Electron-positron pairs in physics and astrophysics: from heavy nuclei to black holes

Based on Physics Reports 487 (2010) 1-140
1 Abstract

Due to the interaction of physics and astrophysics we are witnessing in these years a splendid synthesis of theoretical, experimental and observational results originating from three fundamental physical processes. They were originally proposed by Dirac, by Breit and Wheeler and by Sauter, Heisenberg, Euler and Schwinger. For almost seventy years they have all three been followed by a continued effort of experimental verification on Earth-based experiments. The Dirac process, $e^+e^- \rightarrow 2\gamma$, has been by far the most successful. It has obtained extremely accurate experimental verification and has led as well to an enormous number of new physics in possibly one of the most fruitful experimental avenues by introduction of storage rings in Frascati and followed by the largest accelerators worldwide: DESY, SLAC etc. The Breit–Wheeler process, $2\gamma \rightarrow e^+e^-$, although conceptually simple, being the inverse process of the Dirac one, has been by far one of the most difficult to be verified experimentally. Only recently, through the technology based on free electron X-ray laser and its numerous applications in Earth-based experiments, some first indications of its possible verification have been reached. The vacuum polarization process in strong electromagnetic field, pioneered by Sauter, Heisenberg, Euler and Schwinger, introduced the concept of critical electric field $E_c = m_\gamma^2 c^3/(\epsilon_0 \hbar)$. It has been searched without success for more than forty years by heavy-ion collisions in many of the leading particle accelerators worldwide.

The novel situation today is that these same processes can be studied on a much more grandiose scale during the gravitational collapse leading to the formation of a black hole being observed in Gamma Ray Bursts (GRBs). This report is dedicated to the scientific race. The theoretical and experimental work developed in Earth-based laboratories is confronted with the theoretical interpretation of space-based observations of phenomena originating on cosmological scales. What has become clear in the last ten years is that all the three above mentioned processes, duly extended in the general relativistic framework, are necessary for the understanding of the physics of the gravitational collapse to a black hole. Vice versa, the natural arena where these processes can be observed in mutual interaction and on an unprecedented scale, is indeed the realm of relativistic astrophysics.

We systematically analyze the conceptual developments which have followed the basic work of Dirac and Breit–Wheeler. We also recall how the seminal work of Born and Infeld inspired the work by Sauter, Heisenberg and Euler on effective Lagrangian leading to the estimate of the rate for the
process of electron–positron production in a constant electric field. In addition of reviewing the intuitive semi-classical treatment of quantum mechanical tunneling for describing the process of electron–positron production, we recall the calculations in Quantum Electrodynamics of the Schwinger rate and effective Lagrangian for constant electromagnetic fields. We also review the electron–positron production in both time-alternating electromagnetic fields, studied by Brezin, Itzykson, Popov, Nikishov and Narozhny, and the corresponding processes relevant for pair production at the focus of coherent laser beams as well as electron beam–laser collision. We finally report some current developments based on the general JWKB approach which allows to compute the Schwinger rate in spatially varying and time varying electromagnetic fields.

We also recall the pioneering work of Landau and Lifshitz, and Racah on the collision of charged particles as well as experimental success of AdA and ADONE in the production of electron–positron pairs.

We then turn to the possible experimental verification of these phenomena. We review: (A) the experimental verification of the $e^+e^- \rightarrow 2\gamma$ process studied by Dirac. We also briefly recall the very successful experiments of $e^+e^-$ annihilation to hadronic channels, in addition to the Dirac electromagnetic channel; (B) ongoing Earth based experiments to detect electron–positron production in strong fields by focusing coherent laser beams and by electron beam–laser collisions; and (C) the multiyear attempts to detect electron–positron production in Coulomb fields for a large atomic number $Z > 137$ in heavy ion collisions. These attempts follow the classical theoretical work of Popov and Zeldovich, and Greiner and their schools.

We then turn to astrophysics. We first review the basic work on the energetics and electrodynamical properties of an electromagnetic black hole and the application of the Schwinger formula around Kerr–Newman black holes as pioneered by Damour and Ruffini. We only focus on black hole masses larger than the critical mass of neutron stars, for convenience assumed to coincide with the Rhoades and Ruffini upper limit of $3.2\ M_\odot$. In this case the electron Compton wavelength is much smaller than the spacetime curvature and all previous results invariantly expressed can be applied following well established rules of the equivalence principle. We derive the corresponding rate of electron–positron pair production and introduce the concept of dyadosphere. We review recent progress in describing the evolution of optically thick electron–positron plasma in presence of supercritical electric field, which is relevant both in astrophysics as well as ongoing laser beam experiments. In particular we review recent progress based on the Vlasov–Boltzmann–Maxwell equations to study the feedback of the created electron–positron pairs on the original constant electric field. We evidence the existence of plasma oscillations and its interaction with photons leading to energy and number equipartition of photons, electrons and positrons. We finally review the recent progress obtained by using the Boltzmann equations to study
the evolution of an electron–positron-photon plasma towards thermal equi-
librium and determination of its characteristic timescales. The crucial differ-
ence introduced by the correct evaluation of the role of two and three body
 collisions, direct and inverse, is especially evidenced. We then present some
general conclusions.

The results reviewed in this report are going to be submitted to decisive
tests in the forthcoming years both in physics and astrophysics. To mention
only a few of the fundamental steps in testing in physics we recall the start-
ing of experimental facilities at the National Ignition Facility at the Lawrence
Livermore National Laboratory as well as corresponding French Laser the
Mega Joule project. In astrophysics these results will be tested in galactic
and extragalactic black holes observed in binary X-ray sources, active galac-
tic nuclei, microquasars and in the process of gravitational collapse to a neu-
tron star and also of two neutron stars to a black hole giving origin to GRBs.
The astrophysical description of the stellar precursors and the initial physical
conditions leading to a gravitational collapse process will be the subject of a
forthcoming report. As of today no theoretical description has yet been found
to explain either the emission of the remnant for supernova or the formation
of a charged black hole for GRBs. Important current progress toward the un-
derstanding of such phenomena as well as of the electrodynamical structure
of neutron stars, the supernova explosion and the theories of GRBs will be
discussed in the above mentioned forthcoming report. What is important
to recall at this stage is only that both the supernovae and GRBs processes
are among the most energetic and transient phenomena ever observed in the
Universe: a supernova can reach energy of $\sim 10^{54}$ ergs on a time scale of a few
months and GRBs can have emission of up to $\sim 10^{54}$ ergs in a time scale as
short as of a few seconds. The central role of neutron stars in the description
of supernovae, as well as of black holes and the electron–positron plasma,
in the description of GRBs, pioneered by one of us (RR) in 1975, are widely
recognized. Only the theoretical basis to address these topics are discussed
in the present report.
5.7.3 Effective nonlinear Lagrangian for arbitrary constant electromagnetic field .................................................. 983
5.8 Theory of pair production in an alternating field .................. 988
5.9 Nonlinear Compton scattering and Breit-Wheeler process .... 992
5.10 Quantum description of nonlinear Compton effect ............ 994

6 Semi-classical description of pair production in a general electric field ............................................................... 997
6.1 Semi-classical description of pair production ...................... 998
  6.1.1 JWKB transmission probability for Klein–Gordon Field 999
  6.1.2 Rate of pair production ........................................ 1001
  6.1.3 Including a smoothly varying $B(z)$-field parallel to $E(z)$ 1005
6.2 Time-dependent electric fields ...................................... 1006
6.3 Applications ................................................................ 1007
  6.3.1 Step-like electric field .......................................... 1007
  6.3.2 Sauter electric field ............................................. 1008
  6.3.3 Constant electric field in half space ......................... 1012
6.4 Tunneling into bound States ......................................... 1015
  6.4.1 JWKB transmission probability ............................... 1016
  6.4.2 Energy spectrum of bound states ............................ 1017
  6.4.3 Rate of pair production ........................................ 1017

7 Phenomenology of electron–positron pair creation and annihilation ............................................................. 1019
7.1 $e^+e^-$ annihilation experiments in particle physics ............ 1019
7.2 The Breit–Wheeler process in laser physics ...................... 1021
  7.2.1 Phenomenology of pair production in alternating fields 1023
  7.2.2 Pair production in X-ray free electron lasers ............... 1025
  7.2.3 Pair production by a circularly polarized laser beam .... 1025
  7.2.4 Availability of laser technology for pair production .... 1027
7.3 Phenomenology of pair production in electron beam-laser collisions .......................................................... 1029
  7.3.1 Experiment of electron beam-laser collisions ............ 1029
  7.3.2 Pair production viewed in the rest frame of electron beam1032
7.4 The Breit–Wheeler cutoff in high-energy $\gamma$-rays ............ 1032
7.5 Theory of pair production in Coulomb potential ............... 1036
  7.5.1 The $Z = 137$ catastrophe .................................... 1036
  7.5.2 Semi-Classic description ...................................... 1037
  7.5.3 The critical value of the nuclear charge $Z_{cr} = 173$ ... 1040
  7.5.4 Positron production .......................................... 1042
  7.5.5 Homogeneous and adiabatic approximation ............ 1044
7.6 Pair production in heavy-ion collisions ......................... 1046
  7.6.1 A transient super heavy “quasimolecules” .............. 1046
  7.6.2 Experiments ................................................... 1046
## Contents

8 **The extraction of blackholic energy from a black hole by vacuum polarization processes** 1051

8.1 Test particles in Kerr-Newman geometries 1052

8.2 Reversible and irreversible transformations of a black hole: the Christodoulou-Ruffini mass formula 1054

8.3 Positive and negative root states as limits of quantum field states 1058

8.4 Vacuum polarization in Kerr-Newman geometries 1059

8.5 The “Dyadosphere” in Reissner-Nordström geometry 1062

8.6 The “dyadotorus” 1067

8.7 Geometry of gravitationally collapsing cores

8.7.1 The Tolman-Oppenheimer-Snyder solution 1071

8.7.2 Gravitational collapse of charged and uncharged shells 1073

8.8 The maximum energy extractable from a black hole 1081

8.8.1 The formula of the irreducible mass of a black hole 1081

8.8.2 Extracting electromagnetic energy from a subcritical and overcritical black hole 1083

8.9 A theorem on a possible disagreement between black holes and thermodynamics 1085

8.10 Astrophysical gravitational collapse and black holes 1090

9 **Plasma oscillations in electric fields** 1093

9.1 Semiclassical theory of plasma oscillations in electric fields 1093

9.2 Kinetic theory of plasma oscillations in electric fields 1097

9.3 Plasma oscillations in the color electric field of heavy ions 1098

9.4 Quantum Vlasov equation 1100

9.5 Quantum decoherence in plasma oscillations 1102

9.6 Collision decoherence in plasma oscillations 1104

9.7 $e^+e^-\gamma$ interactions in plasma oscillations in electric fields 1106

9.8 Electro-fluidodynamics of the pair plasma 1116

10 **Thermalization of the mildly relativistic pair plasma** 1131

10.1 Qualitative description of the pair plasma 1133

10.2 The discretization procedure and the computational scheme 1139

10.3 Conservation laws 1140

10.4 Determination of temperature and chemical potentials in kinetic equilibrium 1141

10.5 Binary interactions

10.5.1 Compton scattering 1143

10.5.2 Pair creation and annihilation 1145

10.5.3 Møller scattering of electrons and positrons 1147

10.5.4 Bhaba scattering of electrons on positrons 1149

10.6 Three-body processes 1150

10.7 Cutoff in the Coulomb scattering 1152

10.8 Numerical results 1153
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 Concluding remarks</td>
<td>1157</td>
</tr>
<tr>
<td>12 Notes added in proof</td>
<td>1163</td>
</tr>
<tr>
<td>Bibliography</td>
<td>1167</td>
</tr>
</tbody>
</table>
2 Introduction

The annihilation of electron–positron pair into two photons, and its inverse process – the production of electron–positron pair by the collision of two photons were first studied in the framework of quantum mechanics by P.A.M. Dirac and by G. Breit and J.A. Wheeler in the 1930s [1, 2].

A third fundamental process was pioneered by the work of Fritz Sauter and Oscar Klein, pointing to the possibility of creating an electron–positron pair from the vacuum in a constant electromagnetic field. This became known as the ‘Klein paradox’ and such a process named as vacuum polarization. It would occur for an electric field stronger than the critical value

\[ E_c \equiv \frac{m^2 e^3}{\hbar c} \approx 1.3 \cdot 10^{16} \text{ V/cm.} \]  

(2.0.1)

where \( m_e, e, c \) and \( \hbar \) are respectively the electron mass and charge, the speed of light and the Planck’s constant.

The experimental difficulties to verify the existence of such three processes became immediately clear. While the process studied by Dirac was almost immediately observed [3] and the electron–positron collisions became possibly the best tested and prolific phenomenon ever observed in physics. The Breit–Wheeler process, on the contrary, is still today waiting a direct observational verification. Similarly the vacuum polarization process defied dedicated attempts for almost fifty years in experiments in nuclear physics laboratories and accelerators all over the world, see Section 7.

From the theoretical point of view the conceptual changes implied by these processes became immediately clear. They were by vastness and depth only comparable to the modifications of the linear gravitational theory of Newton introduced by the nonlinear general relativistic equations of Einstein. In the work of Euler, Oppenheimer and Debye, Born and his school it became clear that the existence of the Breit–Wheeler process was conceptually modifying the linearity of the Maxwell theory. In fact the creation of the electron–positron pair out of the two photons modifies the concept of superposition of the linear electromagnetic Maxwell equations and impose the necessity to transit to a nonlinear theory of electrodynamics. In a certain sense the Breit–Wheeler process was having for electrodynamics the same fundamental role of Gedankenexperiment that the equivalence principle had for gravitation. Two different attempts to study these nonlinearities in the electrodynamics were made: one by Born and Infeld [4–6] and one by Euler and
Heisenberg [7]. These works prepared the even greater revolution of Quantum Electro-Dynamics by Tomonaga [8], Feynman [9–11], Schwinger [12–14] and Dyson [15, 16].

In Section 3 we review the fundamental contributions to the electron–positron pair creation and annihilation and to the concept of the critical electric field. In Section 3.1 we review the Dirac derivation [1] of the electron–positron annihilation process obtained within the perturbation theory in the framework of relativistic quantum mechanics and his derivation of the classical formula for the cross-section \( \sigma_{e^+e^-}^{\text{lab}} \) in the rest frame of the electron

\[
\sigma_{e^+e^-}^{\text{lab}} = \pi \left( \frac{\alpha \hbar}{m_e c} \right)^2 (\gamma - 1)^{-1} \left\{ \frac{\hat{\gamma}^2 + 4\hat{\gamma} + 1}{\hat{\gamma}^2 - 1} \ln[\hat{\gamma} + (\hat{\gamma}^2 - 1)^{1/2}] - \frac{\hat{\gamma} + 3}{(\hat{\gamma}^2 - 1)^{1/2}} \right\},
\]

where \( \hat{\gamma} \equiv E_+/m_e c^2 \geq 1 \) is the energy of the positron and \( \alpha = e^2/(\hbar c) \) is as usual the fine structure constant, and we recall the corresponding formula for the center of mass reference frame. In Section 3.2 we recall the main steps in the classical Breit–Wheeler work [2] on the production of a real electron–positron pair in the collision of two photons, following the same method used by Dirac and leading to the evaluation of the total cross-section \( \sigma_{\gamma\gamma} \) in the center of mass of the system

\[
\sigma_{\gamma\gamma} = \frac{\pi}{2} \left( \frac{\alpha \hbar}{m_e c} \right)^2 (1 - \beta^2) \left[ 2\hat{\beta}(\hat{\beta}^2 - 2) + (3 - \hat{\beta}^4) \ln \left( \frac{1 + \hat{\beta}}{1 - \hat{\beta}} \right) \right], \quad \text{with} \quad \hat{\beta} = \frac{|p|}{E},
\]

where \( \hat{\beta} \) is the reduced velocity of the electron or the positron. In Section 3.3 we recall the basic higher order processes, compared to the Dirac and Breit–Wheeler ones, leading to pair creation. In Section 3.4 we recall the famous Klein paradox [17, 18] and the possible tunneling between the positive and negative energy states leading to the concept of level crossing and pair creation by analogy to the Gamow tunneling [19] in the nuclear potential barrier. We then turn to the celebrated Sauter work [20] showing the possibility of creating a pair in a uniform electric field \( E \). We recover in Section 3.5.1 a JWKB approximation in order to reproduce and improve on the Sauter result by obtaining the classical Sauter exponential term as well as the prefactor

\[
\frac{\Gamma_{\text{JWKB}}}{V} \simeq D_s \frac{\alpha E^2}{2\pi^2 \hbar} e^{-\pi E_c/E},
\]

where \( D_s = 2 \) for a spin-1/2 particle and \( D_s = 1 \) for spin-0, \( V \) is the volume. Finally, in Section 3.5.2 the case of a simultaneous presence of an electric and a magnetic field \( B \) is presented leading to the estimate of pair production rate

\[
\frac{\Gamma_{\text{JWKB}}}{V} \simeq \frac{\alpha \beta \epsilon}{\pi \hbar} \coth \left( \frac{\pi \beta}{\epsilon} \right) \exp \left( -\frac{\pi E_c}{\epsilon} \right), \quad \text{spin} - 1/2 \text{ particle}
\]
and
\[
\frac{\Gamma_{\text{JWKB}}}{V} \approx \frac{a \beta e}{2\pi \hbar} \sinh^{-1} \left( \frac{\pi \beta}{e} \right) \exp \left( -\frac{\pi E}{e} \right), \quad \text{spin} - 0 \text{ particle,}
\]
where
\[
\epsilon \equiv \sqrt{(S^2 + P^2)^{1/2} + S},
\]
\[
\beta \equiv \sqrt{(S^2 + P^2)^{1/2} - S},
\]
where the scalar $S$ and the pseudoscalar $P$ are
\[
S \equiv \frac{1}{4} F_{\mu\nu} F^{\mu\nu} = \frac{1}{2} (E^2 - B^2); \quad P \equiv \frac{1}{4} F_{\mu\nu} \tilde{F}^{\mu\nu} = E \cdot B,
\]
where $\tilde{F}^{\mu\nu} \equiv \epsilon^{\mu\nu\lambda\kappa} F_{\lambda\kappa}$ is the dual field tensor.

In Section 4 we first recall the seminal work of Hans Euler [21] pointing out for the first time the necessity of nonlinear character of electromagnetism introducing the classical Euler Lagrangian
\[
\mathcal{L} = \frac{E^2 - B^2}{8\pi} + \frac{1}{\alpha E_0^2} \left[ a_E \left( E^2 - B^2 \right)^2 + b_E \left( E \cdot B \right)^2 \right],
\]
where
\[
a_E = -1/(360\pi^2), \quad b_E = -7/(360\pi^2),
\]
a first order perturbation to the Maxwell Lagrangian. In Section 4.2 we review the alternative theoretical approach of nonlinear electrodynamics by Max Born [5] and his collaborators, to the more ambitious attempt to obtain the correct nonlinear Lagrangian of Electro-Dynamics. The motivation of Born was to attempt a theory free of divergences in the observable properties of an elementary particle, what has become known as ‘unitarian’ standpoint versus the ‘dualistic’ standpoint in description of elementary particles and fields. We recall how the Born Lagrangian was formulated
\[
\mathcal{L} = \sqrt{1 + 2S - P^2} - 1,
\]
and one of the first solutions derived by Born and Infeld [6]. We also recall one of the interesting aspects of the courageous approach of Born had been to formulate this Lagrangian within a unified theory of gravitation and electromagnetism following Einstein program. Indeed, we also recall the very interesting solution within the Born theory obtained by Hoffmann [22, 23]. Still in the work of Born [5] the seminal idea of describing the nonlinear vacuum properties of this novel electrodynamics by an effective dielectric constant and magnetic permeability functions of the field arisen. We then review
in Section 4.3.1 the work of Heisenberg and Euler [7] adopting the general approach of Born and generalizing to the presence of a real and imaginary part of the electric permittivity and magnetic permeability. They obtain an integral expression of the effective Lagrangian given by

\[ \Delta L_{\text{eff}} = \frac{e^2}{16\pi^2\hbar c} \int_0^\infty e^{-s} \frac{ds}{s^3} \left[ is^2 E\bar{B} \cos(s[E^2 - \bar{B}^2 + 2i(\bar{E}\bar{B})]^{1/2}) + \text{c.c.} \right] + \left( \frac{m^2 c^3 e}{\hbar c} \right)^2 + \frac{s^2}{3} (|\bar{B}|^2 - |\bar{E}|^2), \]

where \( \bar{E}, \bar{B} \) are the dimensionless reduced fields in the unit of the critical field \( E_c \),

\[ \bar{E} = \frac{|E|}{E_c}, \quad \bar{B} = \frac{|B|}{E_c}, \]

obtaining the real part and the crucial imaginary term which relates to the pair production in a given electric field. It is shown how these results give as a special case the previous result obtained by Euler (4.1.3). In Section 4.3.2 the work by Weisskopf [24] working on a spin-0 field fulfilling the Klein–Gordon equation, in contrast to the spin 1/2 field studied by Heisenberg and Euler, confirms the Euler-Heisenberg result. Weisskopf obtains explicit expression of pair creation in an arbitrary strong magnetic field and in an electric field described by \( \bar{E} \) and \( \bar{B} \) expansion.

For the first time Heisenberg and Euler provided a description of the vacuum properties by the characteristic scale of strong field \( E_c \) and the effective Lagrangian of nonlinear electromagnetic fields. In 1951, Schwinger [25–27] made an elegant quantum field theoretic reformulation of this discovery in the QED framework. This played an important role in understanding the properties of the QED theory in strong electromagnetic fields. The QED theory in strong coupling regime, i.e., in the regime of strong electromagnetic fields, is still a vast arena awaiting for experimental verification as well as of further theoretical understanding.

In Section 5 after recalling some general properties of QED in Section 5.1 and some basic processes in Section 5.2 we proceed to the consideration of the Dirac and the Breit–Wheeler processes in QED in Secton 5.3. Then we discuss some higher order processes, namely double pair production in Section 5.4, electron-nucleus bremsstrahlung and pair production by a photon in the field of a nucleus in Section 5.5, and finally pair production by two ions in Section 5.6. In Section 5.7 the classical result for the vacuum to vacuum decay via pair creation in uniform electric field by Schwinger is recalled

\[ \frac{\Gamma}{V} = \frac{\alpha E^2}{\pi^2} \sum_{n=1}^\infty \frac{1}{n^2} \exp \left( -\frac{n\pi E_c}{E} \right). \]
This formula generalizes and encompasses the previous results reviewed in our report: the JWKB results, discussed in Section 3.5, and the Sauter exponential factor (3.5.11), and the Heisenberg-Euler imaginary part of the effective Lagrangian. We then recall the generalization of this formula to the case of a constant electromagnetic fields. Such results were further generalized to spatially nonuniform and time-dependent electromagnetic fields by Nikishov [28], Vanyashin and Terent’ev [29], Popov [30–32], Narozhny and Nikishov [33] and Batalin and Fradkin [34]. We then conclude this argument by giving the real and imaginary parts for the effective Lagrangian for arbitrary constant electromagnetic field recently published by Ruffini and Xue [35]. This result generalizes the previous result obtained by Weisskopf in strong fields. In weak field it gives the Euler-Heisenberg effective Lagrangian. As we will see in the Section 7.2 much attention has been given experimentally to the creation of pairs in the rapidly changing electric fields. A fundamental contribution in this field studying pair production rates in an oscillating electric field was given by Brezin and Itzykson [36] and we recover in Section 5.8 their main results which apply both to the case of bosons and fermions. We recall how similar results were independently obtained two years later by Popov [37]. In Section 5.10 we recall an alternative physical process considering the quantum theory of the interaction of free electron with the field of a strong electromagnetic waves: an ultrarelativistic electron absorbs multiple photons and emits only a single photon in the reaction [38]:

$$e + n\omega \rightarrow e' + \gamma.$$ 

This process appears to be of the great relevance as we will see in the next Section for the nonlinear effects originating from laser beam experiments. Particularly important appears to be the possibility outlined by Burke et al. [39] that the high-energy photon $\gamma$ created in the first process propagates through the laser field, it interacts with laser photons $n\omega$ to produce an electron–positron pair

$$\gamma + n\omega \rightarrow e^+ + e^-.$$ 

We also refer to the papers by Narozhny and Popov [40–45] studying the dependence of this process on the status of the polarization of the photons.

We point out the great relevance of departing from the case of the uniform electromagnetic field originally considered by Sauter, Heisenberg and Euler, and Schwinger. We also recall some of the classical works of Brezin and Itzykson and Popov on time varying fields. The space variation of the field was also considered in the classical papers of Nikishov and Narozhny as well as in the work of Wang and Wong. Finally, we recall the work of Khriplovich [46] studying the vacuum polarization around a Reissner–Nordström black hole. A more recent approach using the worldline formalism, sometimes called the string-inspired formalism, was advanced by Dunne and Schubert [47,48].
2 Introduction

In Section 6, after recalling studies of pair production in inhomogeneous electromagnetic fields in the literature by [48–53], we present a brief review of our recent work [54] where the general formulas for pair production rate as functions of either crossing energy level or classical turning point, and total production rate are obtained in external electromagnetic fields which vary either in one space direction $E(z)$ or in time $E(t)$. In Sections 6.1 and 6.2, these formulas are explicitly derived in the JWKB approximation and generalized to the case of three-dimensional electromagnetic configurations. We apply these formulas to several cases of such inhomogeneous electric field configurations, which are classified into two categories. In the first category, we study two cases: a semi-confined field $E(z) \neq 0$ for $z \lesssim \ell$ and the Sauter field

$$E(z) = E_0 / \cosh^2 (z / \ell), \quad V(z) = -\sigma_s m_e c^2 \tanh (z / \ell),$$

where $\ell$ is width in the $z$-direction, and

$$\sigma_s \equiv e E_0 \ell / m_e c^2 = (\ell / \lambda_C) (E_0 / E_c).$$

In these two cases the pairs produced are not confined by the electric potential and can reach an infinite distance. The resultant pair production rate varies as a function of space coordinate. The result we obtained is drastically different from the Schwinger rate in homogeneous electric fields without any boundary. We clearly show that the approximate application of the Schwinger rate to electric fields limited within finite size of space overestimates the total number of pairs produced, particularly when the finite size is comparable with the Compton wavelength $\lambda_C$, see Figs. 6.2 and 6.3 where it is clearly shown how the rate of pair creation far from being constant goes to zero at both boundaries. The same situation is also found for the case of the semi-confined field $z(z) \neq 0$ for $|z| \lesssim \ell$, see Eq. (6.3.34). In the second category, we study a linearly rising electric field $E(z) \sim z$, corresponding to a harmonic potential $V(z) \sim z^2$, see Figs. 6.1. In this case the energy spectra of bound states are discrete and thus energy crossing levels for tunneling are discrete. To obtain the total number of pairs created, using the general formulas for pair production rate, we need to sum over all discrete energy crossing levels, see Eq. (6.4.11), provided these energy levels are not occupied. Otherwise, the pair production would stop due to the Pauli principle.

In Section 7 we focus on the phenomenology of electron–positron pair creation and annihilation experiments. There are three different aspects which are examined: the verification of the process (3.0.1) initially studied by Dirac, the process (3.0.2) studied by Breit and Wheeler, and then the classical work of vacuum polarization process around a supercritical nucleus, following the Sauter, Euler, Heisenberg and Schwinger work. We first recall in Section 7.1 how the process (3.0.1) predicted by Dirac was almost immediately discov-
ered by Klemperer [3]. Following this discovery the electron–positron collisions have become possibly the most prolific field of research in the domain of particle physics. The crucial step experimentally was the creation of the first electron–positron collider the “Anello d’Accumulazione” (AdA) was built by the theoretical proposal of Bruno Touschek in Frascati (Rome) in 1960 [55]. Following the success of AdA (luminosity $\sim 10^{25}/(\text{cm}^2 \text{sec})$, beam energy $\sim 0.25\text{GeV}$), it was decided to build in the Frascati National Laboratory a storage ring of the same kind, Adone. Electron-positron colliders have been built and proposed for this purpose all over the world (CERN, SLAC, INP, DESY, KEK and IHEP). The aim here is just to recall the existence of this enormous field of research which appeared following the original Dirac idea. The main cross-sections (7.1.1) and (7.1.2) are recalled and the diagram (Fig. 7.1) summarizing this very great success of particle physics is presented. While the Dirac process (3.0.1) has been by far one of the most prolific in physics, the Breit–Wheeler process (3.0.2) has been one of the most elusive for direct observations. In Earth-bound experiments the major effort today is directed to evidence this phenomenon in very strong and coherent electromagnetic field in lasers. In this process collision of many photons may lead in the future to pair creation. This topic is discussed in Section 7.2. Alternative evidence for the Breit–Wheeler process can come from optically thick electron–positron plasma which may be created either in the future in Earth-bound experiments, or currently observed in astrophysics, see Section 10. One additional way to probe the existence of the Breit–Wheeler process is by establishing in astrophysics an upper limits to observable high-energy photons, as a function of distance, propagating in the Universe as pioneered by Nikishov [56], see Section 7.4. We then recall in Section 7.3 how the crucial experimental breakthrough came from the idea of John Madey [57] of self-amplified spontaneous emission in an undulator, which results when charges interact with the synchrotron radiation they emit [58]. Such X-ray free electron lasers have been constructed among others at DESY and SLAC and focus energy onto a small spot hopefully with the size of the X-ray laser wavelength $\lambda \approx O(0.1)\text{nm}$ [59], and obtain a very large electric field $E \sim 1/\lambda$, much larger than those obtainable with any optical laser of the same power. This technique can be used to achieve a very strong electric field near to its critical value for observable electron–positron pair production in vacuum. No pair can be created by a single laser beam. It is then assumed that each X-ray laser pulse is split into two equal parts and recombined to form a standing wave with a frequency $\omega$. We then recall how for a laser pulse with wavelength $\lambda$ about $1\mu\text{m}$ and the theoretical diffraction limit $\sigma_{\text{laser}} \approx \lambda$ being reached, the critical intensity laser beam would be

$$I_{\text{laser}}^c = \frac{c}{4\pi} E_{\text{c}}^2 \approx 4.6 \cdot 10^{29} \text{W/cm}^2.$$  

In Section 7.2.1 we recall the theoretical formula for the probability of pair
production in time-alternating electric field in two limiting cases of large frequency and small frequency. It is interesting that in the limit of large field and small frequency the production rate approach the one of the Sauter, Heisenberg, Euler and Schwinger, discussed in Section 5. In the following Section 7.2.2 we recall the actually reached experimental limits quoted by Ringwald [60] for a X-ray laser and give a reference to the relevant literature. In Section 7.2.3 we summarize some of the most recent theoretical estimates for pair production by a circularly polarized laser beam by Narozhny, Popov and their collaborators. In this case the field invariants (3.5.23) are not vanishing and pair creation can be achieved by a single laser beam. They computed the total number of electron–positron pairs produced as a function of intensity and focusing parameter of the laser. Particularly interesting is their analysis of the case of two counter-propagating focused laser pulses with circular polarizations, pair production becomes experimentally observable when the laser intensity $I_{\text{laser}} \sim 10^{26} \text{W/cm}^2$ for each beam, which is about 1 $\sim$ 2 orders of magnitude lower than for a single focused laser pulse, and more than 3 orders of magnitude lower than the critical intensity (7.2.4). Equally interesting are the considerations which first appear in treating this problem that the back reaction of the pairs created on the field has to be taken into due account. We give the essential references and we will see in Section 9 how indeed this feature becomes of paramount importance in the field of astrophysics. We finally review in Section 7.2.4 the technological situation attempting to increase both the frequency and the intensity of laser beams.

