, and relates, in general, the underlying physics at different energy scales.

In classic papers, Dyson [59, 60] has shown not only the equivalence of the Schwinger, Feynman, and Tomonaga approaches,[21] and the finiteness of the so-called renormalized quantum electrodynamics, but also developed a formalism for computations that may be readily applied to other interacting quantum field theories. Theories that are consistently finite when all the different parameters appearing initially in the theory are eliminated in favor of the physically observed ones, which are finite in number, are said to be renormalizable. Dyson's work, had set up:

renormalizability as a condition for generating field theory interactions.

In units of $\hbar = 1$, $\mathrm{c} = 1$, [Mass] $=$ [Length]$^{-1}$. Roughly speaking, in a renormalizable theory, no coupling constants can have the dimensions of negative powers of mass. (Because of dimensional reasons, we note, in particular, that one cannot have too many derivatives of the fields, describing interactions, as every derivative necessitates involving a coupling of dimensionality reduced by one in units of mass.)

The photon as the agent for transmitting the interaction between charged particles, is described by a vector – the vector potential. In quantum electrodynamics, as a theory of the interaction of photons and electrons, for example, the photon is coupled locally to the electromagnetic current. The latter is also a vector, and the interaction is described by their scalar product (in Minkowski space) ensuring the relativistic invariance of the underlying theory. To lowest order in the charge e of the electron e, this coupling may be represented by the diagram Fig. 1.3a. On the other hand, for a spin 0 charged boson φ, say, one encounters two such diagrams, each shown to lowest order in the charge e in Fig. 1.3b, where we note that in the second diagram in the latter part, two photons emerge locally from the same point.

[19] This is well described in their Nobel lectures: Schwinger [201], Feynman [75], Tomonaga [225], as well as in the collection of papers in Schwinger [198, 201].

[20] Stueckelberg and Peterman [209], Gell-Mann and Low [93], Bogoliubov and Shirkov [25], Ovsyannikov [169], Callan [28, 29], Symanzik [213–215], Weinberg [237], and 't Hooft [218].

[21] The best sources for these approaches are their Nobel Lectures: Schwinger [201], Feynman [75], Tomonaga [225], as well as Schwinger [198].

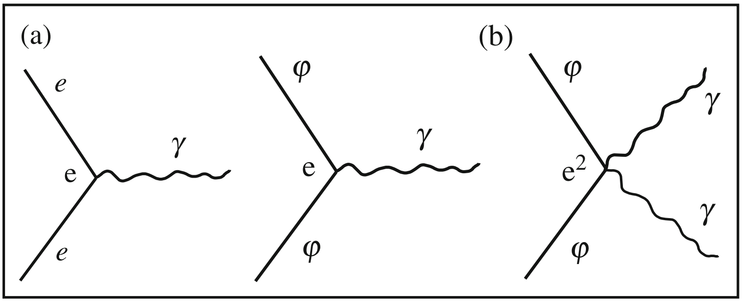

Fig. 1.3 Local couplings for photon emission by an electron, and by a spin 0 charged particle described by the field φ, respectively

Quantum Electrodynamics, was not only the theory of interest. There was also the weak interaction. The preliminary theory of weak interaction dates back to Fermi [69, 70]. Based on weak processes such as β^- decay: $n \rightarrow p + e^- + \tilde{\nu}_e$, he postulated that the weak interactions may be described by local four-point interactions involving a universal coupling parameter $\mathrm{G_F}$. The four particles of the process just mentioned, interact locally at a point with a zero range interaction. The Fermi theory was in good agreement in predicting the energy distribution of the electron. For dimensional reasons, however, the dimensions of the coupling constant $\mathrm{G_F}$ involved in the theory has the dimensions of $[\mathrm{Mass}]^{-2}$, giving rise to a non-renormalizable theory.[22] In analogy to quantum electrodynamics, the situation with this type of interaction may be somehow improved by introducing, in the process, a vector Boson[23] W^- which mediates an interaction[24] between the two pairs (so-called currents), (n, p) and $(e^-, \tilde{\nu}_e)$, with both necessarily described by entities carrying (Lorentz) vector indices, to ensure the invariance of the underlying description. Moreover, in units of $\hbar = 1, \mathrm{c} = 1$, a *dimensionless* coupling g is introduced. The Fermi interaction and its modification are shown, respectively in parts Fig. 1.4a, b.

In order that the process in diagram given in part Fig. 1.4b, be consistent with the "short-range" nature of the Fermi interaction, described by the diagram on the left, the vector particle W^- must not only be massive but its mass, M_W must be quite large. This is because the propagator of a massive vector particle of mass M_W, which mediates an interaction between two spacetime points x and x', as we will discuss below, behaves like $\eta_{\mu\nu}\delta^{(4)}(x - x')/M_W^2$ for a large mass, signifying

[22]It is interesting to point out as one goes to higher and higher orders in the Fermi coupling constant $\mathrm{G_F}$, the divergences increase (Sect. 6.14) without any bound and the theory becomes uncontrollable.

[23] A quantum relativistic treatment of a problem, implies that a theory involving the W^- particle, must also include its antiparticle W^+, having the same mass as of W^-.

[24]Such a suggestion was made, e.g., by Klein [129].

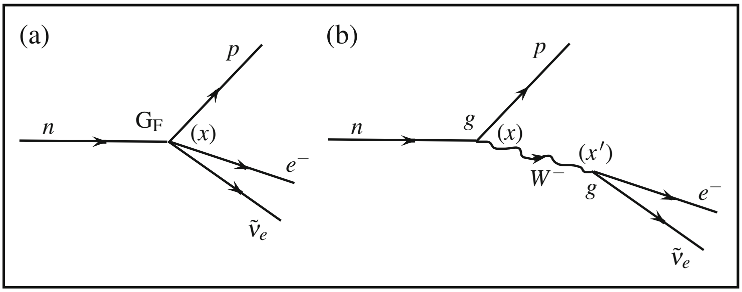

Fig. 1.4 (**a**) The old Fermi theory with a coupling $\mathrm{G_F}$ is replaced by one in (**b**) where the interaction is mediated by a vector boson with a dimensionless coupling g

necessarily a vanishingly small range of the interaction.[25] Upon comparison of both diagrams, one may then infer that

$$\mathrm{G_F} \approx \frac{g^2}{M_W^2}. \tag{1.1}$$

Evidently, the Fourier transform of the propagator in the energy-momentum description, at energies much less than M_W is, due to the $\delta^{(4)}(x-x')$ function given above, simply $\eta_{\mu\nu}/M_W^2$, and (1.1) may be obtained from a low-energy limit.

With some minimal effort, the reader will understand better the above two limits and some of the difficulties encountered with a massive vector boson, in general, if, at this stage, we write down explicitly its propagator between two spacetime points x, x' in describing an interaction carried by the exchange of such a particle which is denoted by[26]:

$$\triangle_+^{\mu\nu}(x-x') = \int \frac{(\mathrm{d}k)}{(2\pi)^4}\, \mathrm{e}^{\mathrm{i}\,k_\nu(x^\nu - x'^\nu)}\, \triangle_+^{\mu\nu}(k), \quad (\mathrm{d}k) = \mathrm{d}k^0\mathrm{d}k^1\mathrm{d}k^1\mathrm{d}k^3, \tag{1.2}$$

$$\triangle_+^{\mu\nu}(k) = \frac{1}{(k^2 + M_W^2 - \mathrm{i}\,0)}\left(\eta^{\mu\nu} + \frac{k^\mu k^\nu}{M_W^2}\right), \tag{1.3}$$

[25] Here $\eta_{\mu\nu}$ is the Minkowski metric.

[26] This expression will be derived in Sect. 4.7. For a so-called virtual particle $k^2 = \mathbf{k}^2 - (k^0)^2 \neq -M_W^2$. The $-\mathrm{i}\,0$ in the denominator in (1.3) just specifies the boundary condition on how the k^0 integration is to be carried out. These things will be discussed in detail later on and are not needed here.

where k^0 is its energy, and $\mathbf{k} = (k^1, k^2, k^3)$ its momentum. Formally for $M_W^2 \to \infty$, $\triangle_+^{\mu\nu}(k) \to \eta^{\mu\nu}/M_W^2$, leading from (1.2) to

$$\triangle_+^{\mu\nu}(x-x') \;\to\; \frac{\eta^{\mu\nu}}{M_W^2}\int \frac{(\mathrm{d}k)}{(2\pi)^4}\,\mathrm{e}^{\mathrm{i}k_\nu(x^\nu - x'^\nu)} = \frac{\eta^{\mu\nu}}{M_W^2}\,\delta^4(x-x'), \tag{1.4}$$

signalling, in a limiting sense, a short range interaction for a heavy-mass particle. On the other hand, for $|k^\nu| \ll M_W$ for each component, one has

$$\triangle_+^{\mu\nu}(k) \approx \frac{\eta^{\mu\nu}}{M_W^2}, \tag{1.5}$$

in the energy-momentum description.