The difficulty of evidencing the Breit–Wheeler process even when the high-energy photon beams have a center of mass energy larger than the energy-threshold $2m_e c^2 = 1.02 \text{ MeV}$ was clearly recognized since the early days. We discuss the crucial role of the effective nonlinear terms originating in strong electromagnetic laser fields: the interaction needs not to be limited to initial states of two photons [61, 62]. A collective state of many interacting laser photons occurs. We turn then in Section 7.3 to an even more complex and interesting procedure: the interaction of an ultrarelativistic electron beam with a terawatt laser pulse, performed at SLAC [63], when strong electromagnetic fields are involved. A first nonlinear Compton scattering process occurs in which the ultrarelativistic electrons absorb multiple photons from the laser field and emit a single photon via the process (5.9.1). The theory of this process has been given in Section 5.10. The second is a drastically improved Breit–Wheeler process (5.9.2) by which the high-energy photon $\gamma$, created in the first process, propagates through the laser field and interacts with laser photons $n\omega$ to produce an electron–positron pair [39]. In Section 7.3.1 we describe the status of this very exciting experiments which give the first evidence for the observation in the laboratory of the Breit–Wheeler process although in a somewhat indirect form. Having determined the theoretical basis as well as attempts to verify experimentally the Breit–Wheeler formula we turn in Section 7.4 to a most important application of the Breit–Wheeler
process in the framework of cosmology. As pointed out by Nikishov [56] the existence of background photons in cosmology puts a stringent cutoff on the maximum trajectory of the high-energy photons in cosmology.

Having reviewed both the theoretical and observational evidence of the Dirac and Breit–Wheeler processes of creation and annihilation of electron–positron pairs we turn then to one of the most conspicuous field of theoretical and experimental physics dealing with the process of electron–positron pair creation by vacuum polarization in the field of a heavy nuclei. This topic has originated one of the vastest experimental and theoretical physics activities in the last forty years, especially by the process of collisions of heavy ions. We first review in Section 7.5 the $Z = 137$ catastrophe, a collapse to the center, in semi-classical approach, following the Pomeranchuk work [64] based on the imposing the quantum conditions on the classical treatment of the motion of two relativistic particles in circular orbits. We then proceed showing in Section 7.5.3 how the introduction of the finite size of the nucleus, following the classical work of Popov and Zeldovich [65], leads to the critical charge of a nucleus of $Z_{\text{cr}} = 173$ above which a bare nucleus would lead to the level crossing between the bound state and negative energy states of electrons in the field of a bare nucleus. We then review in Section 7.5.5 the recent theoretical progress in analyzing the pair creation process in a Coulomb field, taking into account radial dependence and time variability of electric field. We finally recall in Section 7.6 the attempt to use heavy-ion collisions to form transient superheavy “quasimolecules”: a long-lived metastable nuclear complex with $Z > Z_{\text{cr}}$. It was expected that the two heavy ions of charges respectively $Z_1$ and $Z_2$ with $Z_1 + Z_2 > Z_{\text{cr}}$ would reach small inter-nuclear distances well within the electron’s orbiting radii. The electrons would not distinguish between the two nuclear centers and they would evolve as if they were bounded by nuclear “quasimolecules” with nuclear charge $Z_1 + Z_2$. Therefore, it was expected that electrons would evolve quasi-statically through a series of well defined nuclear “quasimolecules” states in the two-center field of the nuclei as the inter-nuclear separation decreases and then increases again. When heavy-ion collision occurs the two nuclei come into contact and some deep inelastic reaction occurs determining the duration $\Delta t_s$ of this contact. Such “sticking time” is expected to depend on the nuclei involved in the reaction and on the beam energy. Theoretical attempts have been proposed to study the nuclear aspects of heavy-ion collisions at energies very close to the Coulomb barrier and search for conditions, which would serve as a trigger for prolonged nuclear reaction times, to enhance the amplitude of pair production. The sticking time $\Delta t_s$ should be larger than $1 \sim 2 \cdot 10^{-21}$ sec [66] in order to have significant pair production. Up to now no success has been achieved in justifying theoretically such a long sticking time. In reality the characteristic sticking time has been found of the order of $\Delta t \sim 10^{-23}$ sec, hundred times shorter than the needed to activate the pair creation process. We finally recall in Section 7.6.2 the Darmstadt-Brookhaven dialogue between the
Orange and the Epos groups and the Apex group at Argonne in which the claim for discovery of electron–positron pair creation by vacuum polarization in heavy-ion collisions was finally retracted. Out of the three fundamental processes addressed in this report, the Dirac electron–positron annihilation and the Breit–Wheeler electron–positron creation from two photons have found complete theoretical descriptions within Quantum Electro-Dynamics. The first one is very likely the best tested process in physical science, while the second has finally obtained the first indirect experimental evidence. The third process, the one of the vacuum polarization studied by Sauter, Euler, Heisenberg and Schwinger, presents in Earth-bound experiments presents a situation “terra incognita”.

We turn then to astrophysics, where, in the process of gravitational collapse to a black hole and in its outcomes these three processes will be for the first time verified on a much larger scale, involving particle numbers of the order of $10^{60}$, seeing both the Dirac process and the Breit–Wheeler process at work in symbiotic form and electron–positron plasma created from the “blackholic energy” during the process of gravitational collapse. It is becoming more and more clear that the gravitational collapse process to a Kerr–Newman black hole is possibly the most complex problem ever addressed in physics and astrophysics. What is most important for this report is that it gives for the first time the opportunity to see the above three processes simultaneously at work under ultrarelativistic special and general relativistic regimes. The process of gravitational collapse is characterized by the timescale $\Delta t_g = GM/c^3 \simeq 5 \cdot 10^{-6} M/M_\odot$ sec and the energy involved are of the order of $\Delta E = 10^{54} M/M_\odot$ ergs. It is clear that this is one of the most energetic and most transient phenomena in physics and astrophysics and needs for its correct description such a highly time varying treatment. Our approach in Section 8 is to gain understanding of this process by separating the different components and describing 1) the basic energetic process of an already formed black hole, 2) the vacuum polarization process of an already formed black hole, 3) the basic formula of the gravitational collapse recovering the Tolman-Oppenheimer-Snyder solutions and evolving to the gravitational collapse of charged and uncharged shells. This will allow among others to obtain a better understanding of the role of irreducible mass of the black hole and the maximum blackholic energy extractable from the gravitational collapse. We will as well address some conceptual issues between general relativity and thermodynamics which have been of interest to theoretical physicists in the last forty years. Of course in these brief chapter we will be only recalling some of these essential themes and refer to the literature where in-depth analysis can be found. In Section 8.1 we recall the Kerr–Newman metric and the associated electromagnetic field. We then recall the classical work of Carter [67] integrating the Hamilton-Jacobi equations for charged particle motions in the above given metric and electromagnetic field. We then recall
in Section 8.2 the introduction of the effective potential techniques in order to obtain explicit expression for the trajectory of a particle in a Kerr–Newman geometry, and especially the introduction of the reversible–irreversible transformations which lead then to the Christodoulou-Ruffini mass formula of the black hole

\[ M^2c^4 = \left( M_{\text{ir}}c^2 + \frac{c^2Q^2}{4GM_{\text{ir}}} \right)^2 + \frac{L^2c^8}{4G^2M_{\text{ir}}^2}, \]

where \( M_{\text{ir}} \) is the irreducible mass of a black hole, \( Q \) and \( L \) are its charge and angular momentum. We then recall in Section 8.3 the positive and negative root states of the Hamilton–Jacobi equations as well as their quantum limit. We finally introduce in Section 8.4 the vacuum polarization process in the Kerr–Newman geometry as derived by Damour and Ruffini [68] by using a spatially orthonormal tetrad which made the application of the Schwinger formalism in this general relativistic treatment almost straightforward. We then recall in Section 8.5 the definition of a dyadosphere in a Reissner–Nordström geometry, a region extending from the horizon radius

\[ r_+ = 1.47 \cdot 10^5 \mu (1 + \sqrt{1 - \xi^2}) \text{ cm} \]

out to an outer radius

\[ r^* = \left( \frac{\hbar}{m_c c} \right)^{1/2} \left( \frac{G M}{\epsilon^2} \right)^{1/2} \left( \frac{m_p}{m_e} \right)^{1/2} \left( \frac{e}{q_p} \right)^{1/2} \left( \frac{Q}{\sqrt{GM}} \right)^{1/2}, \]

\[ = 1.12 \cdot 10^8 \sqrt{\mu \xi} \text{ cm}, \]

where the dimensionless mass and charge parameters \( \mu = \frac{M}{M_{\odot}}, \xi = \frac{Q}{(M \sqrt{G})} \leq 1 \). In Section 8.6 the definition of a dyadotorus in a Kerr–Newman metric is recalled. We have focused on the theoretically well defined problem of pair creation in the electric field of an already formed black hole. Having set the background for the blackholic energy we recall some fundamental features of the dynamical process of the gravitational collapse. In Section 8.7 we address some specific issues on the dynamical formation of the black hole, recalling first the Oppenheimer-Snyder solution [69] and then considering its generalization to the charged nonrotating case using the classical work of W. Israel and V. de la Cruz [70, 71]. In Section 8.7.1 we recover the classical Tolman-Oppenheimer-Snyder solution in a more transparent way than it is usually done in the literature. In the Section 8.7.2 we are studying using the Israel-de la Cruz formalism the collapse of a charged shell to a black hole for selected cases of a charged shell collapsing on itself or collapsing in an already formed Reissner–Nordström black hole. Such elegant and powerful formalism has allowed to obtain for the first time all the analytic equations for such large variety of possibilities of the process of the gravitational collapse. The
theoretical analysis of the collapsing shell considered in the previous section allows to reach a deeper understanding of the mass formula of black holes at least in the case of a Reissner–Nordström black hole. This allows as well to give in Section 8.8 an expression of the irreducible mass of the black hole only in terms of its kinetic energy of the initial rest mass undergoing gravitational collapse and its gravitational energy and kinetic energy $T_+$ at the crossing of the black hole horizon $r_+$

$$M_{ir} = M_0 - \frac{M_0^2}{2r_+} + T_+. $$

Similarly strong, in view of their generality, are the considerations in Section 8.8.2 which indicate a sharp difference between the vacuum polarization process in an overcritical $E \gg E_c$ and undercritical $E \ll E_c$ black hole. For $E \gg E_c$ the electron–positron plasma created will be optically thick with average particle energy 10 MeV. For $E \ll E_c$ the process of the radiation will be optically thin and the characteristic energy will be of the order of $10^{21} \text{ eV}$. This argument will be further developed in a forthcoming report. In Section 8.9 we show how the expression of the irreducible mass obtained in the previous Section leads to a theorem establishing an upper limit to 50% of the total mass energy initially at rest at infinity which can be extracted from any process of gravitational collapse independent of the details. These results also lead to some general considerations which have been sometimes claimed in reconciling general relativity and thermodynamics.

The conditions encountered in the vacuum polarization process around black holes lead to a number of electron–positron pairs created of the order of $10^{60}$ confined in the dyadosphere volume, of the order of a few hundred times to the horizon of the black hole. Under these conditions the plasma is expected to be optically thick and is very different from the nuclear collisions and laser case where pairs are very few and therefore optically thin. We turn then in Section 9, to discuss a new phenomenon: the plasma oscillations, following the dynamical evolution of pair production in an external electric field close to the critical value. In particular, we will examine: (i) the back reaction of pair production on the external electric field; (ii) the screening effect of pairs on the electric field; (iii) the motion of pairs and their interactions with the created photon fields. In Secs. 9.1 and 9.2, we review semi-classical and kinetic theories describing the plasma oscillations using respectively the Dirac-Maxwell equations and the Boltzmann-Vlasov equations. The electron–positron pairs, after they are created, coherently oscillate back and forth giving origin to an oscillating electric field. The oscillations last for at least a few hundred Compton times. We review the damping due to the quantum decoherence. The energy from collective motion of the classical electric field and pairs flows to the quantum fluctuations of these fields. This process is quantitatively discussed by using the quantum Boltzmann-Vlasov
equation in Sections 9.4 and 9.5. The damping due to collision decoherence is quantitatively discussed in Sections 9.6 and 9.7 by using Boltzmann-Vlasov equation with particle collisions terms. This damping determines the energy flows from collective motion of the classical electric field and pairs to the kinetic energy of non-collective motion of particles of these fields due to collisions. In Section 9.7, we particularly address the study of the influence of the collision processes $e^+e^- \rightleftharpoons \gamma\gamma$ on the plasma oscillations in supercritical electric field [72]. It is shown that the plasma oscillation is mildly affected by a small number of photons creation in the early evolution during a few hundred Compton times (see Fig. 9.4). In the later evolution of $10^{3−4}$ Compton times, the oscillating electric field is damped to its critical value with a large number of photons created. An equipartition of number and energy between electron–positron pairs and photons is reached (see Fig. 9.4). In Section 9.8, we introduce an approach based on the following three equations: the number density continuity equation, the energy-momentum conservation equation and the Maxwell equations. We describe the plasma oscillation for both overcritical electric field $E > E_c$ and undercritical electric field $E < E_c$ [73]. In additional of reviewing the result well known in the literature for $E > E_c$ we review some novel result for the case $E < E_c$. It was traditionally assumed that electron–positron pairs, created by the vacuum polarization process, move as charged particles in external uniform electric field reaching arbitrary large Lorentz factors. It is reviewed how recent computations show the existence of plasma oscillations of the electron–positron pairs also for $E \lesssim E_c$. For both cases we quote the maximum Lorentz factors $\gamma_{\text{max}}$ reached by the electrons and positrons as well as the length of oscillations. Two specific cases are given. For $E_0 = 10E_c$ the length of oscillations $10 \hbar / (m_e c)$, and $E_0 = 0.15E_c$ the length of oscillations $10^7 \hbar / (m_e c)$. We also review the asymptotic behavior in time, $t \to \infty$, of the plasma oscillations by the phase portrait technique. Finally we review some recent results which differentiate the case $E > E_c$ from the one $E < E_c$ with respect to the creation of the rest mass of the pair versus their kinetic energy. For $E > E_c$ the vacuum polarization process transforms the electromagnetic energy of the field mainly in the rest mass of pairs, with moderate contribution to their kinetic energy.

We then turn in Section 10 to the last physical process needed in ascertaining the reaching of equilibrium of an optically thick electron–positron plasma. The average energy of electrons and positrons we illustrate is $0.1 < \epsilon < 10$ MeV. These bounds are necessary from the one hand to have significant amount of electron–positron pairs to make the plasma optically thick, and from the other hand to avoid production of other particles such as muons. As we will see in the next report these are indeed the relevant parameters for the creation of ultrarelativistic regimes to be encountered in pair creation process during the formation phase of a black hole. We then review the problem of evolution of optically thick, nonequilibrium electron–positron plasma, to-
wards an equilibrium state, following [74,75]. These results have been mainly obtained by two of us (RR and GV) in recent publications and all relevant previous results are also reviewed in this Section 10. We have integrated directly relativistic Boltzmann equations with all binary and triple interactions between electrons, positrons and photons two kinds of equilibrium are found: kinetic and thermal ones. Kinetic equilibrium is obtained on a timescale of few \((\sigma_T n_{\pm} c)^{-1}\), where \(\sigma_T\) and \(n_{\pm}\) are Thomson’s cross-section and electron–positron concentrations respectively, when detailed balance is established between all binary interactions in plasma. Thermal equilibrium is reached on a timescale of few \((\alpha \sigma_T n_{\pm} c)^{-1}\), when all binary and triple, direct and inverse interactions are balanced. In Section 10.1 basic plasma parameters are illustrated. The computational scheme as well as the discretization procedure are discussed in Section 10.2. Relevant conservation laws are given in Section 10.3. Details on binary interactions, consisting of Compton, Møller and Bhabha scatterings, Dirac pair annihilation and Breit–Wheeler pair creation processes, and triple interactions, consisting of relativistic bremsstrahlung, double Compton process, radiative pair production and three photon annihilation process, are presented in Section 10.5 and 10.6, respectively. In Section 10.5 collisional integrals with binary interactions are computed from first principles, using QED matrix elements. In Section 10.7 Coulomb scattering and the corresponding cutoff in collisional integrals are discussed. Numerical results are presented in Section 10.8 where the time dependence of energy and number densities as well as chemical potential and temperature of electron–positron-photon plasma is shown, together with particle spectra. The most interesting result of this analysis is to have differentiate the role of binary and triple interactions. The detailed balance in binary interactions following the classical work of Ehlers [76] leads to a distribution function of the form of the Fermi-Dirac for electron–positron pairs or of the Bose-Einstein for the photons. This is the reason we refer in the text to such conditions as the Ehlers equilibrium conditions. The crucial role of the direct and inverse three-body interactions is well summarized in fig. 10.1, panel A from which it is clear that the inverse three-body interactions are essential in reaching thermal equilibrium. If the latter are neglected, the system deflates to the creation of electron–positron pairs all the way down to the threshold of 0.5MeV. This last result which is referred as the Cavallo–Rees scenario [77] is simply due to improper neglect of the inverse triple reaction terms.

In Section 11 we present some general remarks.

Here and in the following we will use Latin indices running from 1 to 3, Greek indices running from 0 to 3, and we will adopt the Einstein summation rule.

934
3 The fundamental contributions to the electron–positron pair creation and annihilation and the concept of critical electric field

In this Section we recall the annihilation process of an electron–positron pair with the production of two photons

\[ e^+ + e^- \rightarrow \gamma_1 + \gamma_2, \]  

(3.0.1)

studied by Dirac in [1], the Breit–Wheeler process of electron–positron pair production by light-light collisions [2]

\[ \gamma_1 + \gamma_2 \rightarrow e^+ + e^-. \]  

(3.0.2)

and the vacuum polarization in external electric field, introduced by Sauter [20]. These three results, obtained in the mid-30’s of the last century [78, 79], played a crucial role in the development of the Quantum Electro-Dynamics (QED).

3.1 Dirac’s electron–positron annihilation

Dirac had proposed his theory of the electron [80, 81] in the framework of relativistic quantum theory. Such a theory predicted the existence of positive and negative energy states. Only the positive energy states could correspond to the electrons. The negative energy states had to have a physical meaning since transitions were considered to be possible from positive to negative energy states. It was proposed by Dirac [81] that nearly all possible states of negative energy are occupied with just one electron in accordance with Pauli’s exclusion principle and that the unoccupied states, ‘holes’ in the negative energy states should be regarded as ‘positrons’\(^1\). Historical review of

\(^1\)Actually initially [1, 81] Dirac believed that these ‘holes’ in negative energy spectrum describe protons, but later he realized that these holes represent particles with the same mass as of electron but with opposite charge, ‘anti-electrons’ [82]. The discovery of these anti-electrons was made by Anderson in 1932 [83] and named by him ‘positrons’ [84].
3 The fundamental contributions to the electron–positron pair creation and annihilation and the concept of critical electric field

Adopting his time-dependent perturbation theory [86] in the framework of relativistic Quantum Mechanics Dirac pointed out in [1] the necessity of the annihilation process of electron–positron pair into two photons (3.0.1). He considered an electron under the simultaneous influence of two incident beams of radiation, which induce transition of the electron to states of negative energy, then he calculated the transition probability per unit time, using the well established validity of the Einstein emission and absorption coefficients, which connect spontaneous and stimulated emission probabilities. He obtained the explicit expression of the cross-section of the annihilation process.

Such process is spontaneous, i.e. it occurs necessarily for any pair of electron and positron independently of their energy. The process does not need any previously existing radiation. The derivation of the cross-section, considering the stimulated emission process, was simplified by the fact that the electromagnetic field could be treated as an external classical perturbation and did not need to be quantized [87].

Dirac started from his wave equation [80] for the spinor field $\Psi$:

$$\left\{ \frac{E}{c} + \frac{\epsilon}{c} A_0 + \alpha \cdot \left( p + \frac{\epsilon}{c} A \right) + \beta_D m_e c \right\} \Psi = 0,$$

(3.1.1)

where $m_e$ and $e$ are electron’s mass and charge, $A$ is electromagnetic vector potential, and the matrices $\alpha$ and $\beta_D$ are:

$$\alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix} \quad \text{and} \quad \beta_D = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix},$$

(3.1.2)

where $\sigma$ and $I$ are respectively the Pauli’s and unit matrices. By choosing a gauge in which $A_0$ vanishes he obtained:

$$A = a_1 e^{i \omega_1 [t - l_1 \cdot x / c]} + a_1^* e^{-i \omega_1 [t - l_1 \cdot x / c]} + a_2 e^{i \omega_2 [t - l_2 \cdot x / c]} + a_2^* e^{-i \omega_2 [t - l_2 \cdot x / c]},$$

(3.1.3)

where $\omega_1$ and $\omega_2$ are respectively the frequencies of the two beams, $l_1$ and $l_2$ are the unit vectors in their direction of motion and $a_1$ and $a_2$ are the polarization vectors, the modulus of which are the amplitudes of the two beams.

Dirac solved Eq. (3.1.1) by a perturbation method, finding a solution of the form $\psi = \psi_0 + \psi_1 + \psi_2 + \ldots$, where $\psi_0$ is the solution in the free case, and $\psi_1$ is the first order perturbation containing the field $A$, or, explicitly $- \frac{\epsilon}{c} \alpha \cdot A$. He then computed the explicit expression of the second order expansion term $\psi_2$, which represents electrons that have made the double photon emission process and decay into negative energy states. He evaluated the transition
amplitude for the stimulated transition process, which reads

\[ w_{e^+e^-\rightarrow\gamma_1\gamma_2} = \frac{16e^2|a_1|^2|a_2|^2}{|\mathcal{E}'|m_e c^2} K_{12} \frac{1 - \cos(\delta \mathcal{E}'t/\hbar)}{(\delta \mathcal{E}')^2}, \]  

(3.1.4)

where \( \mathcal{E}' = m_e c^2 - \nu_1 - \nu_2, \nu_1 \) and \( \nu_2 \) are the photons’ frequencies and

\[ K_{12} = -(\mathbf{m}_1 \cdot \mathbf{m}_2)^2 + \frac{1}{4} [1 - (\mathbf{m}_1 \cdot \mathbf{n}_1)(\mathbf{n}_1 \cdot \mathbf{n}_2) + (\mathbf{m}_1 \cdot \mathbf{n}_2)(\mathbf{m}_1 \cdot \mathbf{n}_2)] \frac{v_1 + v_2}{m_e c^2}, \]  

(3.1.5)

is a dimensionless number depending on the unit vectors in the directions of the two photon’s polarization vectors \( \mathbf{m}_1 \) and \( \mathbf{m}_2 \). The quantities \( \mathbf{n}_1 \) and \( \mathbf{n}_2 \) are respectively given by \( \mathbf{n}_{1,2} = \mathbf{l}_{1,2} \times \mathbf{m}_{1,2} \). Introducing the intensity of the two incident beams

\[ I_1 = \frac{v_1^2}{2\pi c} |k_1|^2, \quad I_2 = \frac{v_2^2}{2\pi c} |k_2|^2, \]  

(3.1.6)

where \( k_{1,2} = \omega_{1,2} l_{1,2} \). Dirac obtained from the above transition amplitude the transition probability

\[ P_{e^+e^-\rightarrow\gamma_1\gamma_2} = \frac{8\pi^2 c^2 e^4}{|\mathcal{E}'|m_e c^2 v_1^2 v_2^2} K_{12} \frac{1 - \cos(\delta \mathcal{E}'t/\hbar)}{(\delta \mathcal{E}')^2}. \]  

(3.1.7)

In order to evaluate the spontaneous emission probability Dirac uses the relation between the Einstein coefficients \( A_E \) and \( B_E \) which is of the form

\[ A_E / B_E = 2\pi \hbar / c^2 (v_{1,2} / 2\pi)^3. \]  

(3.1.8)

Integrating on all possible directions of emission he obtains the total probability per unit time in the rest frame of the electron

\[ \sigma_{\text{lab}}^{e^+e^-} = \pi \left( \frac{\alpha \hbar}{m_e c} \right)^2 (\hat{\gamma} - 1)^{-1} \left\{ \frac{\hat{\gamma}^2 + 4\hat{\gamma} + 1}{\hat{\gamma}^2 - 1} \ln[\hat{\gamma} + (\hat{\gamma}^2 - 1)^{1/2}] - \frac{\hat{\gamma} + 3}{(\hat{\gamma}^2 - 1)^{1/2}} \right\}, \]  

(3.1.9)

where \( \hat{\gamma} \equiv \mathcal{E}_+ / m_e c^2 \geq 1 \) is the energy of the positron and \( \alpha = e^2 / (\hbar c) \) is the fine structure constant. Again, historically Dirac was initially confused about the negative energy states interpretation as we recalled. Although he derived the correct formula, he was doubtful about the presence in it of the mass of the electron or of the mass of the proton. Of course today this has been clarified and this derivation is fully correct if one uses the mass of the electron and applied this formula to description of electron–positron annihilation. The
limit for high-energy pairs ($\gamma \gg 1$) is
\[
\sigma_{e^+e^-}^{\text{lab}} \simeq \frac{\pi}{\gamma} \left( \frac{\alpha \hbar}{m_e c} \right)^2 \ln (2\gamma) - 1; \quad (3.1.10)
\]

The corresponding center of mass formula is
\[
\sigma_{e^+e^-} = \frac{\pi}{4\beta^2} \left( \frac{\alpha \hbar}{m_e c} \right)^2 [2\beta(\beta^2 - 2) + (3 - \beta^4) \ln \left( \frac{1 + \beta}{1 - \beta} \right)], \quad (3.1.11)
\]

where $\beta$ is the reduced velocity of the electron or the positron.

### 3.2 Breit–Wheeler pair production

We now turn to the equally important derivation on the production of an electron–positron pair in the collision of two real photons given by Breit and Wheeler [2]. According to Dirac’s theory of the electron, this process is caused by a transition of an electron from a negative energy state to a positive energy under the influence of two light quanta on the vacuum. This process differently from the one considered by Dirac, which occurs spontaneously, has a threshold due to the fact that electron and positron mass is not zero. In other words in the center of mass of the system there must be sufficient available energy to create an electron–positron pair. This energy must be larger than twice of electron rest mass energy.

Breit and Wheeler, following the discovery of the positron by Anderson [84], studied the effect of two light waves upon an electron in a negative energy state, represented by a normalized Dirac wave function $\psi^{(0)}$. Like in the previous case studied by Dirac [1] the light waves have frequencies $\omega_i$, wave vectors $k_i$ and vector potentials (3.1.3). Under the influence of the light waves, the initial electron wave function $\psi^{(0)}$ is changed after some time $t$ into a final wave function $\psi^{(t)}$. The method adopted is the time-dependent perturbation [86] (for details see [88]) to solve the Dirac equation with the time-dependent potential $eA(t)$ (3.1.3). The transition amplitude was calculated by an expansion in powers of $a_{1,2}$ up to $O(a^2)$. The wave function $\psi^{(t)}$ contains a term representing an electron in a positive energy state. The associated density is found to be

\[
w_{\gamma_1 + \gamma_2 \rightarrow e^+ + e^-} = \left( \frac{\alpha \hbar}{m_e c} \right)^2 |a_1|^2 |a_2|^2 K_{12} \frac{|1 - \exp(-i\delta E/\hbar)|^2}{(\delta E)^2}, \quad (3.2.1)
\]
\[
= \left( \frac{\alpha \hbar}{m_e c} \right)^2 2|a_1|^2 |a_2|^2 K_{12} \frac{|1 - \cos(\delta E t/\hbar)|^2}{(\delta E)^2}, \quad (3.2.2)
\]
where \( K_{12} \) is the dimensionless number already obtained by Dirac, Eq. (3.1.5), depending on initial momenta and spin of the wave function \( \psi^{(0)} \) and the polarizations of the quanta. This quantity is actually the squared transition matrix in the momenta and spin of initial and final states of light and electron–positron. The squared amplitudes \( |a_{1,2}|^2 \) in Eq. (3.2.1) are determined by the intensities \( I_{1,2} \) of the two light beams as

\[
|a_{1,2}|^2 = \frac{2\pi c}{\omega_{1,2}} I_{1,2}. \tag{3.2.3}
\]

The quantity \( \delta \mathcal{E} \) in Eq. (3.2.1) is the difference in energies between initial light states and final electron–positron states. Indicating by \( \mathcal{E}^{(-)} = -c(p_1^2 + m_e^2c^2)^{1/2} \), where \( p_1 \) is the four-momentum of the positron, the negative energy of the electron in its initial state and the corresponding quantity for the electron \( \mathcal{E}_2 = -c(p_2^2 + m_e^2c^2)^{1/2} \), where \( p_2 \) is the 4-momentum of the electron, \( \delta \mathcal{E} \) is given by

\[
\delta \mathcal{E} = c(p_2^2 + m_e^2c^2)^{1/2} + \mathcal{E}_1 - \hbar \omega_1 - \hbar \omega_2, \quad \text{where} \quad \mathcal{E}_1 = -\mathcal{E}^{(-)}, \tag{3.2.4}
\]

and \( p_2 = -p_1 + k_1 + k_2 \) is the final momentum of the electron. From this energy and momentum conservation it follows

\[
d(\delta \mathcal{E}) = c^2 \left[ \frac{|p_1|}{\mathcal{E}_1} - \frac{p_1 \cdot p_2}{(|p_1| \mathcal{E}_2)} \right] dp_1. \tag{3.2.5}
\]

It is then possible to sum the probability densities (3.2.1) over all possible initial electron states of negative energy in the volume \( V \). An integral over the phase space \( \int 2|p_1|^2 dp_1 |d\Omega_1 V / (2\pi \hbar)^3 \) must be performed. The effective collision area for the head-on collision of two light quanta was shown by Breit and Wheeler to be

\[
\sigma_{\gamma\gamma} = 2 \left( \frac{\alpha \hbar}{m_e c} \right)^2 \int \frac{c|p_1|^2}{\hbar \omega_1 \hbar \omega_2} K_{12} \left[ \frac{|p_1|}{\mathcal{E}_1} - \frac{p_1 \cdot p_2}{|p_1| \mathcal{E}_2} \right]^{-1} d\Omega_1, \tag{3.2.6}
\]

where \( \Omega_1 \) is the solid angle, which fulfills the total energy conservation \( \delta \mathcal{E} = 0 \).