Although the introduction of the intermediate boson W improves somehow the divergence problem, it is still problematic. The reason is not difficult to understand. In the energy-momentum description, the propagator of a massive vector particle, as given in (1.3), has the following behavior at high energies and momenta

$$\triangle_+^{\mu\nu}(k) \;\to\; \frac{1}{k^2}\,\frac{k^\mu k^\nu}{M_W^2}, \tag{1.6}$$

providing *no damping* in such a limit. Moreover, as one goes to higher orders in perturbation theory the number of integration variables, over energy and momenta arising in the theory, increase, and the divergences in turn increase without bound and the theory becomes uncontrollable.[27] On the other hand, an inherited property of quantum electrodynamics is *gauge symmetry* due to the masslessness of the photon. In the present context of ultraviolet divergences, the photon propagator has a very welcome vanishing property at high energies. This gauge symmetry as well as the related *massless* aspect of the photon, which are key ingredients in the self consistency of quantum electrodynamics, turned out to provide a *guiding principle* in developing the so-called electroweak theory.

In 1956 [138], an important observation was made by Lee and Yang that parity P is violated in the weak interactions. Here we recall that, given a process, its parity transformed (mirror) version, is obtained by reversing the directions of the space variables.[28] This has led Gershtein and Zel'dovich [95], Feynman and Gell-Mann [77], Sudarshan and Marshak [210], and Sakurai [178], to express the currents

[27] The damping provided by the propagators of a massless vector particle, a spin 1/2 particle, and a spin 0 particle, for example, in the ultraviolet region vanish like 1/energy^2, 1/energy, 1/energy^2, respectively.

[28] It was later observed that the product of charge conjugation, where a particle is replaced by its antiparticle, and parity transformation "CP", is also not conserved in a decay mode of K mesons at a small level [38, 39, 82]. As the product "CPT", of charge conjugation, parity transformation, and time reversal "T", is believed to be conserved, the violation of time reversal also follows. For a test of such a violation see CPLEAR/Collaboration [36].

constructed out of the pairs of fields: $(n,p),(e^-,\tilde{\nu}_e),\ \ldots$ in the Fermi theory to reflect, in particular, this property dictated by nature. The various currents were eventually expressed and conveniently parametrized in such a way that the theory was described by the universal coupling parameter $\mathrm{G_F}$. The construction of such fundamental currents together with idea of intermediate vector bosons exchanges to describe the weak interaction led eventually to its modern version.

Quantum Electrodynamics may be considered to arise from local gauge invariance in which the electron field is subjected to a local phase transformation $\mathrm{e}^{\mathrm{i}\vartheta(x)}$. The underlying group of transformations is denoted by $\mathrm{U}(1)$ involving simply the identity as the *single* generator of transformations with which the photon is associated as the single gauge field. In 1954 [247], Yang and Mills, and Shaw in 1955 [203], generalized the just mentioned abelian gauge group of phase transformations, encountered in quantum electrodynamics, to a non-abelian[29] gauge theory, described by the group $\mathrm{SU}(2)$,[30] and turned out to be a key ingredient in the development of the modern theory of weak interactions. This necessarily required the introduction, in addition to the charged bosons $W^{\pm}$, a neutral one. What distinguishes a non-abelian gauge theory from an abelian one, is that in the former theory, direct interactions occur between gauge fields, carrying specific quantum numbers, unlike in the latter, as the gauge field – the photon – being uncharged.

As early as 1956, Schwinger believed that the weak and electromagnetic interactions should be combined into a gauge theory [97, 159, 199]. Here we may pose to note that both in electrodynamics and in the modified Fermi theory, interactions are mediated by vector particles. They are both described by universal dimensionless coupling constants e, and by g (see (1.1)) in the intermediate vector boson description, respectively. In a unified description of electromagnetism and the weak interaction, one expects these couplings to be comparable, i.e.,

$$g^2 \approx \mathrm{e}^2 = 4\pi\,\alpha, \qquad \text{where} \quad \alpha \approx \frac{1}{137}, \quad \mathrm{G_F} \approx 1.166\times 10^{-5}/(\mathrm{GeV})^2. \tag{1.7}$$

in units $\hbar = 1, \mathrm{c} = 1$. From (1.1), we may then estimate the mass of the W bosons to be

$$M_W \approx \sqrt{\frac{4\pi\,\alpha}{\mathrm{G_F}}} \approx 90\ \mathrm{GeV/c^2}, \tag{1.8}$$

[29]Non-abelian refers to the fact that the generators do not commute. In contrast a $\mathrm{U}(1)$ gauge theory, such as quantum electrodynamics, is an abelian one.

[30]$\mathrm{SU}(2)$ consists of 2×2 unitary matrices of determinant one. (The letter S in the group stands for the special property of determinant one.) It involves three generators, with which are associated three gauge fields. This will be studied in detail in Sect. 6.1.

re-inserting the constant c for convenience, in good agreement with the observed mass. We may also estimate the range of the weak interaction to be

$$R_W \approx \frac{\hbar \mathrm{c}}{M_W \mathrm{c}^2} \approx 2.2 \times 10^{-16} \mathrm{cm}. \tag{1.9}$$

Glashow, a former graduate student of Schwinger, eventually realized [96][31] the important fact that the larger group $\mathrm{SU}(2) \times \mathrm{U}(1)$, is needed to include also electrodynamics within the context of a Yang-Mills-Shaw theory. A major problem remained: the local gauge symmetry required that the gauge fields associated with the group must, a priori, be *massless* in the initial formulation of the theory.

The problem of the masslessness of the vector bosons was solved by Weinberg [236, 238] and Salam [182, 183],[32] by making use of a process,[33] referred to as spontaneous symmetry breaking, where a scalar field interacting with the vector bosons, whose expectation value in the vacuum state is non zero, leads to the generation of masses to them.[34] This is referred to as the Higgs[35] mechanism, in which the group $\mathrm{SU}(2) \times \mathrm{U}(1)$ is spontaneously broken to the group $\mathrm{U}(1)$ with the latter associated with the photon, and, in the process, the other bosons, called $W^{\pm}, Z^0$, acquiring masses, thanks to the Higgs boson, and renormalizability may be achieved. The latter particle has been also called the "God Particle".[36] The mere existence of a neutral vector boson Z^0 implies the existence of a weak interaction component in the theory without a charge transfer, the so-called neutral currents. A typical process involving the neutral Z^0 boson exchange is in $\tilde{\nu}_\mu + e^- \rightarrow \tilde{\nu}_\mu + e^-$ shown in Fig. 1.5 not involving the muon itself. Neutral currents[37] have been observed,[38] and all the vector bosons have been observed[39] as well. It turned out that the theory with spontaneous symmetry did not spoil the renormalizability of

[31] See also Salam and Ward [189].

[32] See also Salam and Ward [187–189] and Salam [181].

[33] Some key papers showing how spontaneous symmetry breaking using spin 0 field may generate masses for vector bosons are: Englert and Brout [63], Englert et al. [64], Guralnik et al. [110], and Kibble [127].

[34] Apparently the Legendary Victor Weisskopf was not impressed by this way of generating masses. In his CERN publication [241], on page 7, 11th line from below, he says that this is an awkward way to explain masses and that he believes that Nature should be more inventive, but experiments may prove him wrong.

[35] Higgs [119–121]. This work followed earlier work of Schwinger [200], where he shows, by the exactly solvable quantum electrodynamics in two dimensions, that gauge invariance does not prevent the gauge field to acquire mass dynamically, as well as of the subsequent work of Anderson [7] in condensed matter physics.

[36] This name was given by Lederman and Teresi [137].

[37] Neutral current couplings also appear in Bludman's [23] pioneering work on an $\mathrm{SU}(2)$ gauge theory of weak interactions but did not include electromagnetic interactions.

[38] Hasert et al. [112, 113] and Benvenuti et al. [15].

[39] See, e.g., C. Rubbia's Nobel Lecture [176].

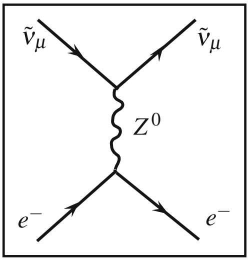

Fig. 1.5 A process involving the exchange of the neutral vector boson Z^0

the resulting theory with massive vector bosons. Proofs of renormalizability were given by 't Hooft [216, 217].[40] It seems that Sydney Coleman used to say that 't Hooft's proof has turned the Weinberg-Salam frog into an enchanted prince.[41] The "Electroweak Theory" turned up to be quite a successful theory.[42]

Another interaction which was also developed in the "image" of quantum electrodynamics was quantum chromodynamics, as a theory of strong interactions based, however, on the non-abelian gauge symmetry group SU(3). Here one notes that a typical way to probe the internal structure of the proton is through electron-proton scattering. The composite nature of the proton, as having an underlying structure, becomes evident when one compares the differential cross sections for elastic electron-proton scattering with the proton described as having a finite extension to the one described as a point-like particle. With a one photon exchange description, the form factors in the differential cross section are seen to vanish rapidly for large momentum transfer (squared) Q^2 of the photon imparted to the proton. As Q^2 is increased further one reaches the so-called resonance region,[43] beyond which, one moves into a deep inelastic region, where experimentally the reaction changes "character", and the corresponding structure functions of the differential cross section have approximate scaling properties (Sect. 6.9), instead of the vanishing properties encountered with elastic form factors, the process of which is depicted in Fig. 1.6. Such properties indicate the presence of approximately free point-like structures within the proton referred as partons, which consist of quarks, gluons together with those emitted[44] from their scattering reactions. This led to the development of the so-called parton model,[45] as a first approximation, in which these point-like particles within the proton are free and the virtual photon interacts

[40] See also 't Hooft and Veltman [221], Lee and Zinn-Justin [139–142], and Becchi et al. [13].