In the center of mass of the system, the momenta of the electron and the positron are equal and opposite \( p_1 = -p_2 \). In that frame the momenta of the photons in the initial state are \( k_1 = -k_2 \). As a consequence, the energies of the electron and the positron are equal: \( \mathcal{E}_1 = \mathcal{E}_2 = \mathcal{E} \), and so are the energies of the photons: \( \hbar \omega_1 = \hbar \omega_2 = \mathcal{E}_\gamma = \mathcal{E} \). The total cross-section of the process is then

\[
\sigma_{\gamma\gamma} = 2 \left( \frac{\alpha \hbar}{m_e c} \right)^2 \frac{c|p|}{\mathcal{E}} \int K_{12} d\Omega_1, \tag{3.2.7}
\]
where \( |\mathbf{p}| = |\mathbf{p}_1| = |\mathbf{p}_2| \), and \( \mathcal{E} = (c^2 |\mathbf{p}|^2 + m_e^2 c^4)^{1/2} \). Therefore, the necessary kinematic condition in order for the process (3.0.2) taking place is that the energy of the two colliding photons be larger than the threshold \( 2m_e c^2 \), i.e.,

\[ \mathcal{E}_\gamma > m_e c^2. \]  

(3.2.8)

From Eq. (3.2.7) the total cross-section in the center of mass of the system is

\[ \sigma_{\gamma\gamma} = \frac{\pi}{2} \left( \frac{\alpha \hbar}{m_e c} \right)^2 (1 - \hat{\beta}^2) \left[ 2\hat{\beta}(\hat{\beta}^2 - 2) + (3 - \hat{\beta}^4) \ln \left( \frac{1 + \hat{\beta}}{1 - \hat{\beta}} \right) \right], \quad \text{with} \quad \hat{\beta} = \frac{c |\mathbf{p}|}{\mathcal{E}}. \]  

(3.2.9)

In modern QED cross-sections (3.1.10) and (3.2.9) emerge from two tree-level Feynman diagrams (see, for example, the textbook [89] and Section 5).

For \( \mathcal{E} \gg m_e c^2 \), the total effective cross-section is approximately proportional to

\[ \sigma_{\gamma\gamma} \simeq \pi \left( \frac{\alpha \hbar}{m_e c} \right)^2 \left( \frac{m_e c^2}{\mathcal{E}} \right)^2. \]  

(3.2.10)

The cross-section in line (3.2.9) can be easily generalized to an arbitrary reference frame, in which the two photons \( k_1 \) and \( k_2 \) cross with arbitrary relative directions. The Lorentz invariance of the scalar product of their 4-momenta \( (k_1 k_2) \) gives \( \omega_1 \omega_2 = \mathcal{E}_\gamma^2 \). Since \( \mathcal{E}_\gamma = \mathcal{E} = m_e c^2 / \sqrt{1 - \hat{\beta}^2} \), to obtain the total cross-section in the arbitrary frame \( \mathcal{K} \), we must therefore make the following substitution [90]

\[ \hat{\beta} \to \sqrt{1 - m_e^2 c^4 / (\omega_1 \omega_2)}, \]  

(3.2.11)

in Eq. (3.2.9).

### 3.3 Collisional \( e^+ e^- \) pair creation near nuclei:
Bethe and Heitler, Landau and Lifshitz, Sauter, and Racah

After having recalled in the previous sections the classical works of Dirac on the reaction (3.0.1) and Breit–Wheeler on the reaction (3.0.2) it is appropriate to return for a moment on the discovery of electron–positron pairs from observations of cosmic rays. The history of this discovery sees as major actors on one side Carl Anderson [84] at Caltech and on the other side Patrick Maynard Stuart Blackett and Giuseppe Occhialini [91] at the Cavendish laboratory. A fascinating reconstruction of their work can be found e.g. in [85]. The scene was however profoundly influenced by a fierce conceptual battle between Robert A. Millikan at Caltech and Arthur Compton at Chicago on the mechanism of production of these cosmic rays. For a refreshing mem-
ory of these heated discussions and a role also of Sir James Hopwood Jeans see e.g. [92]. The contention by Millikan was that the electron–positron pairs had to come from photons originating between the stars, while Jeans located their source on the stars. Compton on the contrary insisted on their origin from the collision of charged particles in the Earth atmosphere. Moreover, at the same time there were indications that similar process of charged particles would occur by the scattering of the radiation from polonium-beryllium, see e.g. Joliot and Curie [93].

It was therefore a natural outcome that out of this scenario two major theoretical developments occurred. One development inquired electron–positron pair creation by the interaction of photons with nuclei following the reaction:

\[
\gamma + Z \rightarrow Z + e^+ + e^-,
\]  

(3.3.1)

major contributors were Oppenheimer and Plesset [94], Heitler [95], Bethe-Heitler [96], Sauter [97] and Racah [98]. Heitler [95] obtained an order of magnitude estimate of the total cross-section of this process

\[
\sigma_{Z\gamma \rightarrow Ze^+e^-} \simeq aZ^2 \left( \frac{e^2}{m_e c^2} \right)^2.
\]  

(3.3.2)

In the ultrarelativistic case \( \epsilon_\pm \gg m_e \) the total cross-section for pair production by a photon with a given energy \( \omega \) is [96]

\[
\sigma = \frac{28}{9} Z^2 \alpha r_e^2 \left( \log \frac{2\omega}{m_e} - \frac{109}{42} \right).
\]  

(3.3.3)

The second development was the study of the reaction

\[
Z_1 + Z_2 \rightarrow Z_1 + Z_2 + e^+ + e^-.
\]  

(3.3.4)

with the fundamental contribution of Landau and Lifshitz [99] and Racah [98, 100]. This process is an example of two photon pair production, see. Fig. 3.1. The 4-momenta of particles \( Z_1 \) and \( Z_2 \) are respectively \( p_1 \) and \( p_2 \). The

\[\text{Figure 3.1:} \text{ The sketch of two photon particle production. Reproduced from [101].} \]
The fundamental contributions to the electron–positron pair creation and annihilation and the concept of critical electric field

total pair production cross-section is [99]

\[ \sigma_{\text{Landau}} = \frac{28}{27 \pi} r_e^2 (Z_1 Z_2 \alpha)^2 L^3_{\gamma}, \quad (3.3.5) \]

where \( L_{\gamma} = \log \gamma \). Racah [100] gives next to leading terms

\[ \sigma_{\text{Racah}} = \frac{28}{27 \pi} r_e^2 (Z_1 Z_2 \alpha)^2 (L^3_{\gamma} - 2.2 L^2_{\gamma} + 3.84 L_{\gamma} - 1.636). \quad (3.3.6) \]

The differential cross-section is given in Section 5.5. The differential distributions of electrons and positrons in a wide energy range was computed by Bhabha in [102].

In parallel progress on the reaction

\[ e^- + Z \rightarrow e^- + Z + \gamma, \quad (3.3.7) \]

was made by Sommerfeld [103], Heitler [95] and later by Bethe and Heitler [96].

Once the exact cross-section of the process (3.3.1) was known, the corresponding cross-section for the process (3.3.7) was found by an elegant method, called the equivalent photons method [101, 104]. The idea to treat the field of a fast charged particle in a way similar to electromagnetic radiation with particular frequency spectrum goes back to Fermi [105]. In such a way electromagnetic interaction of this particle e.g. with a nucleus is reduced to the interaction of this radiation with the nucleus. This idea was successfully applied to the calculation of the cross-section of interaction of relativistic charged particles by Weizsäcker [106] and Williams [107]. In fact, this method establishes the relation between the high-energy photon induced cross-section \( d\sigma_{\gamma X \rightarrow Y} \) to the corresponding cross-section induced by a charged particle \( d\sigma_{e X \rightarrow Y} \) by the relation which is expressed by

\[ d\sigma_{e X \rightarrow Y} = \int \frac{n(\omega)}{\omega} d\sigma_{\gamma X \rightarrow Y} d\omega, \quad (3.3.8) \]

where \( n(\omega) \) is the spectrum of equivalent photons. Its simple generalization,

\[ \sigma_{ee \rightarrow Y} = \int \frac{d\omega_1}{\omega_1} \frac{d\omega_2}{\omega_2} n(\omega_1)n(\omega_2)d\sigma_{\gamma_1\gamma_2 \rightarrow Y}. \quad (3.3.9) \]

Generally speaking, the equivalent photon approximation consists in ignoring that in such a case intermediate (virtual) photons are a) off mass shell and b) no longer transversely polarized. In the early years this spectrum was
3.4 Klein paradox and Sauter work

Every relativistic wave equation of a free particle of mass $m_e$, momentum $p$ and energy $E$, admits “positive energy” and “negative energy” solutions. In Minkowski space such a solution is symmetric with respect to the zero energy and the wave function given by

$$\psi^\pm(x, t) \sim e^{i\hat{E}(k \cdot x - E^\pm t)}$$

(3.4.1)

describes a relativistic particle, whose energy, mass and momentum satisfy,

$$E^2_\pm = m_e^2 c^4 + c^2 |p|^2; \quad E_\pm = \pm \sqrt{m_e^2 c^4 + c^2 |p|^2}.$$  

(3.4.2)

This gives rise to the familiar positive and negative energy spectrum ($E_\pm$) of positive and negative energy states $\psi^\pm(x, t)$ of the relativistic particle, as represented in Fig. 3.2. In such a situation, in absence of external field and at zero temperature, all the quantum states are stable; that is, there is no possibility of “positive” (“negative”) energy states decaying into “negative” (“positive”) energy states since there is an energy gap $2m_e c^2$ separating the negative energy spectrum from the positive energy spectrum. This stability condition was implemented by Dirac by considering all negative energy states as fully filled.

A scalar field described by the wave function $\phi(x)$ satisfies the Klein–Gordon
The fundamental contributions to the electron–positron pair creation and annihilation and the concept of critical electric field

\[
\left\{ \left[ i \hbar \partial_\mu + \frac{e}{c} A_\mu (z) \right]^2 - m_e^2 c^2 \right\} \phi (x) = 0. \tag{3.4.3}
\]

If there is only an electric field \( E(z) \) in the \( z \)-direction and varying only as a function of \( z \), we can choose a vector potential with the only nonzero component \( A_0 (z) \) and potential energy

\[
V (z) = -e A_0 (z) = e \int^z dz' E (z'). \tag{3.4.4}
\]

For an electron of charge \(-e\) by assuming \( \phi (x) = e^{-iEt/\hbar} e^{i p_\perp x_\perp / \hbar} \phi (z) \),

with a fixed transverse momentum \( p_\perp \) in the \( x, y \)-direction and an energy eigenvalue \( \mathcal{E} \), and Eq. (3.4.3) becomes simply

\[
\left[ -\frac{\hbar^2}{d^2} + p_\perp^2 + m_e^2 c^2 - \frac{1}{c^2} [\mathcal{E} - V (z)]^2 \right] \phi (z) = 0. \tag{3.4.5}
\]

Klein studied a relativistic particle moving in an external step function potential \( V (z) \) and in this case Eq. (3.4.5) is modified as

\[
[\mathcal{E} - V_0]^2 = m_e^2 c^4 + c^2 |p|^2; \quad \mathcal{E}_\pm = V_0 \pm \sqrt{m_e^2 c^4 + c^2 |p|^2}, \tag{3.4.6}
\]

where \( |p|^2 = |p_\perp|^2 + p_z^2 \). He solved his relativistic wave equation [109–112] by considering an incident free relativistic wave of positive energy states scattered by the constant potential \( V_0 \), leading to reflected and transmitted waves. He found a paradox that in the case \( V_0 \geq \mathcal{E} + m_e c^2 \), the reflected flux is larger than the incident flux \( j_{\text{ref}} > j_{\text{inc}} \), although the total flux is conserved, i.e. \( j_{\text{inc}} = j_{\text{ref}} + j_{\text{tran}} \). This is known as the Klein paradox [17, 18]. This implies that negative energy states have contributions to both the transmitted flux \( j_{\text{tran}} \) and reflected flux \( j_{\text{ref}} \).

Sauter studied this problem by considering a potential varying in the \( z \)-direction corresponding to a constant electric field \( E \) in the \( \hat{z} = z / |z| \)-direction and considering spin 1/2 particles fulfilling the Dirac equation. In this case the energy \( \mathcal{E} \) is shifted by the amount \( V (z) = -e Ez \). He further assumed an electric field \( E \) uniform between \( z_1 \) and \( z_2 \) and null outside. Fig. 3.3 represents the corresponding sketch of allowed states. The key point now, which is the essence of the Klein paradox [17, 18], is that a level crossing between the positive and negative energy levels occurs. Under this condition the above mentioned stability of the “positive energy” states is lost for sufficiently strong electric fields. The same is true for “negative energy” states.
3.4 Klein paradox and Sauter work

Figure 3.2: The mass-gap $2m_ec^2$ that separates the positive continuum spectrum $\mathcal{E}_+$ from the negative continuum spectrum $\mathcal{E}_-$. 
Figure 3.3: In presence of a strong enough electric field the boundaries of the classically allowed states ("positive" or "negative") can be so tilted that a "negative" is at the same level as a "positive" (level crossing). Therefore a "negative" wave-packet from the left will be partially transmitted, after an exponential damping due to the tunneling through the classically forbidden states, as a "positive" wave-packet outgoing to the right. This figure is reproduced from Fig. II in Ref. [113], and \( \mu = m_e c^2, \epsilon V = V(z), \omega = \check{\epsilon} \).

Some "positive energy" and "negative energy" states have the same energy levels. Thus, these "negative energy" waves incident from the left will be both reflected back by the electric field and partly transmitted to the right as a "positive energy" wave, as shown in Fig. 3.3 [113]. This transmission represents a quantum tunneling of the wave function through the electric potential barrier, where classical states are forbidden. This quantum tunneling phenomenon was pioneered by George Gamow by the analysis of alpha particle emission or capture in the nuclear potential barrier (Gamow wall) [19]. In the latter case however the tunneling occurred between two states of positive energy while in the Klein paradox and Sauter computation the tunneling occurs for the first time between the positive and negative energy states giving rise to the totally new concept of the creation of particle-antiparticle pairs in the positive energy state as we are going to show.

Sauter first solved the relativistic Dirac equation in the presence of the constant electric field by the ansatz,

\[
\psi_s(x, t) = e^{i(k_x x + k_y y - E \pm t)} \chi_s(z) \tag{3.4.7}
\]

where spinor function \( \chi_s(z) \) obeys the following equation \((\gamma_0, \gamma_i\) are Dirac...\]
3.4 Klein paradox and Sauter work

\[
\left[ \hbar c \gamma_3 \frac{d}{dz} + \gamma_0 (V(z) - \varepsilon_\pm) + (m_e c^2 + ic \gamma_2 p_y + ic \gamma_1 p_x) \right] \chi_{3s}(z) = 0, \quad (3.4.8)
\]

and the solution \( \chi_{3s}(z) \) can be expressed in terms of hypergeometric functions [20]. Using this wave function \( \psi_s(x, t) \) (3.4.7) and the flux \( ic \psi_s^\dagger \gamma_3 \psi_s \), Sauter computed the transmitted flux of positive energy states, the incident and reflected fluxes of negative energy states, as well as exponential decaying flux of classically forbidden states, as indicated in Fig. 3.3. Using the regular matching conditions of the wave functions and fluxes at boundaries of the potential, Sauter found that the transmission coefficient \( |T|^2 \) of the wave through the electric potential barrier from the negative energy state to positive energy states:

\[
|T|^2 = \left| \frac{\text{transmission flux}}{\text{incident flux}} \right| \sim e^{-\frac{\pi m_e^2 c^3}{\hbar^2}}. \quad (3.4.9)
\]

This is the probability of negative energy states decaying to positive energy states, caused by an external electric field. The method that Sauter adopted to calculate the transmission coefficient \( |T|^2 \) is indeed the same as the one Gamow used to calculate quantum tunneling of the wave function through nuclear potential barrier, leading to the \( \alpha \)-particle emission [19].

The simplest way to calculate the transmission coefficient \( |T|^2 \) (3.4.9) is the JWKB (Jeffreys–Wentzel–Kramers–Brillouin) approximation. The electric potential \( V(z) \) is not a constant. The corresponding solution of the Dirac equation is not straightforward, however it can be found using the quasi-classical, JWKB approximation. Particle’s energy \( \varepsilon \), momentum \( p \) and mass \( m_e \) satisfy,

\[
[\varepsilon_\pm - V(z)]^2 = m_e^2 c^4 + c^2 |p|^2; \quad \varepsilon_\pm = V(z) \pm \sqrt{m_e^2 c^4 + c^2 |p|^2}, \quad (3.4.10)
\]

where the momentum \( p_z(z) \) is spatially dependent. The momentum \( p_z > 0 \) for both negative and positive energy states and the wave functions exhibit usual oscillatory behavior of propagating wave in the \( \hat{z} \)-direction, i.e. \( \exp i \frac{\hbar}{\hbar} p_z z \). Inside the electric potential barrier where are the classically forbidden states, the momentum \( p_z^2 \) given by Eq. (3.4.10) becomes negative, and \( p_z \) becomes imaginary, which means that the wave function will have an exponential behavior, i.e. \( \exp -\frac{1}{\hbar} \int |p_z| dz \), instead of the oscillatory behavior which characterizes the positive and negative energy states. Therefore the transmission coefficient \( |T|^2 \) of the wave through the one-dimensional potential barrier is given by

\[
|T|^2 \propto \exp -\frac{2}{\hbar} \int_{z_-}^{z_+} |p_z| dz, \quad (3.4.11)
\]
where \( z_- \) and \( z_+ \) are roots of the equation \( p_z(z) = 0 \) defining the turning points of the classical trajectory, separating positive and negative energy states.

### 3.5 A semi-classical description of pair production in quantum mechanics

#### 3.5.1 An external constant electric field

The phenomenon of pair production can be understood as a quantum mechanical tunneling process of relativistic particles. The external electric field modifies the positive and negative energy spectrum of the free Hamiltonian. Let the field vector \( \mathbf{E} \) point in the \( \hat{z} \)-direction. The electric potential is \( A_0 = -|E|z \) where \(-\ell < z < +\ell\) and the length \( \ell \gg \hbar/\left(m_e c\right)\), then the positive and negative continuum energy spectra are

\[
E_{\pm} = |eE|z \pm \sqrt{(cp_z)^2 + c^2p_{\perp}^2 + (m_e c^2)^2},
\]

where \( p_z \) is the momentum in \( \hat{z} \)-direction, \( p_{\perp} \) transverse momenta. The energy spectra \( E_{\pm} \) (3.5.1) are sketched in Fig. 3.4. One finds that crossing energy levels \( E \) between two energy spectra \( E_- \) and \( E_+ \) (3.5.1) appear, then quantum tunneling process occurs. The probability amplitude for this process can be estimated by a semi-classical calculation using JWKB method (see e.g. [54, 88]):

\[
P_{\text{JWKB}}(|p_{\perp}|) \equiv \exp \left\{ -\frac{2}{\hbar} \int_{z_-}^{z_+(E_+)} p_z dz \right\},
\]

where

\[
p_z = \sqrt{p_{\perp}^2 + m_e^2 c^2 - (E - |eE|z)^2/c^2}
\]

is the classical momentum. The limits of integration \( z_\pm(E_{\pm}) \) are the turning points of the classical orbit in imaginary time. They are determined by setting \( p_z = 0 \) in Eq. (3.5.1). The solutions are

\[
z_\pm(E_{\pm}) = c\frac{\left[ p_{\perp}^2 + m_e^2 c^2 \right]^{1/2} + E_{\pm}}{|eE|},
\]

At the turning points of the classical orbit, the crossing energy level

\[
E = E_+ = E_-,
\]
3.5 A semi-classical description of pair production in quantum mechanics

as shown by dashed line in Fig. 3.4. The tunneling length is

\[ z_+(\mathcal{E}_+) - z_-(\mathcal{E}_-) = \frac{2mc^2}{|eE|} = 2 \frac{\hbar}{m_e c} \left( \frac{E_c}{E} \right), \]  

(3.5.6)

which is independent of crossing energy levels \( \mathcal{E} \). The critical electric field \( E_c \) in Eq. (2.0.1) is the field at which the tunneling length (3.5.6) is twice the Compton length \( \lambda_C \equiv \hbar / m_e c \).

Changing the variable of integration from \( z \) to \( y(z) \),

\[ y(z) = \frac{E - |eE|z}{c \sqrt{p_{\perp}^2 + m_e^2 c^2}} \]

(3.5.7)

we obtain

\[ y_-(z) = -1, \quad y_+(z) = +1 \]

(3.5.8)

and the JWKB probability amplitude (3.5.2) becomes

\[ P_{\text{JWKB}}(|p_{\perp}|) = \exp \left[ -\frac{\pi E_c}{E} \right] \frac{1}{\sqrt{1 - y^2}} \int_{-1}^{+1} dy \sqrt{1 - y^2} \]

(3.5.9)

Summing over initial and final spin states and integrating over the transverse phase space \( \int d\mathbf{z}_\perp d\mathbf{p}_\perp / (2\pi \hbar)^2 \) yields the final result

\[ P_{\text{JWKB}} \approx D_s V_{\perp} e^{-\frac{\pi mc^2}{|eE|\hbar} \int \frac{d^2p_{\perp}}{(2\pi \hbar)^2} e^{-\frac{\pi c p_{\perp}^2}{|eE|\hbar}}} = D_s V_{\perp} \frac{|eE|}{4\pi^2 c \hbar} e^{-\frac{\pi E_c}{E}}, \]

(3.5.10)

where the transverse surface \( V_{\perp} = \int d\mathbf{z}_\perp \). For the constant electric field \( E \) in \(-\ell < z < +\ell\), crossing energy levels \( \mathcal{E} \) vary from the maximal energy potential \( V(-\ell) = +eE \ell \) to the minimal energy potential \( V(+\ell) = -eE \ell \). This probability Eq. (3.5.10) is independent of crossing energy levels \( \mathcal{E} \). We integrate Eq. (3.5.10) over crossing energy levels \( \int d\mathcal{E} / m_e c^2 \) and divide it by the time interval \( \Delta t \approx \hbar / m_e c^2 \) during which quantum tunneling occurs, and find the transition rate per unit time and volume

\[ \frac{\Gamma_{\text{JWKB}}}{V} \approx D_s \frac{aE^2}{2\pi^2 \hbar} e^{-\frac{\pi E_c}{E}}, \]

(3.5.11)

where \( D_s = 2 \) for a spin-1/2 particle and \( D_s = 1 \) for spin-0, \( V \) is the volume. The JWKB result contains the Sauter exponential \( e^{-\frac{\pi E_c}{E}} \) [20] and reproduces as well the prefactor of Heisenberg and Euler [7].
The fundamental contributions to the electron–positron pair creation and annihilation and the concept of critical electric field

\[ \mathcal{E}_+ = +\sqrt{m_e^2c^4 + c^2|\mathbf{p}|^2} - |eE|z \]

\[ \mathcal{E}_- = -\sqrt{m_e^2c^4 + c^2|\mathbf{p}|^2} - |eE|z \]

Figure 3.4: Energy-spectra \( \mathcal{E}_\pm \) with an external electric field \( E \) along \( \hat{z} \)-direction (for \(-\ell < z < \ell \) and \( \ell \gg 1 \)). Crossing energy-levels appear, indicated by a dashed line between two continuum energy-spectra \( \mathcal{E}_- \) and \( \mathcal{E}_+ \). The turning points \( z_\pm(\mathcal{E}) \) for the crossing energy-levels \( \mathcal{E} \) of Eq. (3.5.5) are marked. This implies that virtual electrons at these crossing energy-levels in the negative energy-spectrum can quantum-mechanically tunnel toward infinity \([z \gg z_+(\mathcal{E})]\) as real electrons; empty states left over in the negative energy-spectrum represent real positrons. This is how quantum tunneling produces pairs of electrons and positrons.

Let us specify a quantitative condition for the validity of the above “semiclassical” JWKB approximation, which is in fact leading term of the expansion of wave function in powers of \( \hbar \). In order to have the next-leading term be much smaller than the leading term, the de Broglie wavelength \( \lambda(z) \equiv 2\pi\hbar/p_z(z) \) of wave function of the tunneling particle must have only small spatial variations [88]:

\[ \frac{1}{2\pi} \left| \frac{d\lambda(z)}{dz} \right| = \frac{\hbar}{p_z^2(z)} \left| \frac{dp_z(z)}{dz} \right| \ll 1. \quad (3.5.12) \]
3.5 A semi-classical description of pair production in quantum mechanics

with \( p_z(z) \) of Eq. (3.5.3). The electric potential \( A_0 = -e|z|Z \) must satisfy

\[
\frac{\hbar}{2p_z^2} \left| \frac{dA_0}{dz} \right| \approx \frac{E}{E_c} \ll 1. \tag{3.5.13}
\]

so that the result (3.5.11) is valid only for \( E \ll E_c \).

3.5.2 An additional constant magnetic field

The result (3.5.11) can be generalized to include a uniform magnetic field \( B \). The calculation is simplest by going into a Lorentz frame in which \( B \) and \( E \) are parallel to each other, which is always possible for uniform and static electromagnetic field. This frame will be referred to as a *center-of-fields frame*, and the associated fields will be denoted by \( B_{\text{CF}} \) and \( E_{\text{CF}} \). Suppose the initial \( B \) and \( E \) are not parallel, then we perform a Lorentz transformation with a velocity determined by

\[
\frac{\mathbf{v}}{c} = \frac{\mathbf{E} \times \mathbf{B}}{|\mathbf{E}|^2 + |\mathbf{B}|^2}, \tag{3.5.14}
\]

in the direction \( \mathbf{\hat{v}} \equiv \mathbf{v}/|\mathbf{v}| \) as follows

\[
\begin{align*}
E_{\text{CF}} &= (\mathbf{E} \cdot \mathbf{\hat{v}})\mathbf{\hat{v}} + \frac{\mathbf{\hat{v}} \times (\mathbf{E} \times \mathbf{\hat{v}}) + (\mathbf{v}/c) \times \mathbf{B}}{[1 - (|\mathbf{v}|/c)^2]^{1/2}}, \tag{3.5.15} \\
B_{\text{CF}} &= (\mathbf{B} \cdot \mathbf{\hat{v}})\mathbf{\hat{v}} - \frac{\mathbf{\hat{v}} \times (\mathbf{B} \times \mathbf{\hat{v}}) - (\mathbf{v}/c) \times \mathbf{E}}{[1 - (|\mathbf{v}|/c)^2]^{1/2}}. \tag{3.5.16}
\end{align*}
\]

The fields \( B_{\text{CF}} \) and \( E_{\text{CF}} \) are now parallel. As a consequence, the wave function factorizes into a Landau state and into a spinor function, this last one first calculated by Sauter (see Eqs. (3.4.7),(3.4.8)). The energy spectrum in the JWKB approximation is still given by Eq. (3.5.1), but the squared transverse momenta \( p_{\perp}^2 \) is quantized due to the presence of the magnetic field: they are replaced by the Landau energy levels whose transverse energies have the discrete spectrum

\[
c^2 p_{\perp}^2 = 2m_e c^2 \times \frac{p_{\perp}^2}{2m_e} \rightarrow 2m_e c^2 \times \frac{\hbar \omega_L}{2} \left( n + \frac{1}{2} + \hat{\sigma} \right), \ n = 0, 1, 2, \cdots, \tag{3.5.17}
\]

where \( g = 2 + \alpha / \pi + \ldots \) is the anomalous magnetic moment of the electron [11,115–118], \( \omega_L = e|B_{\text{CF}}|/m_ec \) the Landau frequency, \( \hat{\sigma} = \pm 1/2 \) for a spin-1/2 particle (\( \hat{\sigma} = 0 \) for a spin-0 particle) are eigenvalues of spinor operator \( \sigma_z \) in the \((\mathbf{\hat{z}})\)-direction, i.e., in the common direction of \( E_{\text{CF}} \) and \( B_{\text{CF}} \) in the selected frame. The quantum number \( n \) characterizing the Landau levels is associated with harmonic oscillations in the plane orthogonal to \( E_{\text{CF}} \) and \( B_{\text{CF}} \). Apart from the replacement (3.5.17), the JWKB calculation remains the same.
The fundamental contributions to the electron–positron pair creation and annihilation and the concept of critical electric field

as in the case of constant electric field (3.5.11). We must only replace the integration over the transverse phase space $\int dxdyd\mathbf{p}_\perp/(2\pi\hbar)^2$ in Eq. (3.5.10) by the sum over all Landau levels with the degeneracy $V_\perp|e|\mathbf{B}_{\text{CF}}|/(2\pi\hbar c)$ [88]:

$$
\frac{V_\perp|e|\mathbf{B}_{\text{CF}}}{2\pi\hbar c}\sum_{n,\sigma}\exp\left[-\pi\frac{2\hbar|e||\mathbf{B}_{\text{CF}}|(n + 1/2 + \sigma) + (me^2c^2)}{e|\mathbf{E}_{\text{CF}}|}\right].
$$

(3.5.18)

The results are

$$
\frac{V_\perp|e|\mathbf{B}_{\text{CF}}}{2\pi\hbar c}\coth\left(\frac{\pi|e|\mathbf{B}_{\text{CF}}}{|\mathbf{E}_{\text{CF}}|}\right)\exp\left(-\frac{\pi E_c}{|\mathbf{E}_{\text{CF}}|}\right), \text{ spin } -1/2 \text{ particle (3.5.19)}
$$

and

$$
\frac{V_\perp|e|\mathbf{B}_{\text{CF}}}{4\pi\hbar c}\sinh^{-1}\left(\frac{\pi|e|\mathbf{B}_{\text{CF}}}{|\mathbf{E}_{\text{CF}}|}\right)\exp\left(-\frac{\pi E_c}{|\mathbf{E}_{\text{CF}}|}\right), \text{ spin } 0 \text{ particle. (3.5.20)}
$$

We find the pair production rate per unit time and volume

$$
\Gamma_{\text{JWKB}}/V \approx \frac{\alpha|\mathbf{B}_{\text{CF}}||\mathbf{E}_{\text{CF}}|}{\pi\hbar}\coth\left(\frac{\pi|\mathbf{B}_{\text{CF}}|}{|\mathbf{E}_{\text{CF}}|}\right)\exp\left(-\frac{\pi E_c}{|\mathbf{E}_{\text{CF}}|}\right), \text{ spin } -1/2 \text{ particle}
$$

(3.5.21)

and

$$
\Gamma_{\text{JWKB}}/V \approx \frac{\alpha|\mathbf{B}_{\text{CF}}||\mathbf{E}_{\text{CF}}|}{2\pi\hbar}\sinh^{-1}\left(\frac{\pi|\mathbf{B}_{\text{CF}}|}{|\mathbf{E}_{\text{CF}}|}\right)\exp\left(-\frac{\pi E_c}{|\mathbf{E}_{\text{CF}}|}\right), \text{ spin } 0 \text{ particle. (3.5.22)}
$$

We can now go back to an arbitrary Lorentz frame by expressing the result in terms of the two Lorentz invariants that can be formed from the $\mathbf{B}$ and $\mathbf{E}$ fields: the scalar $S$ and the pseudoscalar $P$

$$
S \equiv \frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2); \quad P \equiv \frac{1}{4}F_{\mu\nu}\tilde{F}^{\mu\nu} = \mathbf{E} \cdot \mathbf{B},
$$

(3.5.23)

where $\tilde{F}^{\mu\nu} \equiv e^{\mu\nu\lambda\kappa}F_{\lambda\kappa}$ is the dual field tensor. We define the invariants $\varepsilon$ and $\beta$ as the solutions of the invariant equations

$$
\varepsilon^2 - \beta^2 \equiv \mathbf{E}^2 - \mathbf{B}^2 \equiv 2S, \quad \varepsilon\beta \equiv \mathbf{E} \cdot \mathbf{B} \equiv P,
$$

(3.5.24)

and obtain

$$
\varepsilon \equiv \sqrt{(S^2 + P^2)^{1/2} + S},
$$

(3.5.25)

$$
\beta \equiv \sqrt{(S^2 + P^2)^{1/2} - S}.
$$

(3.5.26)

In the special frame with parallel $\mathbf{B}_{\text{CF}}$ and $\mathbf{E}_{\text{CF}}$, we see that $\beta = |\mathbf{B}_{\text{CF}}|$ and $\varepsilon = |\mathbf{E}_{\text{CF}}|$, so that we can replace (3.5.21) and (3.5.22) directly by the invariant
3.5 A semi-classical description of pair production in quantum mechanics

expressions

\[
\frac{\Gamma_{\text{JWKB}}}{V} \simeq \frac{\alpha \beta \varepsilon}{2 \pi \hbar} \sinh\left(\frac{\pi \beta}{\varepsilon}\right) \exp\left(-\frac{\pi E_c}{\varepsilon}\right), \quad \text{spin} = 0 \text{ particle}, \quad (3.5.28)
\]

and

\[
\frac{\Gamma_{\text{JWKB}}}{V} \simeq \frac{\alpha \beta \varepsilon}{2 \pi \hbar} \coth\left(\frac{\pi \beta}{\varepsilon}\right) \exp\left(-\frac{\pi E_c}{\varepsilon}\right), \quad \text{spin} = \frac{1}{2} \text{ particle} \quad (3.5.27)
\]

which are pair production rates in arbitrary constant electromagnetic fields. We would like to point out that \( S \) and \( P \) in (3.5.23) are identically zero for any field configuration in which

\[
|E| = |B|; \quad E \perp B = 0. \quad (3.5.29)
\]

As example, for a plane wave of electromagnetic field, \( \varepsilon = \beta = 0 \) and no pairs are produced.
4 Nonlinear electrodynamics and rate of pair creation

4.1 Hans Euler and light-light scattering

Hans Euler in his celebrated diplom thesis [21] discussed at the University of Leipzig called attention on the reaction

\[ \gamma_1 \gamma_2 \rightarrow e^+ e^- \rightarrow \gamma'_1 \gamma'_2 \]

He recalled that Halpern [119] and Debye [120] first recognized that Dirac theory of electrons and the Dirac process (3.0.1) and the Breit–Wheeler one (3.0.2) had fundamental implication for the light on light scattering and consequently implied a modifications of the Maxwell equations.