[41] See Salam [183], p. 529.

[42] The basic idea of the renormalizability of the theory rests on the fact that renormalizability may be established for the theory with completely massless vector bosons, as in QED, one may then invoke gauge symmetry to infer that the theory is also renormalizable for massive vector bosons via spontaneous symmetry breaking.

[43] A typical resonance is Δ^+, of mass 1.232 GeV, consisting of a proton p and a π^0 meson.

[44] See, e.g., Fig. 6.7c.

[45] Feynman [73, 74] and Bjorken and Pachos [21].

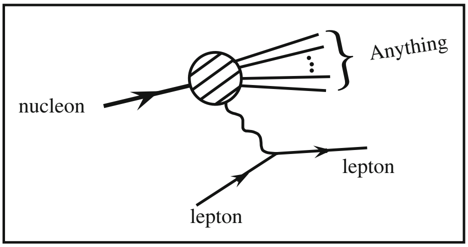

Fig. 1.6 In the process, "Anything" denotes anything that may be created in the process consistent with the underlying conservation laws. The wavy line denotes a neutral particle (γ, Z^0,...) of large momentum transfer

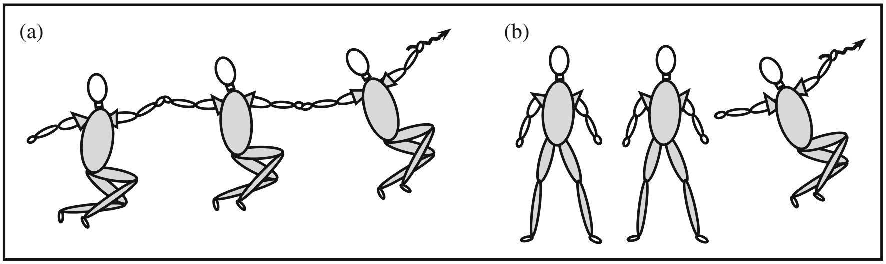

Fig. 1.7 (**a**) If interactions between quarks may be represented, as an analogy, by people holding hands, then pulling one person would drag everybody else along. In the parton model, the situation is represented as in part (**b**) rather than in part (**a**)

with each of its charged constituents independently,[46] instead of interacting with the proton as a whole.

The non-abelian gauge symmetry group SU(3), is needed to accommodate quarks and gluons, involving eight generators with which the gluons are associated. Here, in particular, a quantum number referred to as "color" (three of them)[47] is assigned to the quarks. One of the many reasons for this is that the spin 3/2 particle Δ^{++}, which is described in terms of three identical quarks (the so-called u quarks) as a low lying state with no orbital angular momentum between the quarks, behaves as a symmetric state under the exchange of two of its quarks and would violate the Spin & Statistics connection without this additional quantum number. The color degrees of freedom are not observed in the hadronic states themselves and the latter behave as scalars, that is they are color singlets, under SU(3) transformations. As the group SU(3) involves "color" transformations within each quark flavor, it may be denoted by $\mathrm{SU(3)_{color}}$ or just by $\mathrm{SU(3)_C}$. The gluons also carry "color" and direct gluon-gluon interactions then necessarily occur, unlike the situation with

[46]For an analogy to this, see part (b) of Fig. 1.7b.

[47]Greenberg [104], Han and Nambu [111], Nambu [163], Greenberg and Zwanziger [105], Gell-Mann [92], and Fritzsch and Gell-Mann [87].

photons in quantum electrodynamics since photons do not carry a charge. These gluon-gluon interactions turn out to have an anti-screening effect on a source field which dominate over the screening effect of quark/antiquark interactions leading to the interesting fact that the effective coupling of quark interactions becomes smaller at high-energies, and eventually vanish[48] – a phenomenon referred to as asymptotic freedom. This has far reaching consequences as it allows one to develop perturbation theory at high energies, in the effective coupling, and carry out various applications which were not possible before the development of the theory, and is consistent with deep-inelastic experiments of leptons with nucleons, with the latter described by point-like objects which, at high energies, scatter almost like free particles,[49] as mentioned above, the process of which is shown Fig. 1.6.

A particular experiment which indirectly supports the idea of quarks having spin 1/2 stems from e^-e^+ annihilation to a quark-antiquark, in the center of mass system. One would naïvely expect that the quark and the antiquark will emerge from the process moving in opposite directions[50] on their ways to detectors and will be observed. This is not, however, what happens and instead two narrow jets of hadrons emerge, moving back-yo-back, with the net jet-axis angular distribution consistent with a spin 1/2 character of the quark/antiquark parents sources.

The electroweak theory and quantum chromodynamics together constitute the so-called standard model[51] with underlying gauge symmetry group $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$.

The effective coupling in quantum chromodynamics is expected to become larger at large distances increasing with no bound providing a strong confining force of quarks and gluons restricting them within hadrons – a phenomenon that is sometimes referred to as *infrared slavery*.[52]

From our discussion of quantum electrodynamics, we recall that the effective coupling of a $\mathrm{U}(1)$ gauge theory, as an abelian gauge theory, increases with energies. On the other hand, the asymptotic free nature of non-abelian gauge theories, imply that the effective couplings associated with the groups $\mathrm{SU}(3)$, $\mathrm{SU}(2)$, decrease with energy. Due to the smallness of the $\mathrm{U}(1)$ coupling in comparison to the other two at the present low energies, this gives one the hope that at sufficient high energies these three couplings merge together and the underlying theory would be described by one single force. A theory which attempts to unify the electroweak and the strong interactions is called a *grand unified theory*. Such theories have been developed[53] and the couplings seem to run to merge roughly

[48] This was discovered by Gross and Wilczek [106] and Politzer [172]. See also Vanyashin and Terentyev [229] for preliminary work on vector bosons.

[49] Chromodynamics means Colordynamics, and the name Quantum Chromodynamics is attributed to Gell-Mann, see, e.g., Marciano and Pagels [158].

[50] See, e.g., Fig. 6.7d, and Sect. 6.10.3

[51] The name "Standard Model" is usually attributed to Weinberg.

[52] Unfortunately, no complete proof of this is available.

[53] See, e.g., Georgi et al. [94] for pioneering work. See also Beringer et al. [16] and Olive et al. [167].

somewhere around 10^{15}–10^{16} GeV. This, in turn, gives the hope of the development of a more fundamental theory in which gravitation, which should be effective at energy scale of the order of Planck energy scale $\sqrt{\hbar c/G_N} \simeq 10^{19}\,\text{GeV}$,[54] or even at a lower energy scale, where G_N denotes Newton's gravitational constant, is unified with the electroweak and strong forces. If the standard model is the low energy of such a fundamental theory, then the basic question arises as to what amounts for the enormous difference between the energy scale of such a fundamental theory ($\sim 10^{16}$–$10^{19}\,\text{GeV}$) and the defining energy scale of the standard model ($\sim 300\,\text{GeV}$)? This has been termed as the *hierarchy* problem which will be discussed again later. We will see in Vol. II, in particular, that the above mentioned couplings seem to be unified at a higher energy scale of the order $10^{16}\,\text{GeV}$, when supersymmetry is taken into account, getting it closer to the energy scale at which gravitation may play an important role.

One may generalize the symmetry group of the standard model, and consider transformations which include transformations between quarks and leptons, leading to a larger group such as, for example, to the $\mathrm{SU}(5)$ group, or a larger group, which include $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$. The advantage of having one larger group is that one would have only one coupling parameter and the standard model would be recovered by spontaneous symmetry breaking at lower energies. This opens the way to the realization of processes in which baryon number is not conserved, with a baryon, for example, decaying into leptons and bosons. The experimental[55] bound on lifetime of proton decay seems to be $>10^{33}$ years and is much larger than the age of the universe which is about 13.8 billion years.[56] Such a rare event even if it occurs once will give some support of such grand unified theories.

We have covered quite a large territory and before continuing this presentation, we pose for a moment, at this appropriate stage, to discuss three aspects of importance that are generally expected in order to carry out reliable computations in perturbative quantum field theory. These are:

1. The development of a powerful and simple formalism for doing this.
2. To show how the renormalization process is to be carried, and establish that the resulting expressions are finite to any order of perturbation theory.
3. The physical interpretation will be completed if through the process of renormalization, the initial experimentally unattainable parameters in the theory are eliminated in favor of physically observed ones, which are finite in number, and are generally determined experimentally as discussed earlier in a self consistent manner.

---

[54] The Planck energy (mass) will be introduced in detail later.

[55] See, e.g., Olive et al. [167].

[56] A decay of the proton may have a disastrous effect in the stability of matter over anti-matter itself in the universe. See, however, the discussion given later on the dominance of matter in the visible universe.

Perturbatively renormalizable theories are distinguished from the non-renormalizable ones by involving only a finite number of parameters in the theory that are determined experimentally.

We discuss each of these in turn.

1. A powerful formalism is the *Path Integral* one, pioneered by Feynman,[57] defining a generating functional for so-called Green functions from which physical amplitudes may be extracted, and has the general structure $\int \mathrm{d}\mu[\chi]\, \mathrm{e}^{\mathrm{iAction}}$. Here $\mathrm{d}\mu[\chi]$ defines a measure of integration over classical fields as the counterparts of the quantum fields of the theory.[58] "Action" denotes the classical action. In the simplest case, the measure $\mathrm{d}\mu[\chi]$ takes the form $\Pi_x\, \mathrm{d}\chi(x)$ as a product of all spacetime points. In gauge theories, due to constraints, the determination of the measure of integrations requires special techniques[59] and takes on a much more complicated expression and was successfully carried out by Faddeev and Popov in 1967.[60] The path integral expression as it stands, involves continual integrations to be carried out.

   An equally powerful and quite an elegant formalism is due to Schwinger,[61] referred to as the *Action Principle* or the *Quantum Dynamical Principle*. For quantum field theory computations, the latter gives the variation of the so-called vacuum-to-vacuum transition amplitude (a generating functional): $\delta\langle\, 0_+ \,|\, 0_- \rangle$ as any of the parameters of the theory are made to vary. The latter is then expressed as a differential operator acting on a simple generating functional expressed in closed form. This formalism involves only functional differentiations to be carried out, no functional integrations are necessary, and hence is relatively easier to apply than the path integral. We will learn later, for example, that the path integral may be simply obtained from the quantum dynamical principle by a functional Fourier transform thus involving functional integrals. Again the application of the quantum dynamical principle to the quantization of gauge theories with underlying constraints require special techniques and it was carried out in Manoukian [150].

   All the fundamental interactions in nature are presently described theoretically by gauge theories, involving *constraints*, and the two approaches of their quantization discussed above will be both treated in this book for pedagogical reasons and are developed [62]

[57] See Feynman and Hibbs [78] for the standard pedagogical treatment. See also Feynman [75].

[58] We use a general notation $\chi(x)$ for the fields as functions of spacetime variable and suppress all indices that they may carry to simplify the notation. These fields may include so-called Grassmann fields.

[59] Feynman [72], DeWitt [43, 44], and Faddeev and Popov [66].

[60] *Op. cit.*

[61] Schwinger [194–197, 201].

[62] There is also the canonical formalism, see, e.g., Mohapatra [161, 162] and Utiyama and Sakamoto [228].

via the

Path Integral [66],

or via the

Action Principle (Quantum Dynamical Principle) [150].

2. Historically, Abdus Salam, was the first "architect" of a general theory of renormalization. In 1951, he carried out a systematic study[63] of renormalization [180], introduced and sketched a subtraction scheme in a general form. Surprisingly, this classic paper was not carefully reexamined until much later. This was eventually done in 1976 [147] by Manoukian, and inspired by Salam's work, a subtraction scheme was developed and brought to a mathematically consistent form, and the finiteness of the subtracted, i.e., renormalized, theory was proved by the author[64] to *any* order of perturbation theory.[65] by using, in the process, a power counting theorem established by Weinberg [235] for integrals of a special class of functions, thus completing the Dyson-Salam program. The subtraction was carried out directly in momentum space and no cut-offs were introduced.

   Shortly after the appearance of Salam's work, two other "architects" of a general theory of renormalization theory, Bogoliubov and Parasiuk, in a classic paper in 1957 [24], also developed a subtraction scheme. In 1966 [118], Hepp gave a complete proof of the finiteness of the Bogoliubov-Parasiuk to any order of perturbation theory, by using in the intermediate stages ultraviolet cut-offs, and in 1969 [251], Zimmermann formulated their scheme in momentum space, without cut-offs, and provided a complete proof of finiteness as well, thus completing the Bogoliubov-Parasiuk program. This scheme is popularly known as the BPHZ scheme.

   The equivalence of the Bogoliubov-Parasiuk scheme, in the Zimmermann form, and our scheme was then proved by Manoukian,[66] after some systematic cancelations in the subtractions. This equivalence theorem[67] unifies the two monumental approaches of renormalization.[68]

---

[63] Salam [180], see also Salam [179].

[64] Manoukian [149].

[65] For a pedagogical treatment of all these studies, see my book "Renormalization" [149]. This also includes references to several of my earlier papers on the subject as well as many results related to renormalization theory.

[66] See Manoukian [149] *op. cit.*

[67] This result has been also referred to as "Manoukian's Equivalence Principle", Zeidler [250], p. 972. See also Streater [207].

[68] I was pleased to see that our equivalence theorem has been also considered, by completely different methods, by Figueroa and Gracia-Bondia [80]. For other earlier, and recent, but different, approaches to renormalization theory, see, e.g., Epstein and Glaser [65], Kreimer [131, 132], Connes and Kreimer [34, 35], and Figueroa and Gracia-Bondia [79, 81]. See also Landsman [135] and Aschenbrenner [11].

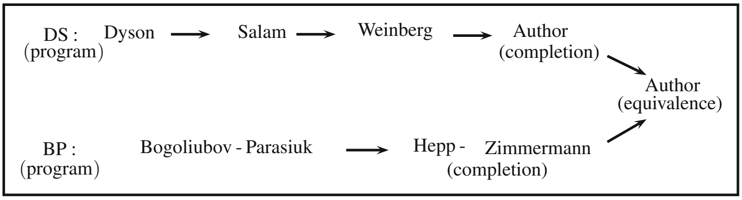

Fig. 1.8 Developments of the general theory of renormalization from the DS and BP programs. The intricacies of this layout also appear in Zeidler [250], pp. 972–975. Regarding the author's work shown in the above layout and of his completion of the renormalization program stemming out of Salam's, Streater [207] writes: "***It is the end of a long chapter in the history of physics***"

The development of the general theory of renormalization from the DS and BP programs may be then summarized as given in Fig. 1.8[69].

3. The physical interpretation of the theory is completed by showing that the subtractions of renormalization are implemented by counterterms in the theory which have the same structures as the original terms in the theory (i.e., in the Lagrangian density),[70] thus establishing the self-consistency involved in the elimination of the initial parameters in the theory in favor of physically observed ones. As mentioned above, for a theory to be renormalizable, i.e., involving only a finite number of parameters determined, in general, experimentally, the counterterms of the theory must be finite in number as well.

All particles due to their energy content experience the gravitational attraction. Einstein's theory of gravitation, also referred to as general relativity (GR), is described by a second rank tensor with the energy-momentum tensor of matter as its source from which the energy density of matter may be defined. It may not be described just by a scalar or just by a vector field as they are inconsistent with experiment. It is easy to see that due to the fact that masses have the same signs (positive) a theory based on a vector field alone will lead to a repulsive rather an attractive gravitational force.[71] GR theory predictions are well supported experimentally in our solar system.

The key observation, referred to as the principle of equivalence, of Einstein is that at any given point in space and any given time, one may consider a frame in which gravity is wiped out at the point in question. For example, in simple Newtonian gravitational physics, a test particle placed at a given point inside a freely falling elevator on its way to the Earth, remains at rest, inside the elevator, for a very short time, depending on the accuracy being sought, and, depending on its position relative to the center of the Earth, eventually moves, in general, from its original position in a given instant. Einstein's principle of equivalence applies only locally

[69] The layout in Fig. 1.8 is based on Manoukian [149], and it also appears in Zeidler [250], p. 974. See also Streater [207] and Figueroa and Gracia-Bondia [80].

[70] For a detailed study of this see, Manoukian ([148]; Appendix, p. 183 in [149]).

[71] Attempts have been made to include such fields as well for generalizations of Einstein's theory, but we will not go into it here.

at a given point and at a given time. At the point in question, in the particular frame in consideration, gravity is wiped out and special relativity survives. The reconciliation between special relativity and Newton's theory of gravitation, then readily leads to GR, where gravity is accounted for by the curvature of spacetime and its departure from the flat spacetime of special relativity one has started out with upon application of the principle of equivalence. By doing this, one is able to enmesh non-gravitational laws with gravity via this principle.

Quantum gravity (QG) is needed in early cosmology, black hole physics, and, in general, to deal with singularities that arise in a classical treatment. QG must also address the problem of the background geometry. A common interest in fundamental physics is to provide a unified description of nature which is applicable from microscopic to cosmological distances. A fundamental constant of unit of length that is expected to be relevant to this end is the Planck length as well as the Planck mass. Out of the fundamental constants of quantum physics $\hbar$, of relativity c, and the Newtonian gravitational one $\mathrm{G_N}$, we may define a unit of length and mass, the Planck length and Planck mass, respectively, relevant in quantum gravity, through the following

$$\ell_{\mathrm{P}} = \sqrt{\frac{\hbar \mathrm{G_N}}{\mathrm{c}^3}} \simeq 1.616 \times 10^{-33}\,\mathrm{cm}, \quad m_{\mathrm{P}} = \sqrt{\frac{\hbar \mathrm{c}}{\mathrm{G_N}}} \simeq 1.221 \times 10^{19}\,\mathrm{GeV/c^2}. \tag{1.10}$$

In units $\hbar = 1$, $\mathrm{c} = 1$, dimensions of physical quantities may be then expressed in powers of mass ([Energy] $=$ [Mass], [Length] $=$ [Mass]$^{-1}$ $=$ [Time], ...), and, as gravitation has a universal coupling to all forms of energy, one may hope that it may be implemented within a unified theory of the four fundamental interactions, with the Planck mass providing a universal mass scale. Unfortunately, it is difficult experimentally to investigate the quantum properties of spacetime as one would be working at very small distances.