If the energy of the photons is high enough then a real electron–positron pair is created, following Breit and Wheeler [2]. Again, if electron–positron pair does exist, two photons are created following [1]. In the case that the sum of energies of the two photons are smaller than the threshold \( 2m_e c^2 \) then the reaction (above) still occurs through a virtual pair of electron and positron.

Under this condition the light-light scattering implies deviation from superposition principle, and therefore the linear theory of electromagnetism has to be substituted by a nonlinear one. Maxwell equations acquire nonlinear corrections due to the Dirac theory of the electron.

Euler first attempted to describe this nonlinearity by an effective Lagrangian representing the interaction term. He showed that the interaction term had to contain the forth power of the field strengths and its derivatives

\[ \mathcal{E}_{\text{int}} = \text{const} \int \left[ FFFFF + \text{const}' \frac{\partial F}{\partial x} \frac{\partial F}{\partial x} FF + \ldots \right], \quad (4.1.1) \]

\( F \) being symbolically the electromagnetic field strength. He also estimated that the constants may be determined from dimensional considerations. Since the interaction \( U_{\text{int}} \) has the dimension of energy density and contains electric charge in the forth power, the constants up to numerical factors are

\[ \text{const} = \frac{\hbar c}{e^2 E_e^2}, \quad \text{const}' = \left( \frac{\hbar}{m_e c} \right)^2, \quad (4.1.2) \]
where $E_e = e \left( \frac{e^2}{m_e c^2} \right)^{-2} = \alpha^{-1} E_c$, namely “the field strength at the edge of the electron”.

From these general qualitative considerations Euler made an important further step taking into account that the Lagrangian (4.1.1) describing such a process had necessarily be built from invariants constructed from the field strengths, such as $E^2 - B^2$ and $E \cdot B$ following a precise procedure indicated by Max Born, see e.g. Pauli’s book [121]. Contrary to the usual Maxwell Lagrangian which is only a function of $F^2_{\mu \nu}$ Euler first recognized that virtual electron–positron loops are represented by higher powers in the field strength corrections to the linear action of electromagnetism and written down the Lagrangian with second order corrections

$$L = \frac{E^2 - B^2}{8\pi} + \frac{1}{a_{E} F^2_0} \left[ a_E \left( E^2 - B^2 \right)^2 + b_E (E \cdot B)^2 \right], \quad (4.1.3)$$

where

$$a_E = -1/(360\pi^2), \quad b_E = -7/(360\pi^2). \quad (4.1.4)$$

The crucial result of Euler has been to determine the values of the coefficients (4.1.4) using time-dependent perturbation technique, e.g. [88] in Dirac theory.

Euler computed only the lowest order corrections in $\alpha$ to Maxwell equations, namely “the 1/137 fraction of the field strength at the edge of the electron”. This perturbation method did not allow calculation of the tunneling rate for electron–positron pair creation in strong electromagnetic field which became the topic of the further work with Heisenberg [7].

### 4.2 Born’s nonlinear electromagnetism

A nonlinear theory of electrodynamics was independently proposed and developed by Max Born [4, 5] and later by Born and Infeld [6]. The main motivation in Born’s approach was the avoidance of infinities in an elementary particle description. Among the classical discussions on the fundamental interactions this topic had attracted attention of a large number of scientists. It was clear in fact from the considerations of J.J. Thomson, Abraham Lorentz that a point-like electron needed to have necessarily an infinite mass. The existence of a finite radius was attempted by Poincare by introduction of non-electromagnetic stresses. Also among the attempts we have to recall the theory of Mie [122–125] modifying the Maxwell theory by nonlinear terms. This theory however had serious difficulty because solutions of Mie field equations depend on the absolute value of the potentials.

Max Born developed his theory in collaboration with Infeld. This alternative to the Maxwell theory is today called the Born-Infeld theory which still finds interest in the framework of subnuclear physics. The coauthorship of
4.2 Born’s nonlinear electromagnetism

Infeld is felt by the general premise of the article in distinguishing the unitarian standpoint versus the dualistic standpoint in the description of particles and fields. "In the dualistic standpoint the particles are the sources of the field, are acted own by the field but are not a part of the field. Their characteristic properties are inertia, measured by specific constant, the mass" [6]. The unitarian theory developed by Thomson, Lorentz and Mie tends to describe the particle as a point-like singularity but with finite mass-energy density fulfilling uniquely an appropriate nonlinear field equations. It is interesting that this approach was later developed in the classical book by Einstein and Infeld [126] as well as in the classical paper by Einstein, Infeld and Hoffmann [127] on equations of motion in General Relativity.

In the Born-Infeld approach the emphasis is directed to a formalism encompassing General Relativity. But for simplicity the field equations are solved within the realm only of the electromagnetic field. A basic tensor $a_{\alpha\beta} = g_{\alpha\beta} + f_{\alpha\beta}$ is introduced. Its symmetric part $g_{\alpha\beta}$ is identified with a metric component and the antisymmetric part $f_{\alpha\beta}$ with the electromagnetic field. Formally therefore both the electromagnetic and gravitational fields are present although the authors explicitly avoided to insert the part of the Lagrangian describing the gravitational interaction and focused uniquely on the following nonlinear Lagrangian

$$\mathcal{L} = \sqrt{1 + 2S - P^2} - 1. \quad (4.2.1)$$

The necessity to have the quadratic form of the $P$ term is due to obtain a Lagrangian invariant under reflections as pointed out by W. Pauli in his classical book [121]. For small field strengths Lagrangian (4.2.1) has the same form as (4.1.3) obtained by Euler.

From the nonlinear Lagrangian (4.2.1) Born and Infeld calculated the fields $D$ and $H$ through a tensor, $P^i_0 = D^i$ and $P^{ij} = -\epsilon^{ijk}H^k$, where

$$P_{\mu\nu} = \frac{\delta\mathcal{L}_{\text{Born}}}{\delta F^{\mu\nu}} = \frac{F_{\mu\nu} - Pf_{\mu\nu}}{\sqrt{1 + 2S - P^2}}, \quad (4.2.2)$$

and introduced therefore an effective electric permittivity and magnetic permeability which are functions of $S$ and $P$. It is very interesting that Born and Infeld managed to obtain a solution for electrostatic field of a point particle ($P = 0$) in which the radial component $D_r = e/r^2$ becomes infinite as $r \to 0$ but the radial component of $E$ field is perfectly finite and is given by the expression

$$E_r = \frac{e}{r_0^2 \sqrt{1 + (r/r_0)^4}}, \quad (4.2.3)$$

where $r_0$ is the “radius” of the electron.

Most important the integral of the electromagnetic energy is finite and
4 Nonlinear electrodynamics and rate of pair creation

\[
\int \mathcal{H}_{\text{Born}} dV = \int (P_{\mu\nu} F^{\mu\nu} - \mathcal{L}_{\text{Born}}) dV = 1.2361 \frac{e^2}{r_0}, \tag{4.2.4}
\]

Equating this energy to \( m_e c^2 \) they obtain \( r_0 = 1.2361 e^2 / (m_e c^2) \).

The attempt therefore is to have a theoretical framework explaining the mass of the electron solely by a modified nonlinear electromagnetic field theory. This approach has not been followed by the current theories in particle physics where the dualistic approach is today adopted in which the charged particles are described by half-integer spin fields and electro-magnetic interactions by integer-spin fields.

The initial goal to develop a fully covariant theory of electrodynamics within General Relativity although not developed by Born himself was not abandoned. Hoffmann found an analytic solution [22] to the coupled system of the Einstein-Born-Infeld equations.

4.3 The Euler-Heisenberg Lagrangian

The two different approaches of Born and Infeld and of Euler present strong analogies and substantial differences. The attempt of Born and Infeld was to obtain at once a new nonlinear Lagrangian for electromagnetic field replacing the Maxwell Lagrangian in order to avoid the appearance of infinite self-energy for a classical point-like electron.

The attempt of Euler [128] was more conservative, to obtain the first order nonlinear perturbation corrections to the Maxwell Lagrangian on the ground of the Dirac theory of the electron.

Born and Infeld in addition introduced an effective dielectric constant and an effective magnetic permeability of the vacuum out of their nonlinear Lagrangian (4.2.1). This approach was adopted as well in the classical work of Heisenberg and Euler [7]. They introduced an effective Lagrangian on the ground of the Dirac theory of the electron and expressed the result in integral form duly taking away infinities, see Section 4.3.1. This integral was explicitly performed in the weak field limit and the special attention was given to the real part, see Section 4.3.2 and the imaginary part, see Section 4.3.3.

A successive work of Weisskopf [24] derived the same equations of Heisenberg and Euler for the real part of the dielectric constant and magnetic permeability by using instead of the spin 1/2 particle of the Dirac equation the scalar relativistic wave equation of Klein and Gordon. The results differ from the one of spin 1/2 particle only by a factor 2 due to the Bose statistics, see Section 4.3.1. The technique used by Weisskopf refers to the case of magnetic field of arbitrary strengths and describes the electric field perturbatively, see Section 4.3.2. As we will see in the following, the Heisenberg and Euler integral can be found in the case of arbitrary large both electric and magnetic fields.
4.3 The Euler-Heisenberg Lagrangian

fields, see Section 5.7.3.

4.3.1 Real part of the effective Lagrangian

We now recall how Heisenberg and Euler adopted the crucial idea of Max Born to describe the nonlinear Lagrangian by the introduction of an effective dielectric constant and magnetic permeability [7]. They further extended this idea by adopting the most general case of a dielectric constant containing real and imaginary part. Such an approach is generally followed in the description of dissipative media. The crucial point was to relate electron–positron pair creation process to imaginary part of the Lagrangian.

Let \( \mathcal{L} \) to be the Lagrangian density of electromagnetic fields \( E, B \), a Legendre transformation produces the Hamiltonian density:

\[
\mathcal{H} = E_i \frac{\delta \mathcal{L}}{\delta E^*_i} - \mathcal{L}.
\]  

(4.3.1)

In Maxwell’s theory, the two densities are given by

\[
\mathcal{L}_M = \frac{1}{8\pi} (E^2 - B^2), \quad \mathcal{H}_M = \frac{1}{8\pi} (E^2 + B^2).
\]  

(4.3.2)

To quantitatively describe nonlinear electromagnetic properties of the vacuum based on the Dirac theory, Heisenberg and Euler introduced an effective Lagrangian \( \mathcal{L}_{\text{eff}} \) of the vacuum state and an associated Hamiltonian density

\[
\mathcal{L}_{\text{eff}} = \mathcal{L}_M + \Delta \mathcal{L}, \quad \mathcal{H}_{\text{eff}} = \mathcal{H}_M + \Delta \mathcal{H}.
\]  

(4.3.3)

Here \( \mathcal{H}_{\text{eff}} \) and \( \mathcal{L}_{\text{eff}} \) are complex functions of \( E \) and \( B \). In Maxwell’s theory, \( \Delta \mathcal{L} \equiv 0 \) in the vacuum, so that \( D = E \) and \( H = B \).

Heisenberg and Euler derived the induced fields \( D, H \) as the derivatives

\[
D_i \equiv \frac{\delta \mathcal{L}_{\text{eff}}}{\delta E^*_i}, \quad H_i \equiv -\frac{\delta \mathcal{L}_{\text{eff}}}{\delta B^*_i}.
\]  

(4.3.4)

Consequently, the vacuum behaves as a dielectric and permeable medium [7,24] in which,

\[
D_i = \epsilon_{ik} E_k, \quad H_i = \mu_{ik} B_k,
\]  

(4.3.5)

where \( \epsilon_{ik} \) and \( \mu_{ik} \) are complex and field-dependent dielectric and permeability tensors of the vacuum.

The discussions on complex dielectric and permeability tensors (\( \epsilon_{ik} \) and \( \mu_{ik} \)) can be found for example in Ref. [129]. The effective Lagrangian and Hamiltonian densities in such a medium are given by

\[
\mathcal{L}_{\text{eff}} = \frac{1}{8\pi} (E \cdot D - B \cdot H), \quad \mathcal{H}_{\text{eff}} = \frac{1}{8\pi} (E \cdot D + B \cdot H).
\]  

(4.3.6)
In this medium, the conservation of electromagnetic energy has the form

\[-\text{div}\mathbf{S} = \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{B} \cdot \frac{\partial \mathbf{H}}{\partial t}, \quad \mathbf{S} = c \mathbf{E} \times \mathbf{B}, \quad (4.3.7)\]

where \(\mathbf{S}\) is the Poynting vector describing the density of electromagnetic energy flux. Consider complex and monochromatic electromagnetic field

\[\mathbf{E} = \mathbf{E}(\omega) \exp(-i\omega t); \quad \mathbf{B} = \mathbf{B}(\omega) \exp(-i\omega t), \quad (4.3.8)\]

of frequency \(\omega\), and dielectric and permeability tensors are frequency-dependent, i.e., \(\epsilon_{ij}(\omega)\) and \(\mu_{ij}(\omega)\). Substituting these fields and tensors into the right-hand side of Eq. (4.3.7), one obtains the dissipation of electromagnetic energy per unit time into the medium

\[Q_{\text{dis}} = \frac{\omega}{2} \left\{ \text{Im} [\epsilon_{ij}(\omega)] E_i E_j^* + \text{Im} [\mu_{ij}(\omega)] B_i B_j^* \right\}. \quad (4.3.9)\]

This is nonzero if \(\epsilon_{ij}(\omega)\) and \(\mu_{ij}(\omega)\) contain an imaginary part. The dissipation of electromagnetic energy in a medium is accompanied by heat production. In the light of the third thermodynamical law of entropy increase, the energy lost \(Q_{\text{dis}}\) of electromagnetic fields in the medium is always positive, i.e., \(Q_{\text{dis}} > 0\). As a consequence, \(\text{Im}[\epsilon_{ij}(\omega)] > 0\) and \(\text{Im}[\mu_{ij}(\omega)] > 0\).

The real parts of \(\epsilon_{ij}(\omega)\) and \(\mu_{ij}(\omega)\) represent an electric and magnetic polarizability of the vacuum and leads, for example, to the refraction of light in an electromagnetic field, or to the elastic scattering of light from light. The reflection index of the medium

\[n_{ij}(\omega) = \sqrt{\epsilon_{ij}(\omega) \mu_{ij}(\omega)} \]

is the reflection index of the medium. The field dependence of \(\epsilon_{ij}\) and \(\mu_{ij}\) implies nonlinear electromagnetic properties of the vacuum as a dielectric and permeable medium.

The effective Lagrangian density (4.3.3) is a relativistically invariant function of the field strengths \(\mathbf{E}\) and \(\mathbf{B}\). Since \((\mathbf{E}^2 - \mathbf{B}^2)\) and \(\mathbf{E} \cdot \mathbf{B}\) are relativistic invariants, one can formally expand \(\Delta \mathcal{L}\) in powers of weak field strengths:

\[\Delta \mathcal{L} = \kappa_{2,0}(\mathbf{E}^2 - \mathbf{B}^2)^2 + \kappa_{0,2}(\mathbf{E} \cdot \mathbf{B})^2 + \kappa_{3,0}(\mathbf{E}^2 - \mathbf{B}^2)^3 + \kappa_{1,2}(\mathbf{E}^2 - \mathbf{B}^2)(\mathbf{E} \cdot \mathbf{B})^2 + \ldots, \quad (4.3.10)\]

where \(\kappa_{i,j}\) are field-independent constants whose subscripts indicate the powers of \((\mathbf{E}^2 - \mathbf{B}^2)\) and \(\mathbf{E} \cdot \mathbf{B}\), respectively. Note that the invariant \(\mathbf{E} \cdot \mathbf{B}\) appears only in even powers since it is odd under parity and electromagnetism is parity invariant. The Lagrangian density (4.3.10) corresponds, via relation (4.3.1), to

\[\Delta \mathcal{H} = \kappa_{2,0}(\mathbf{E}^2 - \mathbf{B}^2)(3\mathbf{E}^2 + \mathbf{B}^2) + \kappa_{0,2}(\mathbf{E} \cdot \mathbf{B})^2
+ \kappa_{3,0}(\mathbf{E}^2 - \mathbf{B}^2)^2(5\mathbf{E}^2 + \mathbf{B}^2) + \kappa_{1,2}(3\mathbf{E}^2 - \mathbf{B}^2)(\mathbf{E} \cdot \mathbf{B})^2 + \ldots. \quad (4.3.11)\]
To obtain $\mathcal{H}_{\text{eff}}$ in Dirac’s theory, one has to calculate

$$
\Delta \mathcal{H} = \sum_{k} \left\{ \psi_k^* \left[ \alpha \cdot (-i\hbar \nabla + eA) + \beta Dm_e e^2 \right] \psi_k \right\},
$$

(4.3.12)

where $\{ \psi_k(x) \}$ are the wave functions of the occupied negative energy states. When performing the sum, one encounters infinities which were removed by Dirac, Heisenberg, and Weisskopf [24, 130–132] by a suitable subtraction.

Heisenberg [131] expressed the Hamiltonian density in terms of the density matrix $\rho(x, x') = \sum_k \psi_k^*(x) \psi_k(x')$ [130]. Heisenberg and Euler [7] calculated the coefficients $\kappa_{ij}$. They did so by solving the Dirac equation in the presence of parallel electric and magnetic fields $E$ and $B$ in a specific direction,

$$
\psi_k(x) \to \psi_{p+x, s} \equiv \psi_k^* (z^2 \epsilon - \bar{\epsilon} t) u_n(y) \chi_s(x), \quad n = 0, 1, 2, \ldots
$$

(4.3.13)

where $\{ u_n(y) \}$ are the Landau states\(^1\) depending on the magnetic field and $\chi_s(x)$ are the spinor functions calculated by Sauter [20]. Heisenberg and Euler used the Euler-Maclaurin formula to perform the sum over $n$ and obtained for the additional Lagrangian in (4.3.3) the integral representation

$$
\Delta \mathcal{L}_{\text{eff}} = \frac{e^2}{16\pi^2 \hbar c} \int_0^\infty \frac{e^{-s}}{s^3} \left[ is^2 \bar{E} \bar{B} \cos(s[\bar{E}^2 - \bar{B}^2 + 2i(\bar{E}\bar{B})]^{1/2}) + \text{c.c.} \right]
$$

$$
+ \left( \frac{m^2 c^3}{\epsilon \hbar} \right)^2 + \frac{s^2}{3} (|\bar{B}|^2 - |\bar{E}|^2),
$$

(4.3.14)

where $\bar{E}, \bar{B}$ are the dimensionless reduced fields in the unit of the critical field $E_c$,

$$
\bar{E} = \frac{|E|}{E_c}, \quad \bar{B} = \frac{|B|}{E_c}.
$$

(4.3.15)

Expanding this expression in powers of $\alpha$ up to $\alpha^3$ yields the following values for the four constants:

$$
\kappa_{2,0} = \frac{\alpha}{360\pi^2} E^{-2}, \quad \kappa_{0,2} = 7\kappa_{2,0}, \quad \kappa_{3,0} = \frac{2\alpha}{315\pi^2} E^{-4}, \quad \kappa_{1,2} = \frac{13}{2} \kappa_{3,0}.
$$

(4.3.16)

The above results will receive higher corrections in QED and are correct only up to order $\alpha^2$. Up to this order, the field-dependent dielectric and permeability tensors $\epsilon_{ik}$ and $\mu_{ik}$ (4.3.5) have the following real parts for weak fields

$$
\text{Re}(\epsilon_{ik}) = \delta_{ik} + \frac{\alpha}{180\pi^2} \left[ 2(\bar{E}^2 - \bar{B}^2) \delta_{ik} + 7\bar{B}_i \bar{B}_k \right] + O(\alpha^2),
$$

$$
\text{Re}(\mu_{ik}) = \delta_{ik} + \frac{\alpha}{180\pi^2} \left[ 2(\bar{E}^2 - \bar{B}^2) \delta_{ik} + 7\bar{E}_i \bar{E}_k \right] + O(\alpha^2).
$$

(4.3.17)

\(^1\)Landau determined the quantum states of a particle in an external magnetic field in 1930 [88,90].
4.3.2 Weisskopf effective Lagrangian

Weisskopf [24] adopted a simpler method. He considered first the special case in which $E = 0, B \neq 0$ and used the Landau states to find $\Delta \mathcal{H}$ of Eq. (4.3.11), extracting from this $\kappa_{2,0}$ and $\kappa_{3,0}$. Then he added a weak electric field $E \neq 0$ to calculate perturbatively its contributions to $\Delta \mathcal{H}$ in the Born approximation (see for example [88]). This led again to the coefficients (4.3.16), (4.3.17). In addition to the weak field expansion of real part of effective Lagrangian, Weisskopf also obtained the leading order term considering very large field strengths $\bar{E} \gg 1$ or $\bar{B} \gg 1$,

$$\Delta L_{\text{eff}} \sim -\frac{e^2}{12\pi^2\hbar c} E^2 \ln \bar{E}; \quad \Delta L_{\text{eff}} \sim \frac{e^2}{12\pi^2\hbar c} \bar{B}^2 \ln \bar{B}, \quad (4.3.18)$$

We shall address this same problem in Section 5.7.3 in the framework of QED [35] and we will compare and contrast our exact expressions with the one given by Weisskopf. The crucial point stressed by Weisskopf is that if one limits to the analysis of the real part of the dielectric constant and magnetic permeability then the nonlinearity of effective electromagnetic Lagrangian represent only small corrections even for field strengths which are much higher than the critical field strength $E_c$. As we will show however, the contribution of the imaginary part of the effective Lagrangian diverges as pointed out by Heisenberg and Euler [7].

4.3.3 Imaginary part of the effective Lagrangian

Heisenberg and Euler [7] were the first to realize that for $E \neq 0$ the powers series expansion (4.3.10) is not convergent, due to singularities of the integrand in (4.3.14) at $s = \pi/\bar{E}, 2\pi/\bar{E}, \ldots$. They concluded that the powers series expansion (4.3.10) does not yield all corrections to the Maxwell Lagrangian, calling for a more careful evaluation of the integral representation (4.3.14). Selecting an integration path that avoids these singularities, they found an imaginary term. Motivated by Sauter’s work [20] on Klein paradox [17, 18], Heisenberg and Euler estimated the size of the imaginary term in the effective Lagrangian as

$$\text{Im} L_{\text{eff}} = -\frac{8}{\pi} E^2 m_e c^2 \left( \frac{m_e c}{\hbar} \right)^3 e^{-\pi/\bar{E}}, \quad (4.3.19)$$

and pointed out that it is associated with pair production by the electric field. The exponential in this expression is exactly reproducing the Sauter result (3.4.9). However, for the first time the pre-exponential factor is determined. This imaginary term in the effective Lagrangian is related to the imaginary parts of field-dependent dielectric $\epsilon$ and permeability $\mu$ of the vacuum.

In 1950’s, Schwinger [25–27] derived the same formula (4.3.14) within the Quantum Electro-Dynamics (QED). In the following sections, our discussions
and computations will focus on the Schwinger formula, the real and imaginary parts of effective Lagrangian for arbitrary values of electromagnetic field strength.

The consideration of Heisenberg and Euler were applied to a uniform electric field. The exponential factor $e^{-\pi/\bar{E}}$ in Eqs. (3.4.9) and (4.3.19) characterizes the transmission coefficient of quantum tunneling, Heisenberg and Euler [7] introduced the critical field strength (4.3.15). They compared it with the field strength $E_e$ of an electron at its classical radius, $E_e = e/r_e^2$ where $r_e = \alpha \hbar / (m_e c)$. They found the field strength $E_e$ is 137 time larger than the critical field strength $E_c$, i.e. $E_e = \alpha^{-1} E_c$. At a critical radius $r_c = \alpha^{1/2} \hbar / (m_e c) < r_e$, the field strength of the electron would be equal to the critical field strength $E_c$. There have been various attempts to reach the critical field: in Secs. 7.5 and 7.6 we will examine the possibility of reaching such value around the bare nucleus. In Section 8.5 we will discuss the possibility of reaching such a field in an astrophysical setting around a black hole.

In conclusion, if an electric field attempts to tear an electron out of the filled state the gap energy must be gained over the distance of two electron radii. The virtual particles give an electron a radius of the order of the Compton wavelength $\lambda_C$. Thus we expect a significant creation of electron–positron pairs if the work done by the electric field $E$ over twice the Compton wave length $\hbar / m_e c$ is larger than $2m_e c^2$

$$eE \left(\frac{2\hbar}{m_e c}\right) > 2m_e c^2.$$  

This condition defines a critical electric field (2.0.1) above which pair creation becomes abundant. To have an idea how large this critical electric field is, we compare it with the value of the electric field required to ionize a hydrogen atom. There the above inequality holds for twice of the Bohr radius and the Rydberg energy

$$eE_{\text{ion}} \left(\frac{2\hbar}{a m_e c}\right) > \alpha^2 m_e c^2,$$

where $E_{\text{ion}} = m_e^2 e^5 / \hbar^4 = 5.14 \times 10^9$ V/cm, so that $E_c = E_{\text{ion}} / \alpha^3$ is about $10^6$ times as large, a value that has so far not been reached in a laboratory on Earth.
5 Pair production and annihilation in QED

5.1 Quantum Electro-Dynamics

Quantum Electro-Dynamics (QED), the quantum theory of electrons, positrons, and photons, was established by by Tomonaga [8], Feynman [9–11], Schwinger [12–14] and Dyson [15, 16] and others in the 1940’s and 1950’s [133]. For decades, both theoretical computations and experimental tests have been developed to great perfection. It is now one of the fundamental pillars of the theory of the microscopic world. Many excellent monographs have been written [25–27, 89, 90, 134–143], so the concepts of the theory and the techniques of calculation are well explained. On the basis of this material, we review some aspects and properties of the QED that are relevant to the subject of the present review.

QED combines a relativistic extension of quantum mechanics with a quantized electromagnetic field. The nonrelativistic system has a unique ground state, which is the state with no particle, the vacuum state. The excited states contain a fixed number of electrons and an arbitrary number of photons. As electrons are allowed to become relativistic, their number becomes also arbitrary, and it is possible to create pairs of electrons and positrons.

In the modern functional integral description, the nonrelativistic system is described by a given set of fluctuating particle orbits running forward in time. If the theory is continued to an imaginary time, in which case one speaks of a Euclidean formulation, the nonrelativistic system corresponds to a canonical statistical ensemble of trajectories.

In the relativistic system, the orbits form worldlines in four-dimensional space-time which may run in any time direction, in particular they may run backwards in time, in which case the backward parts of a line correspond to positrons. The number of lines is arbitrary and the Euclidean formulation corresponds to a grand-canonical ensemble. The most efficient way of describing such an ensemble is by a single fluctuating field [143].

The vacuum state contains no physical particles. It does, however, harbor zero-point oscillations of the electron and photon fields. In the worldline description, the vacuum is represented by a grand-canonical ensemble of interacting closed world lines. These are called virtual particles. Thus the vacuum contains the full complexity of a many-body problem so that one may right-
fully say that the vacuum is the world [144]. In the Fourier decomposition of
the fluctuating fields, virtual particles correspond to Fourier components, or
modes, in which the 4-vectors of energy and momentum $k^\mu \equiv (k^0, \mathbf{k}) \equiv (\mathcal{E}, \mathbf{k})$
do not satisfy the mass-shell relation

$$k^2 \equiv (k^0)^2 - c^2|\mathbf{k}|^2 = \mathcal{E}^2 - c^2|\mathbf{k}|^2 = m^2 c^4,$$

valid for real particles.

The only way to evaluate physical consequences from QED is based on the
smallness of the electromagnetic interaction. It is characterized by the dimen-
sionless fine structure constant $\alpha$. All theoretical results derived from QED
are found in the form of series expansions in powers of $\alpha$, which are expan-
sions around the non-interacting system. Unfortunately, all these expansions
are badly divergent (see e.g. Section 4.62 in [145]). The number of terms con-
tributing to the same order of $\alpha$ grows factorially fast, i.e., faster than any
exponential, leading to a zero radius of convergence. Fortunately, however,
the coupling $\alpha$ is so small that the series possess an apparent convergence
up to order $1/\alpha \approx 137$, which is much higher than will be calculable for a
long time to come (see e.g. Section 4.62 in [145]). With this rather academic
limitation, perturbation expansions are well defined.

In perturbation expansions, all physical processes are expressible in terms
of Feynman diagrams. These are graphic representations of the interacting
world lines of all particles. Among these lines, there are some which run to
infinity. They satisfy the mass shell relation (5.1.1) and describe real particles
observable in the laboratory. Those which remain inside a finite space-time
region are virtual.

The presence of virtual particles in the perturbation expansions leads to
observable effects. Some of these have been measured and calculated with
great accuracy. The most famous examples are

1. the electrostatic polarizability of quantum fluctuations of the QED vac-
uum has been measured in the Lamb shift [146, 147].

2. the anomalous magnetic moment of the electron [11, 115–118].

3. the dependence of the electric charge on the distance. It is observed
by measuring cross-sections of electron–positron collisions, most re-
cently in the L3-experiments at the Large Electron-Positron Collider (LEP)
at CERN [148].

4. the Casimir effect caused by virtual photons, i.e., by the fluctuations of
the electromagnetic field in the QED vacuum [149, 150]. It causes an
attractive force [151–153] between two uncharged conducting plates in
the vacuum (see also [154–160]).
5.2 Basic processes in Quantum Electro-Dynamics

There are, of course, many other discussions of the effects of virtual particles caused either by external boundary conditions or by external classical fields [161–175].

An interesting aspect of virtual particles both theoretically and experimentally is the possibility that they can become real by the effect of external fields. In this case, real particles are excited out of the vacuum. In the previous Section 3.4 and 4.3.1, we have shown that this possibility was first pointed out in the framework of quantum mechanics by Klein, Sauter, Euler and Heisenberg [7, 17, 18, 20] who studied the behavior of the Dirac vacuum in a strong external electric field. Afterward, Schwinger studied this process and derived the probability (Schwinger formula) in the field theory of Quantum Electro-Dynamics, which will be described in this chapter. If the field is sufficiently strong, the energy of the vacuum can be lowered by creating an electron–positron pair. This makes the vacuum unstable. This is the Sauter-Euler-Heisenberg-Schwinger process for electron–positron pair production. There are many reasons for the interest in the phenomenon of pair production in a strong electric field. The most compelling one is that now both laboratory conditions and astrophysical events provide possibilities for observing this process.

In the following chapters, in addition to reviewing the Schwinger formula and QED-effective Lagrangian in constant electromagnetic fields, we will also derive the probability of pair production in an alternating field, and discuss theoretical studies of pair production in (i) electron-beam–laser collisions and (ii) superstrong Coulomb potential. In addition, the plasma oscillations of electron–positron pairs in electric fields will be reviewed in Section 9. The rest part of this chapter, we shall use natural units \( \hbar = c = 1 \).

5.2 Basic processes in Quantum Electro-Dynamics

The total Lagrangian describing the interacting system of photons, electrons, and positrons reads, see e.g. [90]

\[
\mathcal{L} = \mathcal{L}_0^\gamma + \mathcal{L}_0^{e^+e^-} + \mathcal{L}_{\text{int}},
\]

(5.2.1)

where the free Lagrangians \( \mathcal{L}_0^{e^+e^-} \) and \( \mathcal{L}_0^\gamma \) for electrons and photons are expressed in terms of quantized Dirac field \( \psi(x) \) and quantized electromagnetic field \( A_\mu(x) \) as follows:

\[
\mathcal{L}_0^{e^+e^-} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m_e)\psi(x),
\]

(5.2.2)

\[
\mathcal{L}_0^\gamma = -\frac{1}{4} F_{\mu\nu}(x) F^{\mu\nu}(x) + \text{gauge-fixing term}.
\]

(5.2.3)
Here $\gamma^\mu$ are the $4 \times 4$ Dirac matrices, $\bar{\psi}(x) \equiv \psi^\dagger(x)\gamma^0$, and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ denotes the electromagnetic field tensor. Minimal coupling gives rise to the interaction Lagrangian

$$\mathcal{L}_{\text{int}} = -ie^{\mu}(x)A_\mu(x), \quad j^{\mu}(x) = \bar{\psi}(x)\gamma^\mu\psi(x). \quad (5.2.4)$$

After quantization, the photon field is expanded into plane waves as

$$A_\mu(x) = \int \frac{d^3k}{2k_0(2\pi)^3} \sum_{\lambda=1}^{3} \left[ a^{(\lambda)}(k)e_{\mu}^{(\lambda)}(k)e^{-ikx} + a^{(\lambda)^*}(k)e_{\mu}^{(\lambda)^*}(k)e^{ikx} \right],$$

where $e_{\mu}^{(\lambda)}$ are polarization vectors, and $a^{(\lambda)}, a^{(\lambda)^*}$ are annihilation and creation operators of photons. The quantized fermion field $\psi(x)$ has the expansion

$$\psi(x) = \int \frac{d^3k}{(2\pi)^3} \frac{m}{k_0} \sum_{\alpha=1,2} \left[ b_\alpha(k,s_3)u^{(\alpha)}(k,s_3)e^{-ikx} + d_\alpha^\dagger(k,s_3)v^{(\alpha)}(k,s_3)e^{ikx} \right], \quad (5.2.6)$$

where the four-component spinors $u^{(\alpha)}(k,s_3), v^{(\alpha)}(k,s_3)$ are positive and negative energy solutions of the Dirac equation with momentum $k$ and spin component $s_3$. The operators $b(k,s_3), b^\dagger(k,s_3)$ annihilate electrons and create photons, the operators $d(k,s_3)$ and $d^\dagger(k,s_3)$ do the same for positrons [90].

In the framework of QED the transition probability from an initial to a final state for a given process is represented by the imaginary part of the unitary S-matrix squared

$$\mathcal{P}_{f \leftarrow i} = \left| \langle f, \text{out} | \text{Im} S | i, \text{in} \rangle \right|^2, \quad (5.2.7)$$

where

$$\text{Im} S = (2\pi)^4 \delta^4(P_f - P_i) |M_{fi}|, \quad (5.2.8)$$

$M_{fi}$ is called matrix element and $\delta$-function stays for energy-momentum conservation in the process.

When initial state contains two particles with energies $\epsilon_1$ and $\epsilon_2$, and final state contain arbitrary number of particles having 3-momenta $p'_i$, the transition probability per unit time and unit volume is given by

$$\frac{d\mathcal{P}_{f \leftarrow i}}{dVdt} = (2\pi)^4 \delta^4(P_f - P_i) |M_{fi}|^2 \frac{1}{4\epsilon_1\epsilon_2} \prod_i \frac{d^3p'_i}{(2\pi)^3 2\epsilon_i}. \quad (5.2.9)$$

The Lorentz invariant differential cross-section for a given process is then obtained from (5.2.9) by dividing it on the flux density of initial particles

$$d\sigma = (2\pi)^4 \delta^4(P_f - P_i) |M_{fi}|^2 \frac{1}{4I_{\text{kin}}} \prod_i \frac{d^3p'_i}{(2\pi)^3 2\epsilon_i}, \quad (5.2.10)$$

968
5.2 Basic processes in Quantum Electro-Dynamics

where \( p_1 \) and \( p_2 \) are particles’ 4-momenta, \( m_1 \) and \( m_2 \) are their masses respectively, \( I_{\text{kin}} = \sqrt{(p_1p_2)^2 - m_1^2m_2^2} \).

It is useful to work with Mandelstam variables which are kinematic invariants built from particles 4-momenta. Consider the process \( A + B \rightarrow C + D \). Lorentz invariant variables can be constructed in the following way

\[
\begin{align*}
  s &= (p_A + p_B)^2 = (p_C + p_D)^2, \\
  t &= (p_A + p_C)^2 = (p_B + p_D)^2, \\
  u &= (p_B + p_C)^2 = (p_A + p_D)^2.
\end{align*}
\]

(5.2.11)

Since any incoming particle can be regarded as outgoing antiparticle, it gives rise to the crossing symmetry property of the scattering amplitude, which is best reflected in the Mandelstam variables. In fact, reactions \( A + B \rightarrow C + D, A + \bar{C} \rightarrow \bar{B} + D \) or \( A + \bar{D} \rightarrow C + \bar{B} \) where the bar denotes the antiparticle are just different cross-channels of a single general reaction. The meaning of the variables \( s, t, u \) changes, but the amplitude is the same.

The S-matrix is computed through the interaction operator as

\[
S = T \exp \left( i \int L_{\text{int}} d^4x \right),
\]

(5.2.12)

where \( T \) is the chronological operator. Perturbation theory is applied, since the fine structure constant is small, while any additional interaction in collision of particles contains the factor \( \alpha \).

A simple and elegant way of computation of the S-matrix and consequently of the matrix element \( M_{fi} \) is due to Feynman, who discovered a graphical way to depict each QED process, in momentum representation.

In what follows we consider briefly the calculation for the case of Compton scattering process [90], which is given by two Feynman diagrams. Conservation law for 4-momenta is \( p + k = p' + k' \), where \( p \) and \( k \) are 4-momenta of electron and photon respectively, and invariant \( I_{\text{kin}} = \frac{1}{4}(s - m_e^2)^2 \). After the calculation of traces with gamma-matrices, the final result, expressed in Mandelstam variables, is

\[
|M_{fi}|^2 = 2^7 \pi^2 e^4 \left[ \frac{m_e^2}{s - m_e^2} + \frac{m_e^2}{u - m_e^2} + \left( \frac{m_e^2}{s - m_e^2} + \frac{m_e^2}{u - m_e^2} \right)^2 \right. \\
\left. - \frac{1}{4} \left( \frac{s - m_e^2}{u - m_e^2} + \frac{u - m_e^2}{s - m_e^2} \right) \right],
\]

(5.2.13)

\( s = (p + k)^2, t = (p - p')^2 \) and \( u = (p - k')^2 \). Since the differential cross-section is independent of the azimuth of \( p_1' \) relative to \( p_1 \), it is obtained from
(5.2.13) as
\[ d\sigma = \frac{1}{64\pi} |M_{fi}|^2 \frac{dt}{I_{kin}^2}. \]  
\[ (5.2.14) \]

In the laboratory frame, where \( s - m_e^2 = 2m_e\omega, u - m_e^2 = -2m_e\omega' \) and electron is at rest before the collision with photon, the differential cross-section of Compton scattering is thus given by the Klein–Nishina formula \[176\]
\[ d\sigma = \frac{1}{2} \left( \frac{e^2}{m} \right)^2 \left( \frac{\omega'}{\omega} \right)^2 \left( \frac{\omega}{\omega'} + \frac{\omega'}{\omega} - \sin^2 \theta \right), \]
\[ (5.2.15) \]
where \( \omega \) and \( \omega' \) are frequencies of photon before and after the collision, \( \theta \) is the angle at which the photon is scattered.

### 5.3 The Dirac and the Breit–Wheeler processes in QED

We turn now to the formulas obtained within framework of quantum mechanics by Dirac [1] and Breit and Wheeler [2] within QED. The crossing symmetry allows to readily write the matrix element for the pair production (3.0.1) and pair annihilation (3.0.2) processes with the energy-momentum conservation written as \( p_+ + p_- = k_1 + k_2 \), where \( p_+ \) and \( p_- \) are 4-momenta of the positron and the electron, \( k_1 \) and \( k_2 \) are 4-momenta of two photons. It is in fact given by the same formula (5.2.13) with the substitution \( p \rightarrow p_- \), \( p' \rightarrow p_+ \), \( k \rightarrow k_1 \), \( k' \rightarrow k_2 \), but with different meaning of the kinematic invariants \( s = (p_- - k_1)^2 \), \( t = (p_- + p_+)^2 \), \( u = (p_- - k_2)^2 \). Matrix elements for Dirac and Breit–Wheeler processes are the same. The differential cross-section of the Dirac process is obtained from (5.2.13) with the exchange \( s \leftrightarrow t \) and the invariant \( I_{kin} = \frac{1}{4}t(t - 4m_e^2) \), which leads to (3.1.9). For the case of the Breit–Wheeler process with the invariant \( I_{kin} = \frac{1}{4}t^2 \), the result is reduced to (3.2.9).

Since the Dirac pair annihilation process (3.0.1) is the inverse of Breit–Wheeler pair production (3.0.2), it is useful to compare the cross-section of the two processes. We note that the squared transition amplitude \( |M_{fi}|^2 \) must be the same for two processes, due to the CPT invariance. The cross-sections could be different only by kinematics and statistical factors. Let us consider the pair annihilation process in the center of mass system where \( \mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 = \mathcal{E}_1' + \mathcal{E}_2' \) is the total energy, the initial and final momenta are equal and opposite, \( \mathbf{p}_1 = -\mathbf{p}_2 \equiv \mathbf{p} \) and \( \mathbf{p}_1' = -\mathbf{p}_2' \equiv \mathbf{p}' \). The differential cross-section is given by (5.2.14). For the Breit and Wheeler process (3.0.2) of two colliding photons with 4-momenta \( k_1 \) and \( k_2 \), the scalar \( I_{\gamma\gamma}^2 = (k_1k_2)^2 \). For the Dirac process (3.0.1) of colliding electron and positron with 4-momenta \( p_1 \)
and $p_2$, the scalar $I_{e^+e^-}^2 = (p_1p_2)^2 - m_e^4$. As results, one has

$$\frac{d\sigma_{\gamma\gamma}}{d\sigma_{e^+e^-}} = \frac{I_{e^+e^-}^2}{I_{\gamma\gamma}^2} = \frac{2(k_1k_2) - 4m_e^2}{2(k_1k_2)} = \frac{\hat{p}^2 - 2m_e^2}{E^2} = \left(\frac{|\hat{p}|}{E}\right)^2 = \hat{\beta}^2,$$

(5.3.1)

where momenta and energies are related by

$$(p_1 + p_2)^2 = (k_1 + k_2)^2 = 2(k_1k_2) = 2\hat{E}^2.$$

Integrating Eq. (5.3.1) over all scattering angles yields the total cross-section. Whereas the previous $\sigma_{\gamma\gamma}$ required division by a Bose factor 2 for the two identical photons in the final state, the cross-section $\sigma_{e^+e^-}$ has no such factor since the final electron and positron are not identical. Hence we obtain

$$\sigma_{e^+e^-} = \frac{1}{2\hat{\beta}^2} \sigma_{\gamma\gamma}.$$  (5.3.2)

By re-expressing the kinematic quantities in the laboratory frame, one obtains the Dirac cross-section (3.1.9).

As shown in Eq. (5.3.2) in the center of mass of the system, the two cross-sections $\sigma_{e^+e^-}$ and $\sigma_{\gamma\gamma}$ of the above described phenomena differ only in the kinematics and statistical factor $1/(2\hat{\beta}^2)$, which is related to the fact that the resulting particles are massless or massive. The process of electron and positron production by the collision of two photons has a kinematic energy threshold, while the process of electron and positron annihilation to two photons has not such kinematic energy threshold. In the limit of high energy neglecting the masses of the electron and positron, $\hat{\beta} \to 1$, the difference between two cross-sections $\sigma_{e^+e^-}$ and $\sigma_{\gamma\gamma}$ is only the statistical factor 1/2.

The total cross-sections (3.1.11) of Breit–Wheeler’s and Dirac’s process are of the same order of magnitude $\sim 10^{-25}\text{cm}^2$ and have the same energy dependence $1/\hat{E}^2$ above the energy threshold. The energy threshold $(2m_e\hat{c}^2)$ have made until now technically impossible to observe the pair production by the Breit–Wheeler process in laboratory experiments at the intersection of two beams of X-rays. Another reason is of course the smallness of the total cross-section (3.2.10) $(\sigma_{\gamma\gamma} \lesssim 10^{-25}\text{cm}^2)$ and the experimental limitations on the intensities $I_i$ (3.2.3) of the light beams. We shall see however, that this Breit–Wheeler process occurs routinely in the dyadosphere of a black hole. The observations of such phenomena in the astrophysical setting are likely to give the first direct observational test of the validity of the Breit–Wheeler process, see e.g. [177].
5.4 Double-pair production

Following the Breit–Wheeler pioneer work on the process (3.0.2), Cheng and Wu [178–183] considered the high-energy behavior of scattering amplitudes and cross-section of two photon collision, up to higher order $O(\alpha^4)$ [184],

$$\gamma_1 + \gamma_2 \rightarrow e^+ + e^- + e^+ + e^-.$$  \hspace{1cm} (5.4.1)

For this purpose, they calculated the two photon forward scattering amplitude $M_{\gamma\gamma}$ (see Eq. (5.2.9)) by taking into account all relevant Feynman diagrams via two electron loops up to the order $O(\alpha^4)$. The total cross-section $\sigma_{\gamma\gamma}$ for photon-photon scattering is related to the photon-photon scattering amplitude $M_{\gamma\gamma}$ in the forward direction by the optical theorem,

$$\sigma_{\gamma\gamma}(s) = \frac{1}{s} \text{Im} M_{\gamma\gamma},$$  \hspace{1cm} (5.4.2)

where $s$ is the square of the total energy in the center of mass system. They obtained the total cross-section of double pair production (5.4.1) at high energy $s \gg 2m_e$,

$$\lim_{s \rightarrow \infty} \sigma_{\gamma\gamma}(s) = \frac{\alpha^4}{36\pi m_e^2} [175\zeta(3) - 38] \sim 6.5\mu b,$$  \hspace{1cm} (5.4.3)

which is independent of $s$ as well as helicities of the incoming photons. Up to the $\alpha^4$ order, Eq. (5.4.3) is the largest term in the total cross-section for photon-photon scattering at very high energy. This can be seen by comparing Eq. (5.4.3) with the cross-section (3.2.9,3.2.10) of the Breit–Wheeler process (3.0.2), which is the lowest-order process in a photon-photon collision and vanishes as $s \rightarrow \infty$. Thus, although the Breit–Wheeler cross-section is of lower order in $\alpha$, the Cheng-Wu cross-section (5.4.3) is larger as the energy becomes sufficiently high. In particular, the Cheng-Wu cross-section (5.4.3) exceeds the Breit–Wheeler one (3.2.9) as the center of mass energy of the photon $E \geq 0.24\text{GeV}$. Note that in the double pair production (5.4.1), the energy threshold is $E \geq 2m_e$ rather than $E \geq m_e$ in the one pair production of Breit–Wheeler.

In Ref. [185, 186], using the same method Cheng and Wu further calculated other cross-sections for high-energy photon-photon scattering to double pion, muon and electron–positron pairs:

- the process of double muon pair production $\gamma_1 + \gamma_2 \rightarrow \mu^+ + \mu^- + \mu^+ + \mu^-$ and its cross-section can be obtained by replacing $m_e \rightarrow m_\mu$ in Eq. (5.4.3), thus $\sigma_{\gamma\gamma}(s) \sim 1.5 \cdot 10^{-4}\mu b$.

- the process of double pion pair production $\gamma_1 + \gamma_2 \rightarrow \pi^+ + \pi^- + \pi^+ + \pi^-$

\vspace{1cm}

972
\[ \pi^- \] and its extremely small cross-section
\[
\lim_{s \to \infty} \sigma_{\gamma\gamma}(s) = \frac{\alpha^4}{144 \pi m^2_\pi} [7\zeta(3) + 10] \sim 0.23 \cdot 10^{-5} \mu b, \quad (5.4.4)
\]

- the process of one pion pair and one electron–positron pair production
\[
\gamma_1 + \gamma_2 \to e^+ + e^- + \pi^+ + \pi^- \]
and its small cross-section
\[
\lim_{s \to \infty} \sigma_{\gamma\gamma}(s) = \frac{2\alpha^4}{27 \pi m^2_\pi} \left[ \left( \ln \frac{m^2_\pi}{m^2_e} \right)^2 + \frac{16}{3} \ln \frac{m^2_\pi}{m^2_e} + \frac{163}{18} \right] \sim 0.26 \cdot 10^{-3} \mu b, \quad (5.4.5)
\]
which is more than one hundred times larger than (5.4.4).

- the process of one pion pair and one muon-antimuon pair production
\[
\gamma_1 + \gamma_2 \to \mu^+ + \mu^- + \pi^+ + \pi^-
\]
and its small cross-section can be obtained by replacing \(m_e \to m_\mu\) in Eq. (5.4.5) everywhere.

### 5.5 Electron-nucleus bremsstrahlung and pair production by a photon in the field of a nucleus

The other two important QED processes, related by the crossing symmetry are the electron-nucleus bremsstrahlung (3.3.7) and creation of electron–positron pair by a photon in the field of a nucleus (3.3.1). These processes were considered already in the early years of QED. They are of higher order, comparing to the Compton scattering and the Breit–Wheeler processes, respectively, and contain one more vertex connecting fermions with the photon.

The nonrelativistic cross-section for the process (3.3.7) was derived by Sommerfeld [103]. Here we remind the basic results obtained in the relativistic case by Bethe and Heitler [96] and independently by Sauter [97]. The Feynman diagram for bremsstrahlung can be imagined considering for the Compton scattering, but treating one of the photons as virtual one corresponding to an external field. We consider this process in Born approximation, and the momentum recoil of the nucleus is neglected. Integrating the differential cross-section over all directions of the photon and the outgoing electron one has, see e.g. [90]

\[
d\sigma = Z^2 \alpha r^2 \frac{d\omega}{\omega} \frac{d\nu}{\nu} \left\{ \frac{4}{3} - 2\epsilon \epsilon' \frac{\nu^2 + \nu'^2}{\nu^2 \nu'^2} + m_e^2 \left( \frac{\nu'}{\nu^2} + \frac{\nu}{\nu'^2} \right) \right\} + \\
+ L \left[ \frac{8\epsilon \epsilon'}{5 p p'} \left( e^2 e'^2 + \frac{\nu^2 - \nu'^2}{\nu^2 \nu'^2} \right) + \frac{m_e^2}{2 p p'} \left( \epsilon^2 + \epsilon'^2 \right) \right], \quad (5.5.1)
\]
where
\[ L = \log \frac{ee' + pp' - m_e^2}{ee' - pp' - m_e^2}, \quad l = \log \frac{e + p}{e - p}, \quad l' = \log \frac{e' + p'}{e' - p'}, \] (5.5.2)
and \( \omega \) is photon energy, \( p \) and \( p' \) are electron momenta before and after the collision, respectively, \( \epsilon \) and \( \epsilon' \) are its initial and final energies.

The averaged cross-section for the process of pair production by a photon in the field of a nucleus may be obtained by applying the transformation rules relating the processes (3.3.1) and (3.3.7), see e.g. [90]. The result is
\[
d\sigma = Z^2 r_e^2 \frac{p_+ p_-}{\omega_0} d\omega_0 \left\{ -\frac{4}{3} - 2\epsilon_+ \epsilon_- p_+^2 p_-^2 + m_e^2 \left( l_+ \frac{e_+}{p_+} + l_- \frac{e_-}{p_-} \right) + L \left[ \frac{8\epsilon_+ \epsilon_-}{3p_+ p_-} + \frac{\omega_0}{p_+^2 p_-^2} \left( \epsilon_+^2 \epsilon_-^2 + p_+^2 p_-^2 - m_e^2 \epsilon_+ \epsilon_- \right) - \frac{m_e^2 \omega_0}{2p_+ p_-} \left( l_+ \frac{e_+ - e_-}{p_+} + l_- \frac{e_- - e_+}{p_-} \right) \right] \right\},
\] (5.5.3)
where
\[ L = \log \frac{e_+ e_- + p_+ p_- + m_e^2}{e_+ e_- - p_+ p_- + m_e^2}, \quad l_\pm = \log \frac{e_\pm + p_\pm}{e_\pm - p_\pm}, \] (5.5.4)
and \( p_\pm \) and \( e_\pm \) are momenta and energies of positron and electron respectively.

The total cross-section for this process is given in Section 3.3, see (3.3.3). In the ultrarelativistic approximation relaxing the condition \( Z\alpha \ll 1 \) the pair production by the process (3.3.1) was treated by Bethe and Maximon in [187, 188]. The total cross-section (3.3.3) is becoming then
\[
\sigma = \frac{28}{9} Z^2 r_e^2 \left( \log \frac{2\omega}{m_e} - \frac{109}{42} - f(Z\alpha) \right),
\] (5.5.5)
where
\[
f(Z\alpha) = \gamma E + \Re \Psi (1 + iZ\alpha) = (Z\alpha)^2 \sum_{n=1}^{\infty} \frac{1}{n \left[ n^2 + (Z\alpha)^2 \right]},
\] (5.5.6)

5.6 Pair production in collision of two ions

The process of the pair production by two ions (3.3.4) in ultrarelativistic approximation was considered by Landau and Lifshitz [99] and Racah [100]. For the modern review of these topics see [189]. The corresponding differential cross-section with logarithmic accuracy can be obtained from the differential cross-section (5.5.3) taking its ultrarelativistic approximation \( \gamma \gg 1 \) for the
5.6 Pair production in collision of two ions

Lorentz factor of the relative motion of the two nuclei with charges $Z_1$ and $Z_2$ respectively, and treating the real photon line in the process (3.3.1) as a virtual photon corresponding to the external field of the nucleus. One should then multiply the cross-section by the spectrum of these equivalent photons, see Section 3.3, and the result is

$$d\sigma = \frac{8}{\pi} r_c^2 (Z_1 Z_2 \alpha)^2 \frac{d\epsilon_+ d\epsilon_-}{(\epsilon_+ + \epsilon_-)^4} \left( \epsilon_+^2 + \epsilon_-^2 + \frac{2}{3} \epsilon_+ \epsilon_- \right) \log \frac{\epsilon_+ \epsilon_-}{m_e (\epsilon_+ + \epsilon_-)} \log \frac{m_e \gamma}{\epsilon_+ + \epsilon_-}. \quad (5.6.1)$$

The total cross-section is given in Section 3.3, see (3.3.5) and (3.3.6). More recent results containing higher order in $\alpha$ corrections are obtained in [190, 191], see also [192, 193].

Figure 5.1: Classification of the $e^+e^-$ pair production by the number of photons attached to a nucleus. Reproduced from [189].

Lepton pair production in relativistic ion collision to all orders in $Z \alpha$ with logarithmic accuracy is studied in [194] where the matrix elements are separated in different classes, see Fig. 5.1, according to numbers of photon lines attached to a given nucleus

$$M = M_{(i)} + M_{(ii)} + M_{(iii)} + M_{(iv)}. \quad (5.6.2)$$

The Born amplitude $|M_{(i)}|^2$ corresponding to the lowest order in $Z \alpha$ with one photon line attached to each nucleus was computed by Landau and Lifshitz [99] who obtained the famous $L_3^2 \gamma$ dependence of the cross-section in (3.3.5). It should be mentioned that the Racah formula (3.3.6) in contrast with the Landau and Lifshitz result (3.3.5) contains also terms proportional to $L_2^2 \gamma$ and $L_1 \gamma$. These terms come from the absolute square of $M_{(ii)}$ and $M_{(iii)}$ and their interference with the Born amplitude $M_{(i)}$ [191]. Their result for the Coulomb corrections which is defined as the difference between the full cross-section and the Born approximation, in order $L_2^2 \gamma$ is of the “Bethe–Maximon” type [191]

$$\sigma_C = \frac{28}{27 \pi} r_c^2 (Z_1 Z_2 \alpha)^2 \left[ f(Z_1 \alpha) f(Z_2 \alpha) \right] \left( L_2^2 \gamma + O(L_1 \gamma) \right). \quad (5.6.3)$$
The Coulomb corrections here are up to $L_γ^2$-term in the Racah formula (3.3.6).

The calculations mentioned above were all made as early as in the 1930s. Clearly, at that time only $e^+e^-$ pair production was discussed. However, these calculations can be considered for any lepton pair production, for example $\mu^+\mu^-$ pair production, as long as the total energy in the center of mass of the system is large enough. However, simple substitution $m_e \to m_l$, where $l$ stands for any lepton is not sufficient since when the energy reaches the inverse radius $\Lambda = 1/R \sim 30\text{MeV}$ of the nucleus the electric field of the nucleus cannot be approximated as a Coulomb field of a point-like particle [195]. The review of computation for lepton pair production can be found in [196] and [101].

Another effect of large enough collision energy is multiple pair production. Early work on this subject started with the observation that the impact-parameter-dependent total pair production probability computed in the lowest order perturbation theory is larger than one. The analysis of Ref. [190] devoted to the study of the corresponding Feynman diagrams in the high-energy limit leads to the probability of $N$ lepton pair production obeying the Poisson distribution. For a review of this topic see [189].

Two photon particle pair production by collision of two electrons or electron and positron (see Fig. 3.1, where 4-momenta $p_1$ and $p_2$ correspond to their momenta) were studied in storage rings in Novosibirsk ($e^+e^- \to e^+e^-e^+e^-$ [197]) and in Frascati at ADONE ($e^+e^- \to e^+e^-e^+e^-$ [198], $e^+e^- \to e^+e^-\mu^+\mu^-$ [199], $e^+e^- \to e^+e^-\pi^+\pi^-$ [200]), see also [201]. At high energy the total cross-section of two photon production of lepton pairs is given by $[99, 101]$

$$\sigma_{e^e\to l^l} = \frac{28\alpha^4}{27\pi m_l^2} \left( \ln \frac{s}{m_e^2} \right)^2 \ln \frac{s}{m_e^2} ; \ l \equiv e \ or \ \mu. \quad (5.6.4)$$

see c.f. Eq. (3.3.5).