GR predicts the existence of Black holes. Here it is worth recalling of the detection ("Observational waves from a binary black hole merger", Phys. Rev. Lett. 116, 061102 (1–16) (2016)) by B. P. Abbott et al. of gravitational waves from the merger of two black holes 1.3 billion light-years from the Earth. Recall that a black hole (BH) is a region of space into which matter has collapsed and out of which light may not escape. It partitions space into an inner region which is bounded by a surface, referred to as the event horizon which acts as a one way surface for light going in but not coming out. The sun's radius is much larger than the critical radius of a BH which is about 2.5 km to be a black hole for the sun. We will see that for a spherically symmetric BH of mass M, the radius of the horizon is given by $R_{\mathrm{BH}} = 2\mathrm{G_N} M/\mathrm{c}^2$.[72]

[72]This may be roughly inferred from Newton's theory of gravitation from which the escape speed of a particle in the gravitational field of a spherically symmetric massive body of mass M, at a distance r, is obtained from the inequality $v^2/2 - \mathrm{G_N} M/r < 0$, and by formally replacing v by

One may argue that the Planck length may set a lower limit spatial cut-off. The following formal and rough estimates are interesting. Suppose that by means of a high energetic particle of energy E, $\langle E^2\rangle \sim \langle p^2\rangle c^2$, with $\langle p^2\rangle$ very large, one is interested in measuring a field within an interval of size δ around a given point in space. Such form of energy acts as an effective gravitational mass $M \sim \sqrt{\langle E^2\rangle/c^4}$ which, in turn distorts space around it. The radius of the event horizon of such a gravitation mass M is given by $r_{\mathrm{BH}} = 2\mathrm{G_N}M/c^2$. Clearly we must have $\delta > r_{\mathrm{BH}}$, otherwise the region of size δ that we wanted to locate the point in question will be hidden beyond a BH horizon, and localization fails. Also:
$\langle p^2\rangle \geq \langle (p - \langle p\rangle)^2\rangle \geq \hbar^2/4\delta^2$. Hence $M \geq \hbar/2c\delta$,

$$\delta > \frac{2\mathrm{G_N}M}{c^2} \geq \frac{\hbar \mathrm{G_N}}{c^3\delta},$$

which gives $\delta > r_{\mathrm{BH}} = \sqrt{\hbar \mathrm{G_N}/c^3} = \ell_{\mathrm{P}}$.

Interesting investigations by Hawking[73] have shown that a BH is not really a black body, it is a thermodynamic object, it radiates and has a temperature associated with it.[74] In Chapter 7 in Vol. II , we will see, considering a spherically symmetric BH, that its temperature is given by[75]

$$T_{\mathrm{BH}} = \frac{\hbar c^3}{8\,\pi\,\mathrm{G_N}M\,\mathrm{k_B}}. \tag{1.11}$$

where $\mathrm{k_B}$ is the Boltzmann constant. Note that a very massive black hole is cold.

Recall that entropy S represents a measure of the amount of disorder with information encoded in it, and invoking the thermodynamic interpretation of a BH, we may write

$$\frac{\partial S}{\partial (Mc^2)} = \frac{1}{T}, \tag{1.12}$$

which upon integration with boundary condition that for $M \to 0$, $S \to 0$, gives the celebrated result

$$S_{\mathrm{BH}} = \frac{c^3\mathrm{k_B}}{4\hbar \mathrm{G_N}}A = \mathrm{k_B}\frac{A}{4\,\ell_{\mathrm{P}}^2}, \qquad A = 4\pi\left(\frac{2\mathrm{G_N}M}{c^2}\right)^2 \tag{1.13}$$

---

the ultimate speed c to obtain for the critical radius $R_{\mathrm{critical}} = 2\mathrm{G_N}M/c^2$ such that for $r < R_{\mathrm{critical}}$ a particle cannot escape.

[73] Hawking [114, 115].

[74] Particle emission from a BH is formally explained through virtual pairs of particles created near the horizon with one particle falling into the BH while the other becoming free outside the horizon.

[75] A pedestrian approach in determining the temperature is the following. By comparing the expression of energy expressed in terms of the wavelength of radiation λ: $E = hc/\lambda$, with the expression $E = \mathrm{k_B}T$, gives $T = hc/\mathrm{k_B}\lambda$. On dimensional grounds $\lambda \sim 2\,\mathrm{G_N}M/c^2$, which gives $T \sim \pi\hbar c^3/\mathrm{G_N}M\,\mathrm{k_B}$. This is the expression given for the temperature up to a proportionality constant.

referred to as the Bekenstein-Hawking Entropy formula[76] of a BH. This relation relates quantum gravity to information theory. This result is expected to hold in any consistent formulation of quantum gravity, and shows that a BH has entropy unlike what would naïvely expect from a BH with the horizon as a one way classical surface through which information is lost to an external observer. The proportionality of the entropy to the area rather than to the volume of a BH horizon should be noted. It also encompasses Hawking's theorem of increase of the area with time with increase of entropy. We will discuss the Bekenstein-Hawking Entropy formula below in conjunction with loop quantum gravity and string theory.

Now we turn back to the geometrical description of gravitation given earlier, and introduce a gravitational field to account for the departure of the curved spacetime metric from that of the Minkowski one to make contact with the approaches of conventional field theories, dealing now with a field permeating an interaction between all dynamical fields. The quantum particle associated with the gravitational field, the so-called graviton, emerges by considering the small fluctuation of the metric, associated with curved spacetime of GR about the Minkowski metric, as the limit of the full metric, where the gravitational field becomes weaker and the particle becomes identified. This allows us to determine the graviton propagator in the same way one obtains, for example, the photon propagator in QED, and eventually carry out a perturbation theory as a first attempt to develop a quantum theory of gravitation.

In units of $\hbar\ = 1, \mathrm{c} = 1$, Newtons gravitational constant $\mathrm{G_N}$, in 4 dimensional spacetime, has the dimenionality $[\,\mathrm{G_N}\,] = [\,\mathrm{Mass}\,]^{-2}$, which is a dead give away of the non-renormalizabilty of a quantum theory of gravitation based on GR. The non-renormalizability of the theory is easier to understand by noting that the divergences, in general, tend to increase as we go to higher orders in the gravitational coupling constant without a bound, implying the need of an infinite of parameters need to be fixed experimentally[77] and hence is not of any practical value. Some theories which are generalizations of GR, involving higher order derivatives, turn out to be renormalizable[78] but violate, in a perturbative setting, the very sacred principle of positivity condition of quantum theory. Unfortunately, such a theory involves ghosts in a perturbative treatment, due to the rapid damping of the propagator at high energies faster than $1/\mathrm{energy}^2$, and gives rise, in turn, to negative probabilities.[79]

One is led to believe that Einstein's general relativity is a low energy effective theory as the low energy limit of a more complicated theory, and as such it provides a reliable description of gravitation at low energies. Moreover, one may argue that the non-renormalizability of a quantum theory based on GR is due to the fact that one is trying to use it at energies which are far beyond its range of validity. As a

[76]Bekenstein [14].

[77]Manoukian [149] and Anselmi [9].

[78]Stelle [204].

[79]Unitarity (positivity) of such theories in a non-perturbative setting has been elaborated upon by Tomboulis [224].

matter of fact the derivatives occurring in the action, in a momentum description via Fourier transforms, may be considered to be small at sufficiently low energies. In view of applications in the low energy regime, one then tries to separate low energy effects from high energy ones even if the theory has unfavorable ultraviolet behavior such as in quantum gravity.[80] Applications of such an approach have been carried out in the literature as just cited, and, for example, the modification of Newton's gravitational potential at long distances has been determined to have the structure

$$U(r) = -\frac{\mathrm{G_N} m_1 m_2}{r}\Big[1 + \alpha \frac{\mathrm{G_N}(m_1 + m_2)}{\mathrm{c}^2\, r} + \beta \frac{\mathrm{G_N}\hbar}{\mathrm{c}^3\, r^2}\Big], \tag{1.14}$$

for the interaction of two spin 0 particles of masses m_1 and m_2. Here α, β, are dimensionless constants,[81] and the third term represents a quantum correction being proportional to $\hbar$.