### 5.7 QED description of pair production

We turn now to a Sauter-Heisenberg-Euler process in QED. An external electromagnetic field is incorporated by adding to the quantum field $A_\mu$ in (5.2.4) an unquantized external vector potential $A_\mu^e$, so that the total interaction becomes

$$\mathcal{L}_{\text{int}} + \mathcal{L}_{\text{int}}^e = -e\bar{\psi}(x)\gamma^\mu\psi(x) \left[ A_\mu(x) + A_\mu^e(x) \right]. \quad (5.7.1)$$

Instead of an operator formulation, one can derive the quantum field theory from a functional integral formulation, see e.g. [202], in which the quantum
mechanical partition function is described by

\[ Z[A^e] = \int [D\psi D\bar{\psi} D A_\mu] \exp \left[ i \int d^4x (\mathcal{L} + \mathcal{L}^e_{\text{int}}) \right], \tag{5.7.2} \]

to be integrated over all fluctuating electromagnetic and Grassmannian electron fields. The normalized quantity \( Z[A^e] \) gives the amplitude for the vacuum to vacuum transition in the presence of the external classical electromagnetic field:

\[ \langle \text{out}, 0 | 0, \text{in} \rangle = \frac{Z[A^e]}{Z[0]}, \tag{5.7.3} \]

where \( |0, \text{in}\rangle \) is the initial vacuum state at the time \( t = t_- \to -\infty \), and \( \langle \text{out}, 0 | \) is the final vacuum state at the time \( t = t_+ \to +\infty \). By selecting only the one-particle irreducible Feynman diagrams in the perturbation expansion of \( Z[A^e] \) one obtains the effective action as a functional of \( A^e \):

\[ \Delta A_{\text{eff}}[A^e] \equiv -i \ln \langle \text{out}, 0 | 0, \text{in} \rangle. \tag{5.7.4} \]

In general, there exists no local effective Lagrangian density \( \Delta \mathcal{L}_{\text{eff}} \) whose space-time integral is \( \Delta A_{\text{eff}}[A^e] \). An infinite set of derivatives would be needed, i.e., \( \Delta \mathcal{L}_{\text{eff}} \) would have the arguments \( A^e(x), \partial_\mu A^e(x), \partial_\mu \partial_\nu A^e(x), \ldots \), containing gradients of arbitrary high order. With presently available methods it is possible to calculate a few terms in such a gradient expansion, or a semiclassical approximation à la JWKB for an arbitrary but smooth space-time dependence (see Section 3.21ff in Ref. [202]). Under the assumption that the external field \( A^e(x) \) varies smoothly over a finite space-time region, we may define an approximately local effective Lagrangian \( \Delta \mathcal{L}_{\text{eff}}[A^e(x)] \),

\[ \Delta A_{\text{eff}}[A^e] \approx \int d^4x \Delta \mathcal{L}_{\text{eff}}[A^e(x)] \approx V \Delta t \Delta \mathcal{L}_{\text{eff}}[A^e], \tag{5.7.5} \]

where \( V \) is the spatial volume and time interval \( \Delta t = t_+ - t_- \).

For a large time interval \( \Delta t = t_+ - t_- \to \infty \), the amplitude of the vacuum to vacuum transition (5.7.3) has the form,

\[ \langle \text{out}, 0 | 0, \text{in} \rangle = e^{-i(\Delta \varepsilon_0 - i\Gamma/2)\Delta t}, \tag{5.7.6} \]

where \( \Delta \varepsilon_0 = \varepsilon_0(A^e) - \varepsilon_0(0) \) is the difference between the vacuum energies in the presence and the absence of the external field, \( \Gamma \) is the vacuum decay rate, and \( \Delta t \) the time over which the field is nonzero. The probability that the vacuum remains as it is in the presence of the external classical electromagnetic field is

\[ |\langle \text{out}, 0 | 0, \text{in} \rangle|^2 = e^{-2\text{Im}\Delta A_{\text{eff}}[A^e]}. \tag{5.7.7} \]

This determines the decay rate of the vacuum in an external electromagnetic
5 Pair production and annihilation in QED

\[
\frac{\Gamma}{V} = \frac{2 \text{Im} \Delta A_{\text{eff}}[A^e]}{V \Delta t} \approx 2 \text{Im} \Delta \mathcal{L}_{\text{eff}}[A^e]. \tag{5.7.8}
\]

The vacuum decay is caused by the production of electron and positron pairs. The external field changes the energy density by

\[
\frac{\Delta \varepsilon_0}{V} = -\frac{\text{Re} \Delta A_{\text{eff}}[A^e]}{V \Delta t} \approx -\text{Re} \Delta \mathcal{L}_{\text{eff}}[A^e]. \tag{5.7.9}
\]

### 5.7.1 Schwinger formula for pair production in uniform electric field

The Dirac fields appears quadratically in the partition functional (5.7.2) and can be integrated out, leading to

\[
Z[A^e] = \int DA \他知道的 \text{Det} \{i \partial - e[A^e(x)] - m_e + i \eta \}, \quad \partial \equiv \gamma^\mu \partial_\mu, \quad A \equiv \gamma^\mu A_\mu, \tag{5.7.10}
\]

where \text{Det} denotes the functional determinant of the Dirac operator. Ignoring the fluctuations of the electromagnetic field, the result is a functional of the external vector potential \(A^e(x)\):

\[
Z[A^e] \approx \text{const} \times \text{Det} \{i \partial - eA^e(x) - m_e + i \eta \}. \tag{5.7.11}
\]

The infinitesimal constant \(i \eta\) with \(\eta > 0\) specifies the treatment of singularities in energy integrals. From Eqs. (5.7.3)–(5.7.11), the effective action (5.7.7) is given by

\[
\Delta A_{\text{eff}}[A^e] = -i \text{Tr} \ln \left\{ \left[ i \partial - eA^e(x) - m_e + i \eta \right] \frac{1}{i \partial - m_e + i \eta} \right\}, \tag{5.7.12}
\]

where \text{Tr} denotes the functional and Dirac trace. In physical units, this is of order \(\hbar\). The result may be expressed as a one-loop Feynman diagram, so that one speaks of one-loop approximation. More convenient will be the equivalent expression

\[
\Delta A_{\text{eff}}[A^e] = -i \frac{1}{2} \text{Tr} \ln \left( \left[ i \partial - eA^e(x) \right]^2 - m_e^2 + i \eta \right) \frac{1}{-\partial^2 - m_e^2 + i \eta}, \tag{5.7.13}
\]

where

\[
[i \partial - eA^e(x)]^2 = [i \partial_\mu - eA^e_\mu(x)]^2 + \frac{e}{2} \sigma^{\mu \nu} F^e_{\mu \nu}, \tag{5.7.14}
\]

where \(\sigma^{\mu \nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu]\), \(F^e_{\mu \nu} = \partial_\mu A^e_\nu - \partial_\nu A^e_\mu\). Using the identity

\[
\ln \frac{a_2}{a_1} = \int_0^\infty \frac{ds}{s} \left[ e^{is(a_1 + i \eta)} - e^{is(a_2 + i \eta)} \right], \tag{5.7.15}
\]
Eq. (5.7.13) becomes
\[
\Delta A_{\text{eff}}[A^e] = \frac{i}{2} \int_0^\infty \frac{ds}{s} e^{-is(m_e^2 - i\eta)} \text{Tr} \langle x | e^{is \left\{ [i\partial_\mu - e A^e_\mu(x)]^2 + \frac{1}{2} \sigma^{\mu\nu} F^e_{\mu\nu} \right\}} - e^{-is\partial^2} | x \rangle,
\]
where \( \langle x | \cdot \cdot \cdot | x \rangle \) are the diagonal matrix elements in the local basis \( |x\rangle \). This is defined by the matrix elements with the momentum eigenstates \( |k\rangle \) being plane waves: \( \langle x | k \rangle = e^{-ikx} \). The symbol \( \text{Tr} \) denotes integral \( \int d^4x \) in space-time and the trace in spinor space. For constant electromagnetic fields, the integrand in (5.7.16) does not depend on \( x \), and \( \sigma^{\mu\nu} F^e_{\mu\nu} \) commutes with all other operators. This will allows us to calculate the exponential in Eq. (5.7.16) explicitly. The presence of \( -i\eta \) in the mass term ensures convergence of integral for \( s \to \infty \).

If only a constant electric field \( E \) is present, it may be assumed to point along the \( \hat{z} \)-axis, and one can choose a gauge such that \( A^e_z = -Et \) is the only nonzero component of \( A^e_\mu \). Then one finds
\[
\text{tr} \exp is \left[ \frac{e}{2} \sigma^{\mu\nu} F^e_{\mu\nu} \right] = 4 \cosh(seeE),
\]
where the symbol \( \text{tr} \) denotes the trace in spinor space. Using the commutation relation \([\partial_0, x^0] = 1\), where \( x^0 = t \), one computes the exponential term in the effective action (5.7.16) (c.e.g. [89])
\[
\langle x | \exp is \left[ (i\partial_\mu - e A^e_\mu(x))^2 + \frac{e}{2} \sigma^{\mu\nu} F^e_{\mu\nu} \right] | x \rangle = \frac{eE}{(2\pi)^2 is} \coth(eEs). \tag{5.7.18}
\]
The second term in Eq. (5.7.16) is obtained by setting \( E = 0 \) in Eq. (5.7.18), so that the effective action (5.7.16) yields,
\[
\Delta A_{\text{eff}} = \frac{1}{2(2\pi)^2} \int d^4x \int_0^\infty \frac{ds}{s^3} [eEs \coth(eEs) - 1] e^{-is(m_e^2 - i\eta)}. \tag{5.7.19}
\]
Since the field is constant, the integral over \( x \) gives a volume factor, and the effective action (5.7.16) can be attributed to the space-time integral over an effective Lagrangian (5.7.5)
\[
\Delta L_{\text{eff}} = \frac{1}{2(2\pi)^2} \int_0^\infty \frac{ds}{s^3} [eEs \coth(eEs) - 1] e^{-is(m_e^2 - i\eta)}.
\tag{5.7.20}
\]
By expanding the integrand in Eq. (5.7.20) in powers of \( e \), one obtains,
\[
\frac{1}{s^3} [eEs \coth(eEs) - 1] e^{-is(m_e^2 - i\eta)} = \left[ \frac{e^2}{3s} E^2 - \frac{e^4 s}{45} E^4 + O(e^6) \right] e^{-is(m_e^2 - i\eta)}. \tag{5.7.21}
\]
The small-s divergence in the integrand,
\[ \frac{e^2}{3} E^2 \frac{1}{2(2\pi)^2} \int_0^\infty \frac{ds}{s} e^{-is(m_2^2 - i\eta)}, \]  
(5.7.22)
is proportional to the electric term in the original Maxwell Lagrangian. The divergent term (5.7.22) can therefore be removed by a renormalization of the field $E$. Thus we subtract a counterterm in Eq. (5.7.20) and form
\[ \Delta \mathcal{L}_{\text{eff}} = \frac{1}{2(2\pi)^2} \int_0^\infty \frac{ds}{s^3} \left[ eEs \coth(eEs) - 1 - \frac{e^2}{3} E^2 s^2 \right] e^{-is(m_2^2 - i\eta)} . \]  
(5.7.23)

Remembering Eq. (5.7.8), we find from (5.7.23) the decay rate of the vacuum per unit volume
\[ \frac{\Gamma}{V} = \frac{1}{(2\pi)^2} \text{Im} \int_0^\infty \frac{ds}{s^3} \left[ eEs \coth(eEs) - 1 - \frac{e^2}{3} E^2 s^2 \right] e^{-is(m_2^2 - i\eta)} . \]  
(5.7.24)
The integral (5.7.24) can be evaluated analytically by the method of residues. Since the integrand is even, the integral can be extended to the entire $s$-axis. After this, the integration contour is deformed to enclose the negative imaginary axis and to pick up the contributions of the poles of the coth function at $s = n\pi/eE$. The result is
\[ \frac{\Gamma}{V} = \alpha \frac{E^2}{\pi^2} \sum_{n=1}^\infty \frac{1}{n^2} \exp \left( -\frac{n\pi E}{E} \right) . \]  
(5.7.25)
This result, i.e. the Schwinger formula [25–27] is valid to lowest order in $\hbar$ for arbitrary constant electric field strength.

An analogous calculation for a charged scalar field yields
\[ \frac{\Gamma}{V} = \alpha \frac{E^2}{2\pi^2} \sum_{n=1}^\infty \frac{(-1)^{n+1}}{n^2} \exp \left( -\frac{n\pi E}{E} \right) , \]  
(5.7.26)
which generalizes the Weisskopf treatment being restricted to the leading term $n = 1$. These Schwinger results complete the derivation of the probability for pair productions. The leading $n = 1$ -terms of (5.7.25) and (5.7.26) agrees with the JWKB results we discuss in Section 3.5, and thus the correct Sauter exponential factor (3.4.9) and Heisenberg-Euler imaginary part of the effective Lagrangian (4.3.19).

Narozhny and Nikishov [33] have expressed Eq. (5.7.7) through the probability of one pair production $P_1$, of $n$ pair production $P_n$ with $n = 1, 2, 3, \cdots$ as well as the average number of pair productions
\[ |\langle \text{out}, 0 | 0, \text{in} \rangle|^2 = 1 - P_1 - P_2 - P_3 - \cdots , \]  
(5.7.27)
where \( P_n (n = 1, 2, 3, \cdots) \) is the probability of \( n \) pair production, and the probability of one pair production is,

\[
P_1 = V \Delta t \frac{\alpha E^2}{2\pi^2} \ln \left( 1 - e^{-\frac{\alpha n^2}{2\pi}} \right) e^{-2V \Delta t \Im \Delta \mathcal{L}_{\text{eff}}[A^e]},
\]

(5.7.28)

The average number \( \bar{N} \) of pair productions is then given by

\[
\bar{N} = \sum_{n=1}^{\infty} n P_n = V \Delta t \frac{\alpha E^2}{\pi^2} \exp \left( -\frac{\pi E_e}{E} \right),
\]

(5.7.29)

which is the quantity directly related to the experimental measurements.

### 5.7.2 Pair production in constant electromagnetic fields

Since the QED theory is gauge and Lorentz invariant, effective action \( \Delta A_{\text{eff}} \) and Lagrangian \( \Delta \mathcal{L}_{\text{eff}} \) are expressed as functionals of the scalar and pseudoscalar invariants \( S, P \) (3.5.23). Thus they must be invariant under the discrete duality transformation:

\[
|B| \rightarrow -i|E|, \quad |E| \rightarrow i|B|
\]

(5.7.30)
i.e.,

\[
\beta \rightarrow -i\varepsilon, \quad \varepsilon \rightarrow i\beta.
\]

(5.7.31)

This implies that in the case \( E = 0 \) and \( B \neq 0 \), results can be simply obtained by replacing \( |E| \rightarrow i|B| \) in Eqs. (5.7.18), (5.7.23), (5.7.24):

\[
\langle x | \exp is \left[ (i\partial_\mu - eA^e_\mu(x))^2 + \frac{e}{2} \varepsilon^{\mu\nu} F^e_{\mu\nu} \right] | x \rangle = \frac{eB}{(2\pi)^2 is} \cot(eBs),
\]

(5.7.32)

and

\[
\Delta \mathcal{L}_{\text{eff}} = \frac{1}{2(2\pi)^2} \int_0^{\infty} \frac{ds}{s^3} \left[ eBs \cot(eBs) - 1 + \frac{e^2}{3} B^2 s^2 \right] e^{-is(m_2^2 - i\eta)}. \]

(5.7.33)

In the presence of both constant electric and magnetic fields \( E \) and \( B \), we adopt parallel \( E_{\text{CF}} \) and \( B_{\text{CF}} \) pointing along the \( \hat{z} \)-axis in the *center-of-fields frame*, as discussed after Eqs. (3.5.14), (3.5.15), (3.5.16). We can choose a gauge such that only \( A^e_z = -E_{\text{CF}} t, A^e_y = B_{\text{CF}} x_1 \) are nonzero. Due to constant fields, the exponential in the effective action Eq. (5.7.16) can be factorized into a product of the magnetic part and the electric part. Following the same method used to compute the electric part (5.7.17, 5.7.18), one can compute the magnetic part by using the commutation relation \([\partial_1, x^1] = 1\), where \( x^1 = x \). Or one can make the substitution (5.7.30) to obtain the magnetic part, based
on the discrete symmetry of duality. As results, Eqs. (5.7.17), (5.7.18) become
\[
\text{tr} \exp is \left[ \frac{e}{2} \sigma_{\mu
u} F_{\mu\nu} \right] = 4 \cosh(seE_{CF}) \cos(seB_{CF}),
\]
(5.7.34)
and
\[
\langle x | \exp is \left\{ [i \partial_{\mu} - e A_{\mu}^{\nu} (x)]^2 + \frac{e}{2} \sigma_{\mu\nu} F_{\mu\nu} \right\} | x \rangle = \frac{1}{(2\pi)^2} \cdot \frac{eE_{CF}}{is} \cot(seE_{CF}) \frac{eB_{CF}}{s} \cot(seB_{CF}).
\]
(5.7.35)

In this special frame, the effective Lagrangian is then given by
\[
\Delta \mathcal{L}_{\text{eff}} = \frac{1}{2(2\pi)^2} \int_{s_0}^{\infty} ds \left[ e^2 e B_{CF} s^2 \coth(seE_{CF}) \cot(seB_{CF}) 
- 1 - \frac{e^2}{3} (E_{CF}^2 - B_{CF}^2) s^2 \right] \cdot e^{-is(m^2 - i\eta)}.
\]
(5.7.36)

Using definitions in Eqs. (3.5.23), (3.5.24), (3.5.25), we obtain the effective Lagrangian
\[
\Delta \mathcal{L}_{\text{eff}} = \frac{1}{2(2\pi)^2} \int_{s_0}^{\infty} ds \left[ e^2 \epsilon \beta s^2 \coth(\epsilon s) \cot(e\beta s) 
- 1 - \frac{e^2}{3} (\epsilon^2 - \beta^2) s^2 \right] \cdot e^{-is(m^2 - i\eta)};
\]
(5.7.37)
and the decay rate
\[
\frac{\Gamma}{V} = \frac{1}{(2\pi)^2} \text{Im} \int_{s_0}^{\infty} ds \left[ e^2 \epsilon \beta s^2 \coth(\epsilon s) \cot(e\beta s) 
- 1 - \frac{e^2}{3} (\epsilon^2 - \beta^2) s^2 \right] e^{-is(m^2 - i\eta)},
\]
(5.7.38)
in terms of the invariants \(\epsilon\) and \(\beta\) (3.5.25) for arbitrary electromagnetic fields \(E\) and \(B\).

The integral (5.7.38) is evaluated as in Eq. (5.7.25) by the method of residues, and yields [25–27]
\[
\frac{\Gamma}{V} = \frac{\alpha e^2}{\pi^2} \sum_{n=1}^{\infty} \frac{n \pi \beta / \epsilon}{n^2 \tanh n \pi \beta / \epsilon} \exp \left( - \frac{n \pi E_{\epsilon}}{\epsilon} \right),
\]
(5.7.39)
which reduces for \(\beta \to 0\) (\(B = 0\)) correctly to (5.7.25). The \(n = 1\) -term is the JWKB approximation (3.5.27).
The analogous result for bosonic fields is
\[
\frac{\Gamma}{V} = \frac{\alpha e^2}{2\pi^2} \sum_{n=1}^\infty \frac{(-1)^n}{n^2} \frac{n\pi\beta/\varepsilon}{\sinh n\pi\beta/\varepsilon} \exp \left( -\frac{n\pi E_c}{\varepsilon} \right), \tag{5.7.40}
\]
where the first term \( n = 1 \) corresponds to the Euler-Heisenberg result (4.3.19). Note that the magnetic field produces in the fermionic case an extra factor \((n\pi\beta/\varepsilon) / \tanh(n\pi\beta/\varepsilon) > 1\) in each term which enhances the decay rate. The bosonic series (5.7.40), on the other hand, carries in each term a suppression factor \((n\pi\beta/\varepsilon) / \sinh n\pi\beta/\varepsilon < 1\). The average number \(\bar{N} \) (5.7.29) is given by
\[
\bar{N} = \sum_{n=1}^{\infty} nP_n = V\Delta t \frac{\alpha}{\pi} \frac{\alpha\beta\varepsilon}{\tanh \pi\beta/\varepsilon} \exp \left( -\frac{\pi E_c}{\varepsilon} \right). \tag{5.7.41}
\]

The decay rate \(\Gamma/V\) gives the number of electron–positron pairs produced per unit volume and time. The prefactor can be estimated on dimensional grounds. It has the dimension of \(E_c^2/\hbar\), i.e., \(m^4c^5/\hbar^4\). This arises from the energy of a pair \(2m_ec^2\) divided by the volume whose diameter is the Compton wavelength \(\hbar/m_ec\), produced within a Compton time \(\hbar/m_ec^2\). The exponential factor suppresses pair production as long as the electric field is much smaller than the critical electric field \(E_c\), in which case the JWKB results (3.5.27) and (3.5.28) are good approximations.

The general results (5.7.39),(5.7.40) was first obtained by Schwinger [25–27] for scalar and spinor electrodynamics (see also Nikishov [28], Batalin and Fradkin [29]). The method was extended to special space-time-dependent fields in Refs. [30–34]. The monographs [89, 145, 203–205] can be consulted about more detailed calculation, discussion and bibliography.

### 5.7.3 Effective nonlinear Lagrangian for arbitrary constant electromagnetic field

Starting from the integral form of Heisenberg and Euler Lagrangian (5.7.37) we find explicitly real and imaginary parts of the effective Lagrangian \(\Delta\mathcal{L}_{\text{eff}}\) (5.7.37) for arbitrary constant electromagnetic fields \(\mathbf{E}\) and \(\mathbf{B}\) [35]. The essential step is to reach a direct analytic form resulting from performing the integration. We use the expressions [206],
\[
es\cos\th(ees) = \sum_{n=-\infty}^{\infty} \frac{s^2}{(s^2 + \tau_n^2)}, \quad \tau_n \equiv n\pi/ee, \tag{5.7.42}
\]
\[
e\beta s \cot(e\beta s) = \sum_{m=-\infty}^{\infty} \frac{s^2}{(s^2 - \tau_m^2)}, \quad \tau_m \equiv m\pi/e\beta, \tag{5.7.43}
\]
and obtain for the finite effective Lagrangian of Heisenberg and Euler integral representation,

$$\Delta L_{\text{eff}} = \frac{1}{2(2\pi)^2} \sum_{n, m = -\infty}^{\infty} \int_0^\infty ds \frac{s}{\tau_n^2 + \tau_m^2} \left[ \frac{\delta_{n0}}{(s^2 - \tau_n^2)} - \frac{\delta_{n0}}{(s^2 + \tau_n^2)} \right] e^{-is(m_c^2 - i\eta)},$$

(5.7.44)

where divergent terms $n \neq 0, m = 0, n = 0, m \neq 0$ and $n = m = 0$ are excluded from the sum, as indicated by a prime. The symbol $\delta_{ij} \equiv 1 - \delta_{ij}$ denotes the complimentary Kronecker-$\delta$ which vanishes for $i = j$. The divergent term with $n = m = 0$ is eliminated by the zero-field subtraction in Eq. (5.7.37), while the divergent terms $n \neq 0, m = 0$ and $n = 0, m \neq 0$

$$\Delta L_{\text{eff}}^{\text{div}} = \frac{1}{2(2\pi)^2} \int_0^\infty ds \frac{s}{\tau_n^2 + \tau_m^2} \left( \sum_{n = 1}^{\infty} \frac{1}{\tau_m^2} - \sum_{m = 1}^{\infty} \frac{1}{\tau_n^2} \right),$$

(5.7.45)

are eliminated by the second subtraction in Eq. (5.7.37). This can be seen by performing the sums

$$\sum_{n = 1}^{\infty} \frac{1}{\tau_n^2} = \left( \frac{\epsilon e^\eta}{\pi} \right)^k \zeta(k); \quad \sum_{n = 1}^{\infty} \frac{1}{\tau_m^2} = \left( \frac{\epsilon \beta}{\pi} \right)^k \zeta(k),$$

(5.7.46)

where $\zeta(k) = \sum_n 1/n^k$ is the Riemann function.

The infinitesimal $i\eta$ accompanying the mass term in the $s$-integral (5.7.44) is equivalent to replacing $e^{-is(m_c^2 - i\eta)}$ by $e^{-is(1-i\eta)m_c^2}$. This implies that $s$ is to be integrated slightly below (above) the real axis for $s > 0$ ($s < 0$). Equivalently one may shift the $\tau_m$ ($-\tau_m$) variables slightly upwards (downwards) to $\tau_m + i\eta$ ($-\tau_m - i\eta$) in the complex plane.

In order to calculate the finite effective Lagrangian (5.7.44), the factor $e^{-is(1-i\eta)m_c^2}$ is divided into its sin and cos parts:

$$\Delta L_{\text{eff}}^{\sin} = -\frac{i}{4(2\pi)^2} \sum_{n, m = -\infty}^{\infty} \int_{-\infty}^{\infty} ds \frac{s}{\tau_n^2 + \tau_m^2} \left[ \frac{\delta_{m0}}{(s^2 - \tau_n^2)} - \frac{\delta_{m0}}{(s^2 + \tau_n^2)} \right] \sin[s(1 - i\eta)m_c^2];$$

(5.7.47)

$$\Delta L_{\text{eff}}^{\cos} = \frac{1}{2(2\pi)^2} \sum_{n, m = -\infty}^{\infty} \int_0^\infty ds \frac{s}{\tau_n^2 + \tau_m^2} \left[ \frac{\delta_{n0}}{(s^2 - \tau_n^2)} - \frac{\delta_{n0}}{(s^2 + \tau_n^2)} \right] \cos[s(1 - i\eta)m_c^2].$$

(5.7.48)

The sin part (5.7.47) has an even integrand allowing for an extension of the $s$-integral over the entire $s$-axis. The contours of integration can then be closed by infinite semicircles in the half plane, the integration receives contributions.
QED description of pair production

from poles \( \pm \tau_m, \pm i\tau_n \), so that the residue theorem leads to,

\[
\Delta L^\text{sin}_{\text{eff}} = i \frac{\alpha e \beta}{2\pi} \sum_{n=1}^{\infty} \frac{1}{n} \coth \left( \frac{n\pi \beta}{\epsilon} \right) \exp \left( -n\pi E_c / \epsilon \right) \tag{5.7.49}
\]

\[
- i \frac{\alpha e \beta}{2\pi} \sum_{m=1}^{\infty} \frac{1}{m} \coth \left( \frac{m\pi \epsilon}{\beta} \right) \exp \left( im\pi E_c / \beta \right) \tag{5.7.50}
\]

The first part (5.7.49) leads to the exact non-perturbative Schwinger rate (5.7.39) for pair production.

The second term, as we see below, is canceled by the imaginary part of the cos term. In fact, shifting \( s \to s - i\eta \), we rewrite the cos part of effective Lagrangian (5.7.48) as

\[
\Delta L^\text{cos}_{\text{eff}} = \frac{1}{2(2\pi)^2} \sum_{n,m=\infty}^{\infty} \int_0^\infty ds \cos(sm^2_e) \left( \frac{s\delta_{n0}}{s^2 - \tau^2_m - i\eta} - \frac{s\delta_{n0}}{s^2 + \tau^2_n - i\eta} \right). \tag{5.7.51}
\]

In the first term of magnetic part, singularities \( s = \tau_m, (m > 0) \) and \( s = -\tau_m, (m < 0) \) appear in integrating \( s \)-axis. We decompose

\[
\frac{s}{s^2 - \tau^2_m - i\eta} = \frac{i\pi}{2} \delta(s - \tau_m) + \frac{i\pi}{2} \delta(s + \tau_m) + \mathcal{P} \frac{s}{s^2 - \tau^2_m}, \tag{5.7.52}
\]

where \( \mathcal{P} \) indicates the principle value under the integral. The integrals over the \( \delta \)-functions give

\[
\Delta \delta L^\text{cos}_{\text{eff}} = i \frac{\alpha e \beta}{2\pi} \sum_{m=1}^{\infty} \frac{1}{m} \coth \left( \frac{m\pi \epsilon}{\beta} \right) \exp \left( im\pi E_c / \beta \right), \tag{5.7.53}
\]

which exactly cancels the second part (5.7.50) of the sin part \( \Delta L^\text{sin}_{\text{eff}} \).