Conventional quantum field theory is usually formulated in a fixed, i.e., in, a priori, given background geometry such as the Minkowski one. This is unlike the formalism of "Loop Quantum Gravity" (LQG) also called "Quantum Field Theory of Geometry". The situation that we will encounter in this approach is of a *quantum field theory in three dimensional space*, which is a non-perturbative background independent formulation of quantum gravity. The latter means that no specific assumption is made about the underlying geometric structure and, interestingly enough, the latter rather emerges from the theory. Here by setting up an eigenvalue equation of, say, an area operator, in a quantum setting, one will encounter a granular structure of three-dimensional space yielding a discrete spectrum for area measurements with the smallest possible having a non-zero value given to be of the order of the Planck length squared: $\hbar \mathrm{G_N}/c^3 \sim 10^{-66}\,\mathrm{cm}^2$.[82] The emergence of space in terms of "quanta of geometry", providing a granular structure of space, is a major and beautiful prediction of the theory. The 3 dimensional space is generated by a so-called time slicing procedure of spacetime carried out by Arnowitt, Deser and Misner.[83] The basic field variables in the theory is a gravitational "electric" field, which determines the geometry of such a 3 dimensional space and naturally emerges from the definition of the area of a surface in such a space, and its canonical conjugate variable is referred to as the connection. By imposing equal time commutation relation of these two canonically conjugate field variables, the quantum version of the theory arises, and the fundamental problem of the quantization of geometry follows. The basic idea goes to Penrose [171] whose interest was to construct the concept of space from combining angular momenta. It is also interesting that the proportionality of entropy and the surface area of the BH horizon in the Bekenstein-Hawking Entropy formula has been derived in loop quantum gravity.[84]

---

[80]Donoghue [55–57] and Bjerrum-Bohr et al. [19], Bjerrum-Bohr et al. [20].

[81]Recent recorded values are $\alpha = 3$, and $\beta = 41/10\pi$ Bjerrum-Bohr et al. [19].

[82]Rovelli and Smolin [174], Ashtekar and Lewandoski [12], and Rovelli and Vidotto [175].

[83]Arnowitt et al. [10].

[84]See, e.g., Meissner [160] and Ansari [8].

Supersymmetry is now over 40 years old. Supersymmetry provides a symmetry between fermions and bosons. Borrowing a statement made by Dirac, speaking of theories, in general, it is *a theory with mathematical beauty*.[85] The name "Supersymmetry" for this symmetry is attributed to Salam and Strathdee as it seemed to have first appeared in the title of one of their papers [184]. An abbreviated name for it is SUSY, as some refer to this symmetry. The latter is not only beautiful but is also full of thought-provoking surprises. To every degree of freedom associated with a particle in the standard model, in a supersymmetric version, there corresponds a degree of freedom associated with a partner, referred to as a sparticle, with the same mass and with opposite statistics to the particle.[86] Unlike other discoveries, supersymmetry was not, a priori, invented under pressure set by experiments and was a highly intellectual achievement. Theoretically, however, it quickly turned out to be quite important in further developments of quantum field theory. For one thing, in a supersymmetric extention of the standard model, the electroweak and strong effective couplings do merge at energies about 10^{16} GeV, signalling the possibility that these interactions are different manifestations of a single force in support of a grand unified theory of the fundamental interactions. Also gravitational effects are expected to be important at the quantum level at the Planck energy of the order 10^{19} GeV, or possibly at even lower energies, giving the hope of having a unified theory of the four fundamental interactions at high energies. Supersymmetry leads to the unification of coupling constants. SUSY also tends to "soften" divergences of a theory in the sense that divergent contributions originating from fermions tend to cancel those divergent contributions originating from bosons due to their different statistics.

One of the important roles that supersymmetry may play in a supersymmetric extension of the standard problem is in the so-called *hierarchy* problem. The basic idea of a facet of this is the following. A fundamental energy scale arises in the standard model from the vacuum expectation value of the Higgs boson which sets the scale for the masses in the theory, such as for the masses of the vector bosons. It turns out that the masses imparted to the initially massless vector bosons, for example, via the Higgs mechanism, using the parameters in the Lagrangian density are in very good agreement with experimental results. On the other hand, if one introduces a large energy scale cut off $\kappa \sim 10^{15}$ GeV, of the order of a grand unified energy scale, or the Planck energy scale 10^{19} GeV, at which gravitation may play a significant role, to compute the shift of the squared-mass of the Higgs boson, as a scalar particle, due to the dynamics (referred to as radiative corrections), it turns out to be quadratic[87] in κ, which is quite large for such a large cut-off. This requires that the bare mass squared of the Higgs boson to be correspondingly large to cancel

[85] Here we recall the well known statement of Dirac, that a theory with mathematical beauty is more likely to be correct than an ugly one that fits some experimental data [53].

[86] This is such that the total number of fermion degrees of freedom is equal to the total bosonic ones.

[87] See, e.g., Veltman [230].

such a quadratic dependence on κ and obtain a physical mass of the Higgs boson of the order of magnitude of the minute energy[88] ($\sim 10^2\,\text{GeV}$), in comparison, characterizing the standard model, and this seems quite unnatural for the cancelation of such huge quantities.[89] This unnatural cancelation of enormously large numbers has been termed a facet of the hierarchy problem. Supersymmetric theories have, in general, the tendency to cancel out such quadratic divergences, up to possibly of divergences of logarithmic type which are tolerable, thus protecting a scalar particle from acquiring such a large bare mass. So supersymmetry may have an important role to play here.

SUSY relates fermions to bosons, and vice versa, and hence a generator is required which is of fermionic type, that is, it carries a spinor index as in the Dirac field[90] to carry out a transformation fermion $\leftrightarrow$ boson. Since the spins of fermions and bosons are different, this necessarily means that such a generator does *not* commute with the angular (spin) momentum operator as supersymmetry unites particles of the same mass and different spins into multiplets. Bosons and fermions have, in general, different masses, which means that SUSY is to be spontaneously broken if such a symmetry is to have anything to do with nature. If supersymmetry breaking sets at such an energy scale as 1 TeV or so, then some of the lowest mass superpartners may be hopefully discovered.[91]

Of particular interest was also the development of the superspace concept as an extension of the Minkowski one, where one includes an additional degree of freedom usually denoted by[92] $\theta = (\theta_a)$ to the space coordinates $(t, \mathbf{x})$, which turns out to be quite convenient in defining and setting up SUSY invariant integrals such as the action of a dynamical system.[93] To describe dynamics this, in turn, necessitates to introduce superfields of different types[94] as functions of these variables.

The extension of the algebra of the Poincaré group to a superalgebra was first carried out by Gol'fand and Likhtman in [100] to construct supersymmetric field theory models, and with the implementation of spontaneous symmetry breaking

[88] Aad et al. [1] and Chatrchyan et al. [31].

[89] As mentioned earlier, the question, in turn, arises as to what amounts for the enormous difference between the energy scale of grand unification and the energy scale that characterizes the standard model.

[90] This point is of importance because an earlier attempt by Coleman and Mandula [33] to enlarge the Poincaré group did not work. They considered only so-called "Bose" generators (that is tensors, and not spinors) in their analysis.

[91] Perhaps an optimist would argue that since antiparticles corresponding to given particles were discovered, the discovery of superpartners associated with given particles would not be out of the question either. The underlying symmetries involved in these two cases are, however, quite of different nature.

[92] This is called a Grassmann variable.

[93] See Salam and Strathdee [185, 186].

[94] Details on superfields will be given in Vol. II. The *explicit* expression of the pure vector superfield has been recently obtained in Manoukian [155].

by Volkov and Akulov in [232]. In [243], Wess and Zumino also,[95] independently, developed supersymmetric models in 4 dimensions, and this work has led to an avalanche of papers on the subject and to a rapid development of the theory. In particular, supersymmetric extensions of the standard model were developed,[96] supergravity, as a supersymmetric extension of gravitational theory, was also developed.[97] Unfortunately, things do not seem to be much better for supergravity, as far as its renormalizability is concerned.[98]

Now we come to String Theory. String Theory is a theory which attempts to provide a unified description of all the fundamental interactions in Nature and, in particular, give rise to a consistent theory of quantum gravity. A string is a fundamental one dimensional extended object, and if it has to do with quantum gravity, it is, say, of the order of the Plank length $\ell_{\mathrm{P}} = \sqrt{\mathrm{G_N}\,\hbar/\mathrm{c}^3} \sim 10^{-33}\,\mathrm{cm}$, involving the three fundamental constants: Newton's gravitational constant $\mathrm{G_N}$, the quantum unit of action $\hbar$, and the speed of light c. Since no experiments can probe distances of the order of the Planck length, such a string in present day experiments is considered to be point-like. When a string, whether closed or open, moves in spacetime, it sweeps out a two dimensional surface referred to as a worldsheet. String Theory is a *quantum field theory which operates on such a two dimensional worldsheet*. This, as we will see, has remarkable consequences in spacetime itself, albeit in higher dimensions. Particles are identified as vibrational modes of strings, and a single vibrating string may describe several particles depending on its vibrational modes. Strings describing bosonic particles are referred to as a bosonic strings, while those involving fermionic ones as well are referred to as superstrings. The remarkable thing is that the particles needed to describe the dynamics of elementary particles *arise* naturally in the mass spectra of oscillating strings, and are not, a priori, assumed to exist or put in by hand in the underlying theories. The *dimensionality* of the spacetime in which the strings live are *predicted* by the underlying theory as well and are necessarily of higher dimensions than four for consistency with Lorentz invariance of spacetime at the quantum level, consisting of a dimensionality of 26 for the bosonic strings and a spacetime dimensionality of 10 for the superstrings. The extra dimensions are expected to curl up into a space that is too small to be detectable with present available energies. For example the surface of a hollow extended cylinder with circular base is two dimensional, with one dimension along the cylinder, and another one encountered as one moves on its circumference. If the radius of the base of the cylinder is relatively small, the cylinder will appear as one dimensional when viewed from a large distance (low energies). Accordingly, the extra dimensions in string theory are expected to be

---

[95] These basic papers, together with other key ones, are conveniently collected in Ferrara [71].