It remains to find the principle-value integrals in Eq. (5.7.51), which corresponds to the real part of the effective Lagrangian

\[
(\Delta L^\text{cos}_{\text{eff}})_{P} = \frac{1}{2(2\pi)^2} \sum_{n,m=\infty}^{\infty} \int_0^\infty ds \cos(sm^2_e) \left( \frac{s\delta_{n0}}{s^2 - \tau^2_m} - \frac{s\delta_{n0}}{s^2 + \tau^2_n} \right). \tag{5.7.54}
\]

We rewrite the cos function as \( \cos(sm^2_e) = (e^{ism^2_e} + e^{-ism^2_e})/2 \) and make the rotations of integration contours by \( \pm \pi/2 \) respectively,

\[
(\Delta L^\text{cos}_{\text{eff}})_{P} = \frac{1}{2(2\pi)^2} \sum_{n,m=\infty}^{\infty} \int_0^\infty ds \cos(sm^2_e) \times \mathcal{P} \int_0^\infty d\tau \left( \frac{s\delta_{n0} \tau e^{-\tau}}{\tau^2 - (s^2)^2} - \frac{s\delta_{n0} \tau e^{-\tau}}{\tau^2 - (s^2)^2} \right). \tag{5.7.55}
\]
Using the formulas (see Secs. 3.354, 8.211.1 and 8.211.2 in Ref. [206])

\[ J(z) \equiv \mathcal{P} \int_0^\infty ds \frac{s e^{-s}}{s^2 - z^2} = -\frac{1}{2} \left[ e^{-z} \text{Ei}(z) + e^z \text{Ei}(-z) \right], \quad (5.7.56) \]

where \( \text{Ei}(z) \) is the exponential-integral function,

\[ \text{Ei}(z) \equiv \mathcal{P} \int_{-\infty}^z dt \frac{e^t}{t} = \log(-z) + \sum_{k=1}^\infty \frac{z^k}{k^k}, \quad (5.7.57) \]

we obtain the principal-value integrals (5.7.55),

\[ (\Delta \mathcal{L}_\text{eff}^\text{cos})_\mathcal{P} = \frac{1}{2(2\pi)^2} \sum_{n,m=-\infty}^{\infty} \frac{1}{\tau_m^2 + \tau_n^2} \left[ \delta_{m0} J(i\tau_m m_e^2) - \delta_{n0} J(\tau_n m_e^2) \right]. \quad (5.7.58) \]

Having so obtained the real part of an effective Lagrangian for an arbitrary constant electromagnetic field we recover the usual approximate results by suitable expansion of the exact formula. With the help of the series and asymptotic representation (see formula 8.215 in Ref. [206]) of the exponential-integral function \( \text{Ei}(z) \) for large \( z \), corresponding to weak electromagnetic fields \( \varepsilon \ll 1, \beta \ll 1 \),

\[ J(z) = -\frac{1}{z^2} - \frac{6}{z^4} - \frac{120}{z^6} - \frac{5040}{z^8} - \frac{362880}{z^{10}} + \cdots, \quad (5.7.59) \]

and Eq. (5.7.58), we find,

\[ (\Delta \mathcal{L}_\text{eff}^\text{cos})_\mathcal{P} = \frac{1}{2(2\pi)^2} \sum_{n,m=-\infty}^{\infty} \frac{1}{\tau_m^2 + \tau_n^2} \left\{ \delta_{m0} \left[ \frac{1}{\tau_m^4 m_e^4} + \frac{6}{\tau_n^4 m_e^4} + \frac{120}{\tau_n^6 m_e^6} + \cdots \right] + \delta_{n0} \left[ \frac{1}{\tau_n^4 m_e^4} - \frac{6}{\tau_m^4 m_e^4} + \frac{120}{\tau_m^6 m_e^6} + \cdots \right] \right\}. \quad (5.7.60) \]

Applying the summation formulas (5.7.46), the weak field expansion (5.7.60) is seen to agree with the Heisenberg and Euler effective Lagrangian [7],

\[ (\Delta \mathcal{L}_\text{eff})_\mathcal{P} = \frac{2\alpha}{90\pi E_c^2} \left\{ (\mathbf{E}^2 - \mathbf{B}^2)^2 + 7(\mathbf{E} \cdot \mathbf{B})^2 \right\} + \frac{2\alpha}{315\pi^2 E_c^4} \left\{ 2(\mathbf{E}^2 - \mathbf{B}^2)^3 + 13(\mathbf{E}^2 - \mathbf{B}^2)(\mathbf{E} \cdot \mathbf{B})^2 \right\} + \cdots, \quad (5.7.61) \]

which is expressed in terms of a powers series of weak electromagnetic fields up to \( O(\alpha^3) \). The expansion coefficients of the terms of order \( n \) have the general form \( m_e^4/(E_c)^n \). As long as the fields are much smaller than \( E_c \), the expansion converges.

On the other hand, we can address the limiting form of the effective La-
QED description of pair production

5.7 QED description of pair production

Grangian (5.7.58) corresponding to electromagnetic fields ($\epsilon \gg 1, \beta \gg 1$). We use the series and asymptotic representation (formulas 8.214.1 and 8.214.2 in Ref. [206]) of the exponential-integral function $\text{Ei}(z)$ for small $z \ll 1$,

$$J(z) = -\frac{1}{2} \left[ e^z \ln(z) + e^{-z} \ln(-z) \right] + \gamma_E + O(z), \quad (5.7.62)$$

with $\gamma_E = 0.577216$ being the Euler-Mascheroni constant, we obtain the leading terms in the strong field expansion of Eq. (5.7.58),

$$(\Delta L_\text{eff}^\text{cos})_P = \frac{1}{2(2\pi)^2} \sum_{n,m=-\infty}^{\infty'} \frac{1}{\tau_n^2 + \tau_m^2} \left[ \delta_{n0} \ln(\tau_n m^2_e) - \delta_{m0} \ln(\tau_m m^2_e) \right] + \cdots. \quad (5.7.63)$$

In the case of vanishing magnetic field $B = 0$ and $m = 0$ in Eq. (5.7.63), we have,

$$(\Delta L_\text{eff}^\text{cos})_P = \frac{1}{2(2\pi)^2} \sum_{n=1}^{\infty} \frac{1}{\tau_n^2} \ln(\tau_n m^2_e) + \cdots = \frac{aE^2}{2\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \ln \left( \frac{n\pi m_e}{E} \right) + \cdots, \quad (5.7.64)$$

for a strong electric field $E$. In the case of vanishing electric field $E = 0$ and $n = 0$ in Eq. (5.7.63), we obtain for the strong magnetic field $B$,

$$(\Delta L_\text{eff}^\text{cos})_P = -\frac{1}{2(2\pi)^2} \sum_{m=1}^{\infty} \frac{1}{\tau_m^2} \ln(\tau_m m^2_e) + \cdots = -\frac{aB^2}{2\pi^2} \sum_{m=1}^{\infty} \frac{1}{m^2} \ln \left( \frac{m\pi m_e}{B} \right) + \cdots. \quad (5.7.65)$$

The $(m = 1)$ term is the one obtained by Weisskopf [24].

We have presented in Eqs. (5.7.49), (5.7.50), (5.7.53), (5.7.58) closed form results for the one-loop effective Lagrangian $\Delta L_\text{eff}^{}$ (5.7.37) for arbitrary strength of constant electromagnetic fields. The results will receive fluctuation corrections from higher loop diagrams. These carry one or more factors $\alpha, \alpha^2, \ldots$ and are thus suppressed by factors $1/137$. Thus results are valid for all field strengths with an error no larger than roughly 1%. If we include, for example, the two-loop correction, the first term in the Heisenberg and Euler effective Lagrangian (5.7.61) becomes [145]

$$(\Delta L_\text{eff})_P = \frac{2\alpha}{90\pi m_e^2} \left\{ \left( 1 + \frac{40\alpha}{9\pi} \right) (E^2 - B^2)^2 + 7 \left( 1 + \frac{1315\alpha}{252\pi} \right) (E \cdot B)^2 \right\}. \quad (5.7.66)$$

Readers can consult the recent review article [207], where one finds discussions and computations of the effective Lagrangian at the two-loop level, and [208] for discussion of pair production rate.
5.8 Theory of pair production in an alternating field

When the external electromagnetic field $F_{\mu\nu}^e$ is space-time-dependent, i.e., $F_{\mu\nu}^e = F_{\mu\nu}^e(x, t)$ the exponential in Eq. (5.7.16) can no longer be calculated exactly. In this case, JWKB methods have to be used to calculate pair production rates [30, 30, 31, 32, 32, 36, 37, 209, 210]. The aim of this section is to show how one can use a semi-classical JWKB approach to estimate the rate of pair production in an oscillating electric field as first indicated by Brezin and Itzykson in Ref. [36]. They evaluated the production rate of charged boson pairs. The results they obtained can be straightforwardly generalized to charged fermion case, since the spins of charged particles contribute essentially with a counting factor to the final results (see Secs. 3.5 and 5.7.1). Thus, let the electromagnetic potential be $z$ directed, uniform in space and periodic in time with frequency $\omega_0$:

$$A_{\mu}^e(x) = (0, 0, 0, A(t)), \quad A(t) = \frac{E}{\omega_0} \cos \omega_0 t. \quad (5.8.1)$$

Then the electric field is $z$ directed, uniform in space and periodic in time as well. The electric field strength is given by $E(t) = -\dot{A}(t) = E \sin \omega_0 t$. It is assumed that the electric field is adiabatically switched on and damped off in a time $T^e$, which is much larger than the period of oscillation $T_0 = 2\pi/\omega_0$. Suppose also that $T_0$ is much larger than the Compton time $2\pi/\omega$ of the created particle, i.e.,

$$T^e \gg T_0 \gg \frac{2\pi}{\omega} \simeq \frac{2\pi}{m_e}, \quad (5.8.2)$$

where $\omega = \sqrt{|p|^2 + m_e^2}$, $p$ being the 3-momentum of the created particle. Furthermore, $eE$ is assumed to be much smaller than $m_e^2$, i.e., $E \ll E_c$ (see Eq. (3.5.13)).

We have to study the time evolution of a scattered wave function $\psi(t)$ representing the production of particle and antiparticle pairs in the electromagnetic potential (5.8.1). As usual, an antiparticle can be thought of as a wave-packet moving backward in time. Therefore, for large positive time (forward) only positive energy modes ($\sim e^{-i\omega t}$) contribute to $\psi(t)$. Similarly, for large negative times both positive energy and negative energy modes ($\sim e^{i\omega t}$) contribute to $\psi(t)$ which satisfies the differential equation [36]:

$$\left[ \frac{d^2}{dt^2} + \omega^2(t) \right] \psi(t) = 0, \quad (5.8.3)$$
where the “variable frequency” is defined as
\[
\omega(t) \equiv \left\{ m^2_e + p_{\perp}^2 + \left[ p_z - eA(t) \right]^2 \right\}^{1/2}.
\] (5.8.4)

The JWKB method suggests a general solution of the form
\[
\psi(t) = \alpha(t)e^{-i\chi(t)} + \beta(t)e^{i\chi(t)}, \quad \chi(t) \equiv \int_0^t dt' \omega(t'),
\] (5.8.5)

where the boundary conditions at large positive and negative times are:
\[
\alpha(-\infty) = 1, \quad \beta(+\infty) = 0; \quad \dot{\chi}(\pm \infty) = \omega.
\] (5.8.6)

The backward scattering amplitude \(\beta(t)\) for large negative time \((t \rightarrow -\infty)\) represents the probability of antiparticle production.

The normalization condition \(|\psi(t)|^2 = 1\) implies
\[
\dot{\alpha}(t)e^{-i\chi(t)} + \dot{\beta}(t)e^{i\chi(t)} = 0. \tag{5.8.7}
\]

Eq. (5.8.3) can be written in terms of the scattering amplitudes as
\[
\dot{\alpha}(t)e^{-i\chi(t)} - \dot{\beta}(t)e^{i\chi(t)} = -\frac{\dot{\omega}(t)}{\omega(t)} \left[ \alpha(t)e^{-i\chi(t)} - \beta(t)e^{i\chi(t)} \right], \tag{5.8.8}
\]
or, which is the same,
\[
\dot{\alpha}(t) = -\frac{\dot{\omega}(t)}{2\omega(t)} \left[ \alpha(t) - \beta(t)e^{i2\chi(t)} \right], \tag{5.8.9}
\]
\[
\dot{\beta}(t) = -\frac{\dot{\omega}(t)}{2\omega(t)} \left[ \beta(t) - \alpha(t)e^{-i2\chi(t)} \right]. \tag{5.8.10}
\]

It follows from assumption (5.8.2) that \(\dot{\omega}(t)\) vanishes as \(|t| \rightarrow \infty\), i.e.,
\[
\frac{\dot{\omega}(t)}{\omega^2(t)} = \frac{eE[p_z - eA(t)]}{\{m^2_e + p_{\perp}^2 + [p_z - eA(t)]^2\}^{3/2}} \ll 1. \tag{5.8.11}
\]

More precisely
\[
\left| \frac{\dot{\omega}(t)}{\omega^2(t)} \right| < \frac{eE}{m^2_e + p_{\perp}^2} < \frac{eE}{m^2_e} \ll 1. \tag{5.8.12}
\]

Therefore, \(\alpha(t)\) and \(\beta(t)\) slowly vary in time and tend to constants as \(|t| \rightarrow \infty\). The phase \(e^{i2\chi(t)}\) oscillates very rapidly as compared to the variation of \(\alpha(t)\) and \(\beta(t)\), for \(\chi(t) = \omega(t) \gg |\dot{\omega}(t)/\omega(t)|\). In the zeroth order the oscillating
terms in Eqs. (5.8.9), (5.8.10) are negligible and one finds
\[ \alpha^{(0)}(t) = \left[ \frac{\omega}{\omega(t)} \right]^{1/2} \approx 1; \quad \beta^{(0)}(t) = 0, \] (5.8.13)
which duly satisfy the boundary conditions, and
\[ \beta^{(1)}(t) = \int_t^\infty dt' \frac{\dot{\omega}(t')}{{\omega(t')}} e^{-i2\chi(t')}, \] (5.8.14)
where (5.8.2) and (5.8.12) have been used. \(|\beta^{(1)}(-\infty)|^2\) gives information about the probability of particle-antiparticle pair production. Namely, the probability of pair production per unit volume and time is given by
\[ \tilde{P} = \lim_{T_e \to \infty} \frac{1}{T_e} \int \frac{d^3k}{(2\pi)^3} |\beta^{(1)}(-T_e)|^2 \] (5.8.15)
Since \(\omega(t)\) is a periodic function with the same frequency \(\omega_0\) as \(A(t)\) one can make a Fourier series expansion:
\[ \frac{\dot{\omega}(t)}{2\omega(t)} = \sum_{n=-\infty}^{+\infty} c_n e^{in\omega_0}. \] (5.8.16)
Defined a renormalized frequency \(\Omega\) via \(\chi(t) = t\Omega\) one finds
\[ \Omega \equiv \int_0^{2\pi} dx \frac{1}{2\pi} \left[ m_e^2 + k_{\perp}^2 + \left( k_3 - \frac{eE}{\omega_0} \cos x \right)^2 \right]^{1/2}. \] (5.8.17)
so that
\[ \lim_{T_e \to \infty} \frac{1}{T_e} \left| \int_{-T_e/2}^{T_e/2} dt' \frac{\dot{\omega}(t')}{2\omega(t')} e^{-i2\chi(t')} \right|^2 = 2\pi \sum_n \delta(n\omega_0 - 2\Omega) |c_n|^2. \] (5.8.18)
Consequently, the probability of pair production (5.8.15) is,
\[ \tilde{P} = \int \frac{d^3k}{(2\pi)^3} \left( \sum_n \delta(n\omega_0 - 2\Omega) |c_n|^2 \right) = \int \frac{d^3k}{(2\pi)^3} \omega_0 |c_n|^2, \] (5.8.19)
5.8 Theory of pair production in an alternating field

where \( n^o = 2\Omega/\omega_0 \) and \( c_{n^o} \) are determined via Eq. (5.8.16) as

\[
c_{n^o} = \left\{ \frac{2i}{\omega_0} \int_0^\infty dx' \left[ m_e^2 + p_\perp^2 + \left( p_z - \frac{eE}{\omega_0} \cos(x') \right)^2 \right] \right\}^{1/2}.
\]

(5.8.20)

The expression for \( c_{n^o} \) contains a very rapidly oscillating phase factor with frequency of the order of \( m_e/\omega_0 \), and it decreases very rapidly in terms of imaginary time \( \tau = -it \). Its evaluation requires the application of the steepest-descent method in the complex time \( x = \omega_0 t \) plane. This is done by selecting a proper contour turning in a neighborhood of the saddle point and following the steepest-descent line, so as to find the main contributions to the integral in Eq. (5.8.20). The saddle point originates from branch points and poles in Eq. (5.8.20), which are the zeros of \( \omega(x) \). Mathematical details can be found in Ref. [36]. One finds

\[
\Phi \sim \frac{\omega_0}{9} \int \frac{d^3k}{(2\pi)^3} e^{-2A} \cos^2 B,
\]

(5.8.21)

where

\[
-A + iB = \frac{2i}{\omega_0} \int_0^{x_0} dx' \left[ m_e^2 + p_\perp^2 + \left( p_z - \frac{eE}{\omega_0} \cos(x') \right)^2 \right]^{1/2},
\]

(5.8.22)

and the saddle point is \( x_0 = 1/\pi + i \sinh^{-1}[(\omega_0/eE)(m_e^2 + p_\perp^2)^{1/2}] \).

The exponential factor \( e^{-2A} \) in Eq. (5.8.21) indicates that particle-antiparticle pairs tend to be emitted with small momenta. This allows one to estimate the right-hand side of Eq. (5.8.21) as follow: (i) \( p_z \) is set equal to zero, moreover, the range of the \( p_z \)-integration is of the order of \( 2eE/\omega_0 \) as suggested by the classical equation of motion (5.8.3); (ii) \( \cos^2 B \) is replaced by its average value \( 1/2 \). As a result, one obtains [36],

\[
\Phi \sim \frac{(eE)^3}{18\pi \omega_0^2} \int_{\eta^{-1}}^\infty du u \exp \left[ -\frac{\pi eE}{\omega_0^2} u^2 g(u) \right],
\]

(5.8.23)

where \( \eta^{-1} = m_e \omega_0/(eE) \), \( u = (m_e^2 + p_\perp^2) \omega_0^2/(eE)^2 \) and

\[
g(z) = \frac{4}{\pi} \int_0^1 dy \left[ \frac{1 - y^2}{1 + z^{-2}y^2} \right]^{1/2} = F \left( \frac{1}{2}, \frac{1}{2} ; 2, \frac{1}{2} ; -z^{-2} \right),
\]

(5.8.24)

where \( F(1/2, 1/2; 2; -z^{-2}) = {}_2F_1(1/2, 1/2; 2; -z^{-2}) \) is the Gauss hypergeo-
metrical function. The function $u^2g(u)$ is monotonically increasing:

$$\frac{eE}{\omega_0^2}u^2g(u) \geq \frac{eE}{\omega_0}\eta^{-2}g(\eta) = \frac{m_e^2}{eE}g(\eta) \gg 1,$$

(5.8.25)

which indicates that the integral in (5.8.23) is strongly dominated by values in a neighborhood of $u = \eta^{-1}$. This allows one to approximately perform the integration and leads to the rate of pair production of charged bosons [36],

$$\tilde{P}_{\text{boson}} \simeq \frac{\alpha E^2}{2\pi} \frac{1}{g(\eta) + \frac{1}{\pi}g'(\eta)} \exp \left[ -\frac{\pi m_e^2}{eE}g(\eta) \right].$$

(5.8.26)

Analogously, the rate of pair production of charged fermions can be approximately obtained from Eq. (5.8.26) by taking into account two helicity states of fermions (see Secs. 3.5 and 5.7.1),

$$\tilde{P}_{\text{fermion}} \simeq \frac{\alpha E^2}{\pi} \frac{1}{g(\eta) + \frac{1}{\pi}g'(\eta)} \exp \left[ -\frac{\pi m_e^2}{eE}g(\eta) \right].$$

(5.8.27)

This formula has played an important role in recent studies of electron and positron pair production by laser beams, which we will discuss in some details in Section 7.2. Momentum spectrum of electrons and positrons, produced from the vacuum, was calculated in [30–32, 37]. For $\eta \gg 1$ this distribution is concentrated along the direction of electric field, while for $\eta \ll 1$ it approaches isotropic one.

Unfortunately, it appears very difficult to produce a macroscopic electric field with strength of the order of the critical value (2.0.1) and lifetime long enough ($\gg \hbar/(m_e c^2)$) in any ground laboratory to directly observe the Sauter-Euler-Heisenberg-Schwinger process of electron–positron pair production in vacuum. The same argument applies for the production of any other pair of fermions or bosons. In the following Section, we discuss some ideas to experimentally create a transient electric field $E \ll E_c$ in Earth-bound laboratories, whose lifetime is expected to be long enough (larger than $\hbar/m_e c^2$) for the pair production process to take place.

### 5.9 Nonlinear Compton scattering and Breit-Wheeler process

In Section 3.2, we have discussed the Breit–Wheeler process [2] in which an electron–positron pair is produced in the collision of two real photons $\gamma_1 + \gamma_2 \rightarrow e^+ + e^-$ (3.0.2). The cross-section they obtained is $O(r_e^2)$, where $r_e$ is the classical electron radius, see Eq. (3.2.10). This lowest order photon-photon
pair production cross-section is so small that it is difficult to observe creation of pairs in the collision of two high-energy photon beams, even if their center of mass energy is larger than the energy-threshold \(2m_e c^2 = 1.02 \text{ MeV}\).

In the previous Sections we have seen that in strong electromagnetic fields in lasers the effective nonlinear terms (5.7.61) become significant and therefore, the interaction needs not to be limited to initial states of two photons [61,62]. A collective state of many interacting laser photons occurs.

We turn now to two important processes [38,39] emerging in the interaction of an ultrarelativistic electron beam with a terawatt laser pulse, performed at SLAC [63], when strong electromagnetic fields are involved. The first process is the nonlinear Compton scattering, in which an ultrarelativistic electron absorbs multiple photons from the laser field, but emits only a single photon via the process

\[
e + n\omega \rightarrow e' + \gamma,
\]

where \(\omega\) represents photons from the strong electromagnetic wave of the laser beam (its frequency being \(\omega\)), \(n\) indicates the number of absorbed photons and \(\gamma\) represents a high-energy emitted photon (see Eq. (5.9.2) for cross symmetry). The theory of this nonlinear Compton effect (5.9.1) is given in Section 5.10. The same process (5.9.1) has been expressed by Bamber et al. [211] in a semi-classical framework. The second is the nonlinear Breit–Wheeler process

\[
\gamma + n\omega \rightarrow e^+ + e^-.
\]

between this very high-energy photon \(\gamma\) and multiple laser photons: the high-energy photon \(\gamma\), created in the first process, propagates through the laser field and interacts with laser photons \(n\omega\) to produce an electron–positron pair [39].

In the electric field \(E\) of an intense laser beam, an electron oscillates with the frequency \(\omega\) of the laser and its maximum velocity in unit of the speed of light is given by

\[
v_{\text{max}}\gamma_{\text{max}} = \frac{eE}{m\omega}, \quad \gamma_{\text{max}} = \frac{1}{\sqrt{1 - \frac{v_{\text{max}}^2}}},
\]

In the case of weak electric field, \(v_{\text{max}} \ll 1\) and the nonrelativistic electron emits the dipole radiation well described in linear and perturbative QED. On the other hand, in the case of strong electric fields, \(v_{\text{max}} \rightarrow 1\) and the ultrarelativistically oscillating electron emits multi-pole radiation. The radiated power is then a nonlinear function of the intensity of the incident laser beam. Using the maximum velocity \(v_{\text{max}}\) of oscillating electrons in the electric field of laser beam, one can characterize the effect of nonlinear Compton scattering by the dimensionless parameter

\[
\eta = v_{\text{max}}\gamma_{\text{max}} = \frac{eE_{\text{rms}}}{m\omega} = \frac{m_e c^2 E_{\text{rms}}}{\omega \hbar E_c},
\]
where the subscript ‘rms’ means root-mean-square, with respect to the number of interacting laser photons with scattered electron. The parameter $\eta$ can be expressed as a Lorentz invariant,

$$\eta^2 = \frac{e^2 |\langle A_\mu A^\mu \rangle|}{m_e^2}, \quad (5.9.5)$$

where $A_\mu$ is the gauge potential of laser wave, $\partial^\mu A_\mu = 0$ and the time-average is taken over one period of laser wave, $\langle A_\mu \rangle = 0$ and

$$\langle A_\mu A^\mu \rangle = \langle (A_\mu - \langle A_\mu \rangle)^2 \rangle. \quad (5.9.6)$$

Eq. (5.9.5) shows that $\eta^2$ is the intensity parameter of laser fields, and $\eta$ in (5.9.4) coincides with the parameter $\eta$ introduced in Eq. (5.8.23) for the pair production in an alternating electric field (see Section 5.8).

### 5.10 Quantum description of nonlinear Compton effect

In Refs. [40–45, 61, 62, 90, 212–216], the quantum theory of the interaction of free electrons with the field of a strong electromagnetic wave has been studied. The application of quantum perturbation theory to such interaction requires not only that the interaction constant $\alpha$ should be small but also that field should be sufficiently weak. The characteristic quantity in this respect is the dimensionless invariant ratio $\eta$, see (5.9.5). The photon emission processes occurring in the interaction of an electron with the field of a strong electromagnetic wave have been discussed in Ref. [90] for any $\eta$ value. The method used is based on an exact treatment of this interaction, while the interaction of the electron with the newly emitted photons regarded as a perturbation.

Laser beam is considered as a monochromatic plane wave, described by the gauge potential $A_\mu(\phi)$ and $\phi = kx$, where wave vector $k = (\omega, k)$ ($k^2 = 0$) (see Eq. (3.1.3)). The Dirac equation can be exactly solved [217] for an electron moving in this field of electromagnetic plane wave of an arbitrary polarization and the normalized wave function of the electron with momentum $p$ is given by (c.e.g. [90]),

$$\psi_p = \left[ 1 + \frac{e}{2(kp)} k A \right] \frac{u(p)}{\sqrt{2q_0}} e^{ip\phi}, \quad (5.10.1)$$

$$\Phi = -px - \int_0^{kx} \left[ \frac{e}{(kp)} (pA) - \frac{e^2}{2(kp)} A^2 \right] d\phi, \quad (5.10.2)$$
where \( u(p) \) is the solution of free Dirac equation \( (\not{p} - m_e)u(p) = 0 \) and the time-average value of 4-vector,

\[
q = p - \frac{e^2 \langle A^2 \rangle}{2(kp)} k,
\]

is the kinetic momentum operator in the electron state \( \psi_p \) (5.10.1) and the "effective mass" \( m_* \) of the electron in the field is

\[
q^2 = m_*^2, \quad m_* = m_e \sqrt{1 + \eta^2},
\]

where \( \eta^2 \) is given by (5.9.5). The electron becomes “heavy” in an oscillating electromagnetic field.

The \( S \)-matrix element for a transition of the electron from the state \( \psi_p \) to the state \( \psi_{p'} \), with emission of a photon having momentum \( k' \) and polarization \( e' \) is given by (c.e.g. [90])

\[
S_{fi} = -ie \int \bar{\psi}_{p'}(\gamma e') \psi_p e^{ik'x} \sqrt{2\omega'} d^4x
\]

\[
= \frac{1}{2\omega' \cdot 2q_0 \cdot 2q'_0} \sum_n M_{fi}^{(n)} (2\pi)^4 i\delta(4)(nk + q - q' - k'),
\]

where the integrand in Eq. (5.10.5) is expanded in Fourier series and expansion coefficients are in terms of Bessel functions \( J_n \), the scattering amplitude \( M_{fi}^{(n)} \) in Eq. (5.10.6) is obtained\(^1\) by integrating over \( x \). Eq. (5.10.6) shows that \( S_{fi} \) is an infinite sum of terms, each corresponds to an energy-momentum conservation law \( nk + q = q' + k' \), indicating an electron \( (q) \) absorbs \( n \)-photons \( (nk) \) and emits another photon \( (k') \) of frequency

\[
\omega' = \frac{nw}{1 + (n\omega/m_*)(1 - \cos \theta)},
\]

in the frame of reference where the electron is at rest \( (q = 0, q_0 = m_*) \), and \( \theta \) is the angle between \( k \) and \( k' \). Given the \( n \)th term of the \( S \)-matrix \( S_{fi} \) (5.10.6), the differential probability per unit volume and unit time yields,

\[
d\gamma_e^{(n)} = \frac{d^3k' d^3q'}{(2\pi)^6 \cdot 2\omega' \cdot 2q_0 \cdot 2q'_0} |M_{fi}^{(n)}|^2 (2\pi)^4 i\delta(4)(nk + q - q' - k').
\]

Integrating over the phase space of final states \( \int d^3k' d^3q' \), one obtains the total probability of emission from unit volume in unit time (circular polari-

\(^1\)The explicit expression \( M_{fi}^{(n)} \) is not given here for its complexity, see for example [90].
5 Pair production and annihilation in QED

\[ P_{e \gamma} = \frac{e^2 m_e^2}{4 q_0} \sum_{n=1}^{\infty} \int_{0}^{\nu_n} \frac{d\nu}{(1 + \nu)^2} \left[ -4 J_n^2(z) + \eta^2 \left( 2 + \frac{\kappa^2}{1 + \kappa} \right) (J_{n+1}^2 + J_{n-1}^2 - 2 J_n^2) \right], \]

where \( \kappa = (k k') / (k p') \), \( \kappa_n = 2 n (k p) / m_e^2 \) and Bessel functions \( J_n(z) \),

\[ z = 2 m_e^2 \frac{\eta}{(1 + \eta^2)^{1/2}} \left[ \frac{\nu}{\nu_n} \left( 1 - \frac{\nu}{\nu_n} \right) \right]^{1/2}, \]

for any \( \eta \) value. A systematic investigation of various quantum processes in the field of a strong electromagnetic wave can be found in [40–45], in particular photon emission and pair production in the field of a plane wave with various polarizations are discussed.

We now turn to the Breit–Wheeler process for multi-photons (5.9.2). In this process, the pair production is attributed to the interaction of a high-energy photon with many laser photons in the electromagnetic laser wave. Actually, the Breit–Wheeler process for multi-photons, see Eq. (5.9.2), is related to the nonlinear Compton scattering process, see Eq. (5.9.1), by crossing symmetry.

By replacement \( p \rightarrow -p \) and \( k' \rightarrow -l \) and reverse the common sign of the expression in Eq. (5.10.6), one obtains the probability of pair production (5.9.2) by a photon \( \gamma \) (momentum \( l \)) colliding with \( n \) laser photons (momentum \( k \)) per unit volume in unit time (circular polarization) [40–45],

\[ P_{\gamma \gamma} = \frac{e^2 m_e^2}{16 l_0} \sum_{n>n_0}^{\infty} \int_{1}^{\nu_n} \frac{d\nu}{\nu^{3/2}(1 + \nu)^{1/2}} \times \]

\[ \times \left[ 2 J_n^2(z) + \eta^2 (2\nu - 1) (J_{n+1}^2 + J_{n-1}^2 - 2 J_n^2) \right], \]

where \( \nu = (k l^2 / 4 k q) (k q') \), \( \nu_n = n / n_0, n_0 = 2 m_e^2 / (k l) \) and Bessel functions \( J_n(z) \),

\[ z = 4 m_e^2 \eta (1 + \eta^2)^{1/2} \left[ \frac{\nu}{\nu_n} \left( 1 - \frac{\nu}{\nu_n} \right) \right]^{1/2}. \]

In Eq. (5.10.11), the number \( n \) of laser photons must be larger than \( n_0 (n > n_0) \), which is the energy threshold \( n_0 (k l) = 2 m_e^2 \) for the process (5.9.2) of pair production to occur.
As shown in previous sections, the rate of pair production may be split into an exponential and a pre-exponential factor. The exponent is determined by the classical trajectory of the tunneling particle in imaginary time which has the smallest action. It plays the same role as the activation energy in a Boltzmann factor with a “temperature” $\hbar$. The pre-exponential factor is determined by the quantum fluctuations of the path around that trajectory. At the semi-classical level, the latter is obtained from the functional determinant of the quadratic fluctuations. It can be calculated in closed form only for a few classical paths \cite{202}. An efficient technique for doing this is based on the JWKB wave functions, another on solving the Heisenberg equations of motion for the position operator in the external field \cite{202}.

Given the difficulties in calculating the pre-exponential factor, only a few nonuniform electric fields in space or in time have led to analytic results for the pair production rate: (i) the electric field in the $z$-direction is confined in the space $x < x_0$, i.e., $E = E(x)\hat{z}$ where $E(x) = E_0\Theta(x_0 - x)$ \cite{218,219}; (ii) the electric field in the $z$-direction depends only on the light-cone coordinate $z_+ = (t + z)/\sqrt{2}$, i.e., $E = E(z_+)\hat{z}$ \cite{220,221}. If the nonuniform field has the form $E(z) = E_0/\cosh^2(z)$, which will be referred as a Sauter field, the rate was calculated by solving the Dirac equation \cite{33} in the same way as Heisenberg and Euler did for the constant electric field. For general space and time dependencies, only the exponential factor can be written down easily — the fluctuation factor is usually hard to calculate \cite{65}. In the Coulomb field of heavy nucleus whose size is finite and charge $Z$ is supercritical, the problem becomes even more difficult for bound states being involved in pair production, and a lot of effort has been spent on this issue \cite{65,203,222}.

If the electric field has only a time dependence $E = E(t)$, both exponential and pre-exponential factors were approximately computed by Brezin and Itzykson using JWKB methods for the purely periodic field $E(t) = E_0 \cos \omega_0 t$ \cite{36}. The result was generalized by Popov in Ref. \cite{209,223} to more general time-dependent fields $E(t)$. After this, several time-independent but space-dependent fields were treated, for instance an electric field between two conducting plates \cite{224}, and an electric field around a Reissner–Nordström black
hole [46].

The semi-classical expansion was carried beyond the JWKB approximation by calculating higher order corrections in powers of $\hbar$ in Refs. [51,52] and [53]. Unfortunately, these terms do not comprise all corrections of the same orders $\hbar$ as explained in [54].

An alternative approach to the same problems was recently proposed by using the worldline formalism [47], sometimes called the “string-inspired formalism”. This formalism is closely related to Schwinger’s quantum field theoretic treatment of the tunneling problem, where the evaluation of a fluctuation determinant is required involving the fields of the particle pairs created from the vacuum. The worldline approach is a special technique for calculating precisely this functional determinant. Within the worldline formalism, Dunne and Schubert [48] calculated the exponential factor and Dunne et al. [49] gave the associated prefactor for various field configurations: for instance a spatially uniform, and single-pulse field with a temporal Sauter shape $\propto 1/\cosh^2 \omega t$. For general $z$-dependencies, a numerical calculation scheme was proposed in Ref. [225–228] and applied further in [229]. For a multidimensional extension of the techniques see Ref. [50].

In this Section, a general expression is derived for the pair production rate in nonuniform electric fields $E(z)$ pointing in the $z$-direction recently derived in [54,230]. A simple variable change in all formulas leads to results for electric fields depending also on time rather than space. As examples, three cases will be treated: (i) a nonzero electric field confined to a region of size $\ell$, i.e., $E(z) \neq 0, |z| \lesssim \ell$ (Sauter field see Eq. (6.3.4)); (ii) a nonzero electric field in a half space, i.e., $E(z) \neq 0, z \gtrsim 0$ (see Eq. (6.3.25)); (iii) an electric field increasing linearly like $E(z) \sim z$. In addition, the process of negative energy electrons tunneling into the bound states of an electric potential with the emission of positrons will be studied, by considering the case: the electric field $E(z) \sim z$ of harmonic potential $V(z) \sim z^2$.

### 6.1 Semi-classical description of pair production

The phenomenon of pair production in an external electric field can be understood as a quantum mechanical tunneling process of Dirac electrons [1, 80]. In the original Dirac picture, the electric field bends the positive and negative energy levels of the Hamiltonian, leading to a level crossing and a tunneling of the electrons in the negative energy band to the positive energy band. Let the field vector $E(z)$ point in the $z$-direction. In the one-dimensional potential energy (3.4.4) the classical positive and negative energy spectra are

$$
E_{\pm}(p_z, p_\perp; z) = \pm \sqrt{(cp_z)^2 + c^2 p_\perp^2 + (m_e c^2)^2 + V(z)}, \quad (6.1.1)
$$
where $p_z$ is the momentum in the $z$-direction, $p_\perp$ the momentum orthogonal to it, and $p_\perp \equiv |p_\perp|$. For a given energy $\mathcal{E}$, the tunneling takes place from $z_-$ to $z_+$ determined by $p_z = 0$ in Eq. (6.1.1)

$$\mathcal{E} = \mathcal{E}_+(0, p_\perp; z_+) = \mathcal{E}_-(0, p_\perp; z_-).$$

(6.1.2)

The points $z_\pm$ are the turning points of the classical trajectories crossing from the positive energy band to the negative one at energy $\mathcal{E}$. They satisfy the equations

$$V(z_\pm) = \mp \sqrt{c^2 p^2_\perp + m^2 e c^4 + \mathcal{E}}.$$  

(6.1.3)

This energy level crossing $\mathcal{E}$ is shown in Fig. 6.1 for the Sauter potential $V(z) \propto \tanh(z/\ell)$.

![Figure 6.1: Positive- and negative-energy spectra $\mathcal{E}_\pm(z)$ of Eq. (6.1.1) in units of $m_e c^2$, with $p_z = p_\perp = 0$ as a function of $\hat{z} = z/\ell$ for the Sauter potential $V_\pm(z)$ (6.3.4) for $\sigma_s = 5$.](image)

6.1.1 JWKB transmission probability for Klein–Gordon Field

The probability of quantum tunneling in the $z$-direction is most easily studied for a scalar field which satisfies the Klein–Gordon equation (3.4.3). If there is only an electric field in the $z$-direction which varies only along $z$, a vector potential with the only nonzero component (3.4.4) is chosen, and the ansatz

$$\phi(x) = e^{-iEt/\hbar} e^{ip_\perp x_\perp/\hbar} \phi_{p_\perp, \mathcal{E}}(z),$$

is made with a fixed momentum $p_\perp$ in the $x, y$-direction and an energy $\mathcal{E}$, and Eq. (3.4.3) becomes simply

$$\left[-\hbar^2 \frac{d^2}{dz^2} + p^2_\perp + m^2 e c^2 - \frac{1}{c^2} [\mathcal{E} - V(z)]^2\right] \phi_{p_\perp, \mathcal{E}}(z) = 0.$$  

(6.1.4)
By expressing the wave function \( \phi_{p_{\perp},E}(z) \) as an exponential

\[
\phi_{p_{\perp},E}(z) = C e^{iS_{p_{\perp},E}/\hbar},
\]

where \( C \) is some normalization constant, the wave equation becomes a Riccati equation for \( S_{p_{\perp},E} \):

\[
-ih\partial_z^2 S_{p_{\perp},E}(z) + [\partial_z S_{p_{\perp},E}(z)]^2 - p_{z}^2(z) = 0.
\]

(6.1.6)

where the function \( p_z(z) \) is the solution of the equation

\[
p_z^2(z) = \frac{1}{c^2} \left[ \mathcal{E} - V(z) \right]^2 - p_{\perp}^2 - m_e^2 c^2.
\]

(6.1.7)

The solution of Eq. (6.1.6) can be found iteratively as an expansion in powers of \( \hbar \):

\[
S_{p_{\perp},E}(z) = S_{p_{\perp},E}^{(0)}(z) - ihS_{p_{\perp},E}^{(1)}(z) + (-ih)^2 S_{p_{\perp},E}^{(2)}(z) + \ldots.
\]

(6.1.8)

Neglecting the expansion terms after \( S_{p_{\perp},E}^{(1)}(z) = -\log p_z^{1/2}(z) \) leads to the JWKB approximation for the wave functions of positive and negative energies (see e.g. [88,202])

\[
\phi_{p_{\perp},E}^{JWKB}(z) = \frac{C}{p_z^{1/2}(z)} e^{iS_{p_{\perp},E}^{(0)}(z)/\hbar}.
\]

(6.1.9)

where \( S_{p_{\perp},E}^{(0)}(z) \) is the eikonal

\[
S_{p_{\perp},E}^{(0)}(z) = \int^z p_z(z')dz'.
\]

(6.1.10)

Between the turning points \( z_- < z < z_+ \), whose positions are illustrated in Fig. 6.1, the momentum \( p_z(z) \) is imaginary and it is useful to define the positive function

\[
\kappa_z(z) \equiv \sqrt{p_{\perp}^2 + m_e^2 c^2 - \frac{1}{c^2} \left[ \mathcal{E} - V(z) \right]^2} \geq 0.
\]

(6.1.11)

The tunneling wave function in this regime is the linear combination

\[
\frac{C}{2(\kappa_z)^{1/2}} \exp \left[ -\frac{1}{\hbar} \int^z \kappa_z dz \right] + \frac{\bar{C}}{2(\kappa_z)^{1/2}} \exp \left[ +\frac{1}{\hbar} \int^z \kappa_z dz \right].
\]

(6.1.12)

Outside the turning points, i.e., for \( z < z_- \) and \( z > z_+ \), there exist negative energy and positive energy solutions for \( \mathcal{E} < \mathcal{E}_- \) and \( \mathcal{E} > \mathcal{E}_+ \) for positive \( p_z \).
6.1 Semi-classical description of pair production

On the left-hand side of $z_-$, the general solution is a linear combination of an incoming wave running to the right and outgoing wave running to the left:

$$\frac{\mathcal{C}_+}{(p_z)^{1/2}} \exp \left( \frac{i}{\hbar} \int_{z_z}^{z} p_z dz \right) + \frac{\mathcal{C}_-}{(p_z)^{1/2}} \exp \left( -\frac{i}{\hbar} \int_{z_z}^{z} p_z dz \right). \quad (6.1.13)$$

On the right-hand of $z_+$, there is only an outgoing wave

$$\frac{\mathcal{T}}{(p_z)^{1/2}} \exp \left( \frac{i}{\hbar} \int_{z_z}^{z} p_z dz \right), \quad (6.1.14)$$

The connection equations can be solved by

$$\mathcal{\bar{C}} = 0, \quad \mathcal{C}_\pm = e^{\pm i\pi/4} \mathcal{C}_{\mp}/2, \quad \mathcal{T} = \mathcal{C}_+ \exp \left( -\frac{1}{\hbar} \int_{z_+}^{z_-} \kappa_z dz \right). \quad (6.1.15)$$

The incident flux density is

$$j_z \equiv \frac{\hbar}{2m_e} \left[ \phi^* \partial_z \phi - (\partial_z \phi^*) \phi \right] = \frac{p_z}{m_e} \phi^* \phi = \frac{|\mathcal{C}_+|^2}{m_e}, \quad (6.1.16)$$

which can be written as

$$j_z(z) = v_z(z)n_-(z), \quad (6.1.17)$$

where $v_z(z) = p_z(z)/m_e$ is the velocity and $n_-(z) = \phi^*(z)\phi(z)$ the density of the incoming particles. Note that the $z$-dependence of $v_z(z)$ and $n_-(z)$ cancel each other. By analogy, the outgoing flux density is $|\mathcal{T}|^2/m_e$.

6.1.2 Rate of pair production

From the considerations given above, the transmission probability

$$P_{\mathrm{JWKB}} \equiv \frac{\text{transmitted flux}}{\text{incident flux}} \quad (6.1.18)$$

is found to be the simple exponential

$$P_{\mathrm{JWKB}}(p_\perp, \mathcal{E}) = \exp \left[ -\frac{2}{\hbar} \int_{z_-}^{z_+} \kappa_z dz \right]. \quad (6.1.19)$$

In order to derive from (6.1.18) the total rate of pair production in the electric field, it must be multiplied with the incident particle flux density at the entrance $z_-$ of the tunnel. The particle velocity at that point is $v_z = \partial \mathcal{E}/\partial p_z$, \hfill 1001
where the relation between \( E \) and \( z \) is given by Eq. (6.1.3):

\[
-1 = \frac{E - V(z_\perp)}{\sqrt{(cp_\perp)^2 + m_e^2 c^4}}.
\]

(6.1.20)

This must be multiplied with the particle density which is given by the phase space density \( d^3p / (2\pi \hbar)^3 \). The incident flux density at the tunnel entrance is therefore

\[
j_z(z_\perp) = D_s \int \frac{\partial E}{\partial p_\perp} \frac{d^2 p_\perp}{2\pi \hbar} \int \frac{dE}{2\pi \hbar} \frac{d^2 p_\perp}{(2\pi \hbar)^2},
\]

(6.1.21)

and the extra factor \( D_s \) is equal to 2 for electrons with two spin orientations\(^1\).

It is useful to change the variable of integration from \( z \) to \( \zeta(z) \) defined by

\[
\zeta(p_\perp, E; z) = \frac{E - V(z)}{\sqrt{(cp_\perp)^2 + m_e^2 c^4}}.
\]

(6.1.22)

and to introduce the notation for the electric field \( E(p_\perp, E; \zeta) \equiv E[z(p_\perp, E; \zeta)] \), where \( z(p_\perp, E; \zeta) \) is the inverse function of (6.1.22), the equations in (6.1.3) reduce to

\[
\zeta_-(p_\perp, E; z_-) = -1, \quad \zeta_+(p_\perp, E; z_+) = +1.
\]

(6.1.23)

In terms of the variable \( \zeta \), the JWKB transmission probability (6.1.19) can be rewritten as

\[
P_{\text{JWKB}}(p_\perp, E) = \exp \left\{ -\frac{2m_e^2 c^3}{\hbar E_0} \left[ 1 + \frac{(cp_\perp)^2}{m_e^2 c^4} \right] \int_{-1}^{1} d\zeta \frac{\sqrt{1 - \zeta^2}}{E(p_\perp, E; \zeta) / E_0} \right\}.
\]

(6.1.24)

Here a standard field strength \( E_0 \) has been introduced to make the integral in the exponent dimensionless, which is abbreviated by

\[
G(p_\perp, E) \equiv \frac{2}{\pi} \int_{-1}^{1} d\zeta \frac{\sqrt{1 - \zeta^2}}{E(p_\perp, E; \zeta) / E_0}.
\]

(6.1.25)

The first term in the exponent of (6.1.24) is equal to \( 2E_c / E_0 \).

At the semi-classical level, tunneling takes place only if the potential height is larger than \( 2m_e c^2 \) and for energies \( E \) for which there are two real turning points \( z_\pm \). The total tunneling rate is obtained by integrating over all incoming momenta and the total area \( V_\perp = \int dxdy \) of the incoming flux. The

\(^1\)By setting \( D_s \) equal to 1 one can obtain the tunneling result also for spin-0 particles although the Dirac picture is no longer applicable.
6.1 Semi-classical description of pair production

JWKB-rate per area is

\[
\frac{\Gamma_{\text{JWKB}}}{V_\perp} = D_s \int \frac{d\mathcal{E}}{2\pi \hbar} \int \frac{d^2p_\perp}{(2\pi \hbar)^2} \mathcal{P}_{\text{JWKB}}(p_\perp, \mathcal{E}). \tag{6.1.26}
\]

Using the relation following from (6.1.20)

\[
d\mathcal{E} = eE(z_-)dz_-,
\]

the alternative expression is obtained

\[
\frac{\Gamma_{\text{JWKB}}}{V_\perp} = D_s \int dz_- \int \frac{d^2p_\perp}{2\pi \hbar} (2\pi \hbar)^2 eE(z_-) \mathcal{P}_{\text{JWKB}}(p_\perp, \mathcal{E}(z_-)), \tag{6.1.28}
\]

where \(\mathcal{E}(z_-)\) is obtained by solving the differential equation (6.1.27).

The integral over \(p_\perp\) cannot be done exactly. At the semi-classical level, this is fortunately not necessary. Since \(E_c\) is proportional to \(1/\hbar\), the exponential in (6.1.24) restricts the transverse momentum \(p_\perp\) to be small of the order of \(\sqrt{\hbar}\), so that the integral in (6.1.28) may be calculated from an expansion of \(G(p_\perp, \mathcal{E})\) up to the order \(p_\perp^2\):

\[
G(p_\perp, \mathcal{E}) \approx \frac{2}{\pi} \int_{-1}^{1} d\zeta \frac{\sqrt{1 - \zeta^2}}{E(0, \mathcal{E}; \zeta)/E_0} \left[ 1 - \frac{1}{2} \frac{dE(0, \mathcal{E}, \zeta)/d\zeta}{E(0, \mathcal{E}, \zeta)} \delta + \ldots \right] = G(0, \mathcal{E}) + G_\delta(0, \mathcal{E}) \delta + \ldots, \tag{6.1.29}
\]

where \(\delta \equiv \delta(p_\perp) \equiv (cp_\perp)^2/(m^2 c^4)\), and

\[
G_\delta(0, \mathcal{E}) \equiv -\frac{1}{\pi} \int_{-1}^{1} d\zeta \frac{\zeta \sqrt{1 - \zeta^2}}{E'^2(0, \mathcal{E}; \zeta)/E_0} E'(0, \mathcal{E}, \zeta) = -\frac{1}{2} G(0, \mathcal{E}) + \frac{1}{\pi} \int_{-1}^{1} d\zeta \frac{\zeta^2}{\sqrt{1 - \zeta^2}} E(0, \mathcal{E}, \zeta)/E_0. \tag{6.1.30}
\]

The integral over \(p_\perp\) in (6.1.28) is approximately performed as follows:

\[
\int \frac{d^2p_\perp}{(2\pi \hbar)^2} e^{-\pi(E_c/E_0)(1+\delta)[G(0,\mathcal{E})+G_\delta(0,\mathcal{E})]} = \tag{6.1.31}
\]

\[
= \frac{m^2 c^2}{4\pi \hbar^2} \int_0^\infty d\delta e^{-\pi(E_c/E_0)[G(0,\mathcal{E})+\delta G(0,\mathcal{E})]} \approx \frac{eE_0}{4\pi^2 hcG(0,\mathcal{E})} e^{-\pi(E_c/E_0)G(0,\mathcal{E})},
\]

where

\[
\tilde{G}(0, \mathcal{E}) \equiv G(0, \mathcal{E}) + G_\delta(0, \mathcal{E}). \tag{6.1.32}
\]

The electric fields \(E(p_\perp, \mathcal{E}; \zeta)\) at the tunnel entrance \(z_-\) in the prefactor of (6.1.28) can be expanded similarly to first order in \(\delta\). If \(z^0_-\) denotes the solu-
tions of (6.1.20) at \( p_\perp = 0 \), it is found that for small \( \delta \):

\[
\Delta z_\perp \equiv z_\perp - z_0^\perp \approx \frac{m_e c^2 \delta}{E(z_0^\perp)}.
\] (6.1.33)

so that

\[
E(z_\perp) \approx E(z_0^\perp) - m_e c^2 \frac{E'(z_0^\perp) \delta}{E(z_0^\perp)}.
\] (6.1.34)

Here the extra term proportional to \( \delta \) can be neglected in the semi-classical limit since it gives a contribution to the prefactor of the order \( \hbar \). Thus the JWKB-rate (6.1.28) of pair production per unit area is obtained

\[
\frac{\Gamma_{\text{JWKB}}}{V_\perp} = \int dz \frac{\partial_z \Gamma_{\text{JWKB}}(z)}{V_\perp} \approx \int dz \frac{D_s c^2 E_0 E(z)}{8 \pi^2 \hbar^2 c} e^{-\pi \left( \frac{E_c}{E_0} \right) G(0,E(z))} (6.1.35)
\]

where \( E \) is short for \( z_0^\perp \). At this point it is useful to return from the integral \( \int dz_\perp e E(z_\perp) \) introduced in (6.1.28) to the original energy integral \( \int d\varepsilon \) in (6.1.26), so that the final result is

\[
\frac{\Gamma_{\text{JWKB}}}{V_\perp} \equiv \int d\varepsilon \frac{\partial_\varepsilon \Gamma_{\text{JWKB}}(z)}{V_\perp} \approx D_s \frac{e E_0}{4 \pi^2 \hbar c} \int d\varepsilon \frac{1}{2 \pi \hbar} e^{-\pi \left( \frac{E_c}{E_0} \right) G(0,E)} (6.1.36)
\]

where \( \varepsilon \)-integration is over all crossing energy levels.

These formula can be approximately applied to the three-dimensional case of electric fields \( \mathbf{E}(x,y,z) \) and potentials \( \mathbf{V}(x,y,z) \) at the points \( (x,y,z) \) where the tunneling length (3.5.6) is much smaller than the variation lengths \( \delta x_\perp \) of electric potentials \( V(x,y,z) \) in the \( xy \)-plane,

\[
\frac{1}{z_+ - z_-} \gg \frac{1}{V \delta x_\perp}.
\] (6.1.37)

At these points \( (x,y,z) \), one can arrange the tunneling path \( dz \) and momentum \( p_z(x,z,z) \) in the direction of electric field, corresponding perpendicular area \( d^2 V_\perp \equiv dxdy \) for incident flux and perpendicular momentum \( p_\perp \).

It is then approximately reduced to a one-dimensional problem in the region of size \( O(a) \) around these points. The surfaces \( z_\perp(x_\perp, y_\perp, \varepsilon) \) and \( z_+ = (x_+, y_+, \varepsilon) \) associated with the classical turning points are determined by Eqs. (6.1.22) and Eqs. (6.1.23) for a given energy \( \varepsilon \). The JWKB-rate of pair production (6.1.35) can then be expressed as a volume integral over the rate density per volume element

\[
\Gamma_{\text{JWKB}} = \int dx dy dz \frac{d^3 \Gamma_{\text{JWKB}}}{dx dy dz} = \int dt dx dy dz \frac{d^4 N_{\text{JWKB}}}{dt dx dy dz}.
\] (6.1.38)
On the right-hand side it is useful to rewrite the rate $\Gamma_{JWKB}$ as the time derivative of the number of pair creation events $dN_{JWKB}/dt$, so that one obtains an event density in four-space

$$
\frac{d^4N_{JWKB}}{dt \, dx \, dy \, dz} \approx D_s \frac{e^2 E_0 E(z)}{8\pi^3 \hbar G(0, E(z))} e^{-\pi(\epsilon_c/E_0)g(0, E(z))},
$$

(6.1.39)

Here $x, y$ and $z$ are related by the function $z = z_-(x, y, \epsilon)$ which is obtained by solving (6.1.27).

It is now useful to observe that the left-hand side of (6.1.39) is a Lorentz invariant quantity. In addition, it is symmetric under the exchange of time and $z$, and this symmetry will be exploited in the next section to relate pair production processes in a $z$-dependent electric field $E(z)$ to those in a time-dependent field $E(t)$.

Attempts to go beyond the JWKB results (6.1.35) or (6.1.36) require a great amount of work. Corrections will come from three sources:

I from the higher terms of order in $(\hbar)^n$ with $n > 1$ in the expansion (6.1.8) solving the Riccati equation (6.1.6).

II from the perturbative evaluation of the integral over $p_\perp$ in Eqs. (6.1.26) or (6.1.28) when going beyond the Gaussian approximation.

III from perturbative corrections to the Gaussian energy integral (6.1.36) or the corresponding $z$-integral (6.1.35).

All these corrections contribute terms of higher order in $\hbar$.

### 6.1.3 Including a smoothly varying $B(z)$-field parallel to $E(z)$

The results presented above can easily be extended for the presence of a constant magnetic field $B$ parallel to $E(z)$. Then the wave function factorizes into a Landau state and a spinor function first calculated by Sauter [20]. In the JWKB approximation, the energy spectrum is still given by Eq. (6.1.1), but the squared transverse momenta $p_\perp^2$ is quantized and must be replaced by discrete values corresponding to the Landau energy levels. From the known nonrelativistic levels for the Hamiltonian $p_\perp^2/2m_e$ one extracts immediately the replacements (3.5.17). Apart from the replacement (3.5.17), the JWKB calculations remain the same. Thus one must only replace the integration over the transverse momenta $\int d^2p_\perp/(2\pi\hbar)^2$ in Eq. (6.1.31) by the sum over all Landau levels with the degeneracy $eB/(2\pi\hbar c)$. Thus, the right-hand side becomes

$$
\frac{eB}{2\pi\hbar c} e^{-\pi(\epsilon_c/E_0)G(0, \epsilon)} \sum_{n, \sigma} e^{-\pi(B/E_0)(n+1/2+g\sigma)}\tilde{G}(0, \epsilon),
$$

(6.1.40)
where $g$ and $\hat{\sigma}$ are as in (3.5.17). The result is, for spin-0 and spin-1/2:

$$\frac{eE_0}{4\pi^2\hbar \tilde{G}(0, \mathcal{E})} e^{-\pi(E_0/E_0)\tilde{G}(0, \mathcal{E})} f_{0,1/2}(B \tilde{G}(0, \mathcal{E})/E_0)$$

(6.1.41)

where

$$f_0(x) \equiv \frac{\pi x}{\sinh \pi x}, \quad f_{1/2}(x) \equiv 2\frac{\pi x}{\sinh \pi x} \cosh \frac{\pi g x}{2}$$

(6.1.42)

In the limit $B \to 0$, Eq. (3.5.20) reduces to Eq. (6.1.31).

The result remains approximately valid if the magnetic field has a smooth $z$-dependence varying little over a Compton wavelength $\lambda_C$.

In the following only nonuniform electric fields without a magnetic field are considered.

### 6.2 Time-dependent electric fields

The semi-classical considerations given above can be applied with little change to the different physical situation in which the electric field along the $z$-direction depends only on time rather than $z$. Instead of the time $t$ itself it is better to work with the zeroth length coordinate $x_0 = ct$, as usual in relativistic calculations. As an intermediate step consider for a vector potential

$$A_\mu = (A_0(z), 0, 0, A_z(x_0)),$$

(6.2.1)

with the electric field

$$E = -\partial_z A_0(z) - \partial_0 A_z(x_0), \quad x_0 = ct.$$ (6.2.2)

The associated Klein–Gordon equation (3.4.3) reads

$$\left\{ \left[ i\hbar \partial_0 + \frac{e}{c} A_0(z) \right]^2 + \hbar^2 \partial^2_{z\perp} - \left[ i\hbar \partial_z + \frac{e}{c} A_z(x_0) \right]^2 - m^2_c c^2 \right\} \phi(x) = 0.$$ (6.2.3)

The previous discussion was valid under the assumption $A_z(x_0) = 0$, in which case the ansatz

$$\phi(x) = e^{-i\mathcal{E}t/\hbar} e^{ip_{\perp} \cdot x_{\perp}} \phi_{p_{\perp}, \mathcal{E}}(z),$$

led to the field equation (6.1.4). For the present discussion it is useful to write the ansatz as $\phi(x) = e^{-i p_0 x_0/\hbar} e^{ip_{\perp} \cdot x_{\perp}/\hbar} \phi_{p_{\perp}, p_0}(z)$ with $p_0 = \mathcal{E}/c$, and Eq. (6.1.4) in the form

$$\left\{ \frac{1}{c^2} \left[ \mathcal{E} - c \int^z dz' E(z') \right]^2 - p^2_{\perp} - m^2_c c^2 + \hbar^2 \frac{d^2}{dz^2} \right\} \phi_{p_{\perp}, p_0}(z) = 0.$$ (6.2.4)
6.3 Applications

Now assume that the electric field depends only on $x_0 = ct$. Then the ansatz $\phi(x) = e^{ipz/c}e^{ip_\perp x_\perp/\hbar} \phi_{p_\perp, p_z}(x_0)$ leads to the field equation

$$\left\{ -\hbar^2 \frac{d^2}{dx_0^2} - p_\perp^2 - m^2c^2 - \left[-p_z - \frac{e}{c} \int_{x_0}^{x_0} dx'_0 E(x'_0) \right]^2 \right\} \phi_{p_\perp, p_z}(x_0) = 0. \tag{6.2.5}$$

If Eq. (6.2.5) is compared with (6.2.4) it can be seen that one arises from the other by interchanging

$$z \leftrightarrow x_0, \quad p_\perp \rightarrow ip_{\perp}, \quad c \rightarrow ic, \quad E \rightarrow -iE. \tag{6.2.6}$$

With these exchanges, it may easy to calculate the decay rate of the vacuum caused by a time-dependent electric field $E(x_0)$ using the formulas derived above.

6.3 Applications

Now formulas (6.1.36) or (6.1.35) are applied to various external field configurations capable of producing electron–positron pairs.

6.3.1 Step-like electric field

First one checks the result for the original case of a constant electric field $E(z) \equiv eE_0$ where the potential energy is the linear function $V(z) = -eE_0z$. Here the function (6.1.25) becomes trivial

$$G(0, \mathcal{E}) = \frac{2}{\pi} \int_{-1}^{1} d\zeta \sqrt{1 - \zeta^2} = 1, \quad G_\delta(0, \mathcal{E}) = 0, \tag{6.3.1}$$

which is independent of $\mathcal{E}$ (or $z$). The JWKB-rate for pair production per unit time and volume is found from Eq. (6.1.35) to be

$$\frac{\Gamma_{\text{JWKB}}^{\text{EH}}}{V} \simeq D_s \frac{\alpha E_0^2}{2\pi^2\hbar} e^{-\pi E_0/c/E_0}. \tag{6.3.2}$$

where $V \equiv dz_\perp V_\perp$. This expression contains the exponential $e^{-\pi E_0/c/E_0}$ found by Sauter [20], and the correct prefactor as calculated by Heisenberg and Euler [7], and by Schwinger [25–27].

In order to apply the transformation rules (6.2.6) to obtain the analogous result for the constant electric field in time, one can rewrite Eq. (6.3.2) as

$$\frac{dN_{\text{JWKB}}}{dx_0 V} \simeq D_s \frac{\alpha E_0^2}{2\pi^2\hbar c} e^{-\pi E_0/c/E_0}. \tag{6.3.3}$$
where \( \frac{dN}{dx_0} = \Gamma_{\text{JWKB}}^E/c \) and \( N_{\text{JWKB}} \) is the number of pairs produced. Applying the transformation rules (6.2.6) to Eq. (6.3.3), one obtains the same formula as Eq. (6.3.2).

### 6.3.2 Sauter electric field

Let us now consider the nontrivial Sauter electric field localized within finite slab in the \( xy \)-plane with the width \( \ell \) in the \( z \)-direction. A field of this type can be produced, e.g., between two opposite charged conducting plates. The electric field \( E(z)\hat{z} \) in the \( z \)-direction and the associated potential energy \( V(z) \) are given by

\[
E(z) = E_0 / \cosh^2 \left( z / \ell \right), \quad V(z) = -\sigma_s m_e c^2 \tanh \left( z / \ell \right), \quad (6.3.4)
\]

where

\[
\sigma_s = e E_0 \ell / m_e c^2 = (\ell / \lambda_C) (E_0 / E_c). \quad (6.3.5)
\]

From now on natural units, in which energies are measured in units of \( m_e c^2 \), are used. Figure 6.1 shows the positive and negative energy spectra \( E_{\pm}(z) \) of Eq. (6.1.1) for \( p_z = p_\perp = 0 \) in particular the energy gap and energy level crossings. From Eq. (6.1.3) one finds the classical turning points

\[
z_{\pm} = \ell \arctanh \left( \frac{\varepsilon \pm \sqrt{1 + \delta}}{\sigma_s} \right) = \ell \frac{\ln \left( \