[96] See Fayet [67], Dimopoulos and Georgi [45].

[97] See Freedman, van Nieuwenhuizen and Ferrara [84], Deser and Zumino [42].

[98] See, e.g., Deser et al. [41], Deser [40], Stelle [205, 206], and Howe and Stelle [124].

small and methods, referred to as compactifications,[99] have been developed to deal with them thus ensuring that the "observable" dimensionality of spacetime is four.

Superstring theories involve fermions and are thus relevant to the real world, but there are, however, several superstring theories. Also unlike the loop quantum gravity, which provides a background independent formulation of spacetime with the latter emerging from the theory itself, as discussed earlier, the strings in string theories are assumed to move in a pre-determined spacetime, and thus spacetime plays a passive role in them.[100] A theory, referred to as M-Theory,[101] based on non-perturbative methods, is envisaged to unify the existing superstrings theories into one single theory, instead of several ones, and be related to them by various limiting and/or transformation rules, referred to as dualities,[102] and is of 11 dimensional spacetime. M-Theory is believed to be approximated by 11 dimensional supergravity,[103] and the spacetime structure is envisaged to emerge from the theory as well. Bosonic strings involve tachyonic states. This is unlike the situation in superstring theories in which supersymmetry plays a key role in their definitions, and a process referred to as a GSO projection method, ensuring the equality of the degrees of freedom of bosonic and fermionic states, as required by supersymmetry, in turn implies that no tachyonic states appear in the theory.[104]

String theory was accidentally discovered through work carried out by Veneziano in 1968 when he attempted to write down consistent explicit expressions of meson-meson scattering amplitudes in strong interactions physics.[105] This was, of course before the discovery of QCD. With the many excited states of mesons and baryons (resonances), it was observed experimentally that there exists a linear relationship between spin J and the mass M squared of a resonance given by a linear relationship

$$\text{with a universal slope :} \quad \frac{\mathrm{d}J}{\mathrm{d}M^2} = \alpha', \quad \alpha' \cong 1\,\mathrm{GeV}^{-2}, \tag{1.15}$$

defining so-called Regge trajectories. Veneziano postulated and wrote down a scattering amplitude of meson – meson scattering: $p_1(m_1) + p_2(m_2) \to p_3(m_3) + p_4(m_4)$, which, in particular, showed that the amplitude involves the exchange of an infinite number of particles (corresponding to arbitrary integer spins). This is unlike the situation in conventional field theory as QED or the standard model, where they involve the exchange of a finite number of particles to any given order. String theory shares this property of the Veneziano amplitude. As a matter of fact the Veneziano

[99] An idea used by Kaluza and Klein in their attempt to unify gravity and electromagnetism in a 5 dimensional generalization of general relativity.

[100] See also Horowitz [122].

[101] Townsend [227], Witten [244], and Duff [58].

[102] Duff [58] and Schwarz [191].

[103] Cremmer et al. [37].

[104] The GSO method of projection was proposed in Gliozzi et al. [98, 99].

[105] Veneziano [231], see also Lovelace and Squires [145] and Di Vecchia [54].

amplitude may be derived from string theory. Nambu [164], Nielsen [166] and Susskind [211] have shown that the famous expression of the amplitude postulated by Veneziano may be interpreted as a quantum theory of scattering of relativistic strings. Although, a priori, this was assumed to describe a strong interaction process, Yoneya [248], and Scherk and Schwarz [190] made use of the fact that string theory (involving closed strings) contains a spin 2 massless state, which was identified with the elusive graviton, in addition to a whole spectrum of other excitation modes, to propose that string theory provides a framework for the unification of general relativity and quantum mechanics. As early as 1971, Neveu and Schwarz [165], and Raymond [173] included fermions in their analyses, which eventually led to the notion of superstrings, and during a short period of time, several types[106] of superstrings were introduced in the literature.

Due to the assumed non-zero extensions of strings, it is hoped that they provide, naturally, an ultraviolet cut-off $\Lambda \sim (\ell_P)^{-1}$ and render all processes involving strings ultraviolet finite. This is unlike conventional quantum field theory interactions where all the quantum fields are multiplied locally at the same spacetime points, like multiplying distributions at the same point, and are, in this sense, quite troublesome.

In string theory, two strings with given vibrational modes, identifying two given particles, may combine forming one string with an arbitrary number of different vibrational modes associated with a myriad number of particles, defining generalized 3-vertices. The combined string may again split into two strings with associated vibrational modes, identified appropriately with two more particles, describing a scattering process of 2 particles $\rightarrow$ 2 particles. Thus interactions involve string worldsheets of various topologies arise.

Other extended objects are also encountered in string theory called branes which, in general, are of higher spatial dimensions than one, with the string defined as a one dimensional brane. For example, an open string, satisfying a particular boundary condition, referred to as a Dirichlet boundary condition, specifies a hypersurface, referred to as a D brane, on which the end points of the open string reside. On the other hand, the graviton corresponds to a vibrational mode of closed strings, and since the latter, having no ends, may not be restricted to a brane and moves away from it. This might explain the weakness of the gravitational field, if our universe is a 3 dimensional brane embedded in a higher dimensional spacetime. Massless particles encountered in string theory are really the physically relevant ones because of the large unit of mass $(\ell_P)^{-1} \sim 10^{19}$ GeV in attributing masses to the spectrum of massive particles.[107] As we will see a massless particle may acquire mass if,

---

[106] Green and Schwarz [102, 103] and Gross et al. [107, 108].

[107] A systematic analysis of all the massless field excitations encountered in both bosonic and superstrings are investigated in Manoukian [152–154], in their respective higher dimensional spacetimes, and include the determinations of the degrees of freedom associated with them. Note that in four dimensional spacetime the number of degrees of freedom (spin states) of non-scalar fields is always two. This is not true in higher dimensional spacetime. For example, the degrees of freedom associated with a massless vector particle is 8 in 10 dimensions, while for the graviton is

for example, the end points of the open string are attached to two different branes, instead of a single brane.

We will learn the remarkable facts that Einstein's general relativity as well as Yang-Mills field theory may be obtained from string theory.

Interesting high energy scattering amplitudes have been computed in string theory over the years,[108] which provide a hint that space may not be probed beyond the Planck length – a result shared with "loop quantum gravity". It is worth mentioning that the Bekenstein-Hawking Entropy relation has been also derived in string theory.[109]

In recent years much work has been done, which is worth mentioning here but rather briefly, indicating that general relationships may exist between field theories and string theories, and consequently considerable attention was given trying to make such a statement more and more precise, with the ultimate hope of providing, in turn, a consistent and acceptable quantum theory of gravitation relevant to our world but much work still remains to be done. In particular, much study has been made to study the equivalence relation between certain four dimensional gauge theories and superstring theories, referred to as the AdS/CFT correspondence, where AdS space stands for anti-de-Sitter space, and CFT stands for conformal field theory.[110] Such correspondences have been also referred to as Gauge/Gravity duality, as well as Maldacena duality, a duality which was first proposed by Maldacena.[111] Without going into technical details, the aim of this work is to show, for example, the existence of an equivalence relation between a certain supersymmetric $\mathrm{SU}(N)$ Yang-Mills field theory in 4 dimensional Minkowski spacetime, and a superstring theory in a 5 dimensional AdS space, having one additional dimension to the Minkowski one, and with the 5 dimensions of the AdS space supplemented by 5 extra dimensions defined by a five-sphere, making up the 10 dimensions of superstrings mentioned earlier. The interest in this work is that it deals with a connection between string theory (involving gravity) and

35, as shown later in Chapter 3 of Vol. II. In 4 dimensions, their degrees of freedom are, of course, two.

[108] See, e.g., Amati et al. [3, 4] and 't Hooft [219].

[109] See, e.g., Strominger and Fava [208] and Horowitz et al. [123].

[110] AdS space and CFT symmetry may be introduced as follows. AdS space, in D dimensions, may be defined in terms of coordinates $z=(z^0,z^1,\dots,z^{D-1},z^D)$ satisfying a quadratic equation $\sum_{k=1}^{D-1}(z^k)^2-(z^0)^2-(z^D)^2=-R^2$, for a given constant R^2, embedded in a $(D+1)$ dimensional space with interval squared $\mathrm{d}s^2=\sum_{j=1}^{D-1}\mathrm{d}z^{j\,2}-\mathrm{d}z^{0\,2}-\mathrm{d}z^{D\,2}$. On the other hand a $D-$Sphere is defined in terms of coordinates $y^1,\dots,y^{D+1}$ satisfying a quadratic equation $\sum_{j=1}^{D+1}(y^j)^2=\rho^2$ for a given constant ρ. The conformal group, as applied in 4 dimensional Minkowski spacetime, is defined by a scale transformation $x^\mu\rightarrow\lambda\,x^\mu$, and a so-called special (conformal) transformation

$$\frac{x'^\mu}{x'^2}=\frac{x^\mu}{x^2}+a^\mu,$$

for a constant 4-vector a^μ, in addition to the Poincaré ones.

[111] Maldacena [146]. See also Witten [245], Gubser et al. [109], and Aharoni et al. [2].

supersymmetric gauge theories. This brings us into contact with the holographic principle, in analogy to holography in capturing 3 dimensional images of objects on a 2 dimensional (holographic) plate,[112] showing that an equivalence relation exists between the 3 and the 2 dimensional set-ups. The 4 dimensional quantum field theory is like a hologram capturing information about the higher dimensional quantum gravity theory. In this case the SU(N) theory provides a holographic description of gravitational field. This is in analogy to black hole entropy with its encoded information being proportional to the area rather than to the volume of the region enclosed by the horizon. Perhaps holography is a basic property of string theory and one expects that much has to be done before developing a realistic quantum gravity, and in turn provide a background independent formulation for string theory. The holographic principle was first proposed by 't Hooft.[113]

We close this chapter by commenting on two symmetries which seem to be observed in Nature, that is of the CPT symmetry and of the Spin & Statistics connection and of their relevance to our own existence. We will see how these symmetries arise from quantum field theory in Sect. 4.10 and Sect. 4.5, respectively.

CPT taken in any order, seems to be an observed symmetry in Nature, where C stands for charge conjugation with which particles are replaced by their antiparticles and vice versa, P represents space reflection, while T denotes time reversal.

Local Lorentz invariant quantum field theories preserve (Sect. 4.10) the CPT symmetry. Experimentally, symmetry violations are well known to occur when one restricts to one or to the product of two transformations in CPT in dealing with some fundamental processes. For example the violation of parity was already established in 1957[114] as well as the violation of charge symmetry.[115] Later, in 1964 CP violation, and hence also of T, was observed in neutral Kaon decays.[116] The CP transformation and C, provide the fundamental relations between matter and anti-matter. The question then arises as to why we observe, apart in accelerator experiments, only one form (matter) than the other form in the visible universe – a key criterion for our own *existence*. If an equal amount of matter and anti-matter was produced at some stage then why, our visible universe is matter dominated. Sakharov in 1967 [177] proposed that a key reason for this is CP violation. In more details to explain this asymmetry, he proposed that (1) baryon number is not conserved. (This is supported by recent grand unified field theories,) (2) CP and C are violated, (3) the universe has gone through a phase of extremely rapid expansion to avoid the pairing of matter and anti-matter. The violation of such symmetries, at

[112]Recall that the *two* dimensional holographic plate which registers the interference of reflected light off an object and an unperturbed Laser beam stores information of the shape of the *three* dimensional object. As one shines a Laser beam on it an image of the three dimensional object emerges.

[113]'t Hooft [220], see also especially Thorn [223], as well as the analysis with further interpretations by Susskind [212]. See also Bousso [27].

[114]Wu et al. [246], Garwin et al. [91], and Friedman and Telegdi [86].

[115]Garwin et al. [91].

[116]Christenson et al. [32].

the microscopic level, and their consequences on a macroscopic scale is certainly intriguing.

Clearly, the "Spin & Statistics" connection, of which the Pauli exclusion principle is a special case applicable to spin $1/2$ particles, is important not only in physics but in all of the sciences, and is relevant to our own existence. For one thing, the periodic table of elements in chemistry is based on the exclusion principle. In simplest terms, the upshot of this is that half-odd-integer spin fields are quantized by anti-commutators, while integer spins fields are quantized by commutators. This result is of utmost significance for our existence. As a matter of fact the Pauli exclusion principle is not only sufficient for the stability of matter[117] in our world but it is also necessary.[118] In the problem of stability of neutral matter, with a finite number of electrons per atom, but involving several nuclei, and correspondingly a large number of electrons N, the stability of matter, based on the Pauli exclusion principle, or instability of so-called "bosonic matter", in which the exclusion principle is abolished, rests rather on the following. For "bosonic matter", the ground state energy $E_N \sim -N^{\alpha}$, with $\alpha > 1$,[119] where $(N + N)$ denotes the number of the negatively charged particles plus an equal number of positively charged particles. This behavior for "bosonic matter" is unlike that of matter, with the exclusion principle, for which $\alpha = 1$.[120] A power law behavior with $\alpha > 1$ implies instability as the formation of a single system consisting of $(2N + 2N)$ particles is favored over two separate systems brought together each consisting of $(N + N)$ particles, and the energy released upon collapse of the two systems into one, being proportional to $[(2N)^{\alpha} - 2(N)^{\alpha}]$, will be overwhelmingly large for realistic large N, e.g., $N \sim 10^{23}$. Dyson [61], has estimated that without the exclusion principle, the assembly of two macroscopic objects would release energy comparable to that of an atomic bomb, and such "matter" in bulk would collapse into a condensed high-density phase and our world will cease to exist.[121] Ordinary matter, due to the exclusion principle, occupies a very large volume. This point was emphasized by Ehrenfest in a discussion with Pauli in 1931[122] on the occasion of the Lorentz medal to this effect: "We take a piece of metal, or a stone. When we think about it, we are astonished that this quantity of matter should occupy so large a volume". He went on by stating that the exclusion principle is the reason: "Answer: only the Pauli principle, no two electrons in the same state". In this regard,

---

[117]For a pedagogical treatment of the problem of "stability of matter" and related intricacies, see Manoukian [151], Chapter 14.

[118]Lieb and Thirring [144] and Thirring [222].

[119]Dyson [61], Lieb [143], and Manoukian and Muthaporn [156].

[120]Lieb and Thirring [144] and Thirring [222].

[121]In the Preface of Tomonaga's book on spin [226], one reads: "The existence of spin, and the statistics associated with it, is the most subtle and ingenious design of Nature – without it the whole universe would collapse".

[122]See Ehrenfest [62].

a rigorous treatment[123] shows that the extension of matter radially grows not any slower than $N^{1/3}$ for large N. No wonder why matter occupies so large a volume. The importance of the "Spin & Statistics" connection and the role it plays in our world cannot be overemphasized. Needless to say, no quantum field theory treatment is complete without the CPT Theorem and the Spin & Statistics Connection.

The present volume deals with the foundations of quantum field theory and with the intricacies of abelian and non-abelian gauge theories. Volume II deals with quantum gravity, supersymmetry, and string theory.

References

1. Aad, G. et al. (2012). Observation of a new particle in the search for the Standard Model Higgs Boson with the ATLAS detector at the LHC. *Physics Letters, B716*, 1–29.
2. Aharoni, O. et al. (2008). $N = 6$ superconformal Chern-Simons matter theories, M2-branes and their gravity duals. *JHEP, 0810*, 091.
3. Amati, D., Ciafaloni, M., & Veneziano, G. (1987). Superstring collisions at planckian energies. *Physics Letters, B197*, 81–88.
4. Amati, D., Ciafaloni, M., & Veneziano, G. (1988). Classical and Quantum effects from Planckian energy superstring collisions. *International Journal of Modern Physics, 3*, 1615–1661.
5. Anderson, C. D. (1932). The apparent existence of easily deflectable positives. *Science, 76*, 238–239.
6. Anderson, C. D. (1933). The positive electron. *Physical Review, 43*, 491–494.
7. Anderson, P. W. (1963). Plasmons, gauge invariance, and mass. *Physical Review, 130*, 439–442.
8. Ansari, M. H. (2008). Generic degeneracy and entropy in loop quantum gravity. *Nuclear Physics, B795*, 635–644.
9. Anselmi, D. (2003). Absence of higher derivatives in the renormalization of propagators in quantum field theory with infinitely many couplings. *Class. Quantum Grav., 20*, 2344–2378.
10. Arnowitt, R. S., Deser, S., & Misner, W. (2008). The dynamics of general relativity. *General Relativity and Gravitation, 40*, 1997–2027. Reprinted from *Gravitation: An Introduction to current research* (Chap.7), Edited by L. Witten. John Wiley & Sons Inc., New York/London, 1962.
11. Aschenbrenner, M. (1996). A decoupling theorem for the BPHZL-scheme. *Annals of Physics, 250*, 320–351.
12. Ashtekar, A., & Lewandoski, J. (1997). Quantum theory of gravity I: Area operators. *Classical Quantum Gravity, 14*, A55–A81.
13. Becchi, C., Rouet, A., & Stora, R. (1976). Renormalization of gauge theories. *Annals of Physics, 98*, 287–321.
14. Bekenstein, J. D. (1973). Black holes and entropy. *Physical Review, D7*, 2333–2346.
15. Benvenuti, A. et al. (1974). Observation of Muonless neutrino-induced inelastic interactions. *Physical Review Letters, 32*, 800–803.
16. Beringer, J. et al. (2012). Particle data group. *Physical Review D, 86*, 010001.
17. Bethe, H. (1947). The Electromagnetic shift of energy levels. *Physical Review, 72*, 339–341.
18. Bethe, H., & Fermi, E. (1932). Über die Wechselwirkung von Zwei Elektronen. *Zeitschrift fur Physik, 77*, 296–306.

[123]Manoukian and Sirininlakul [157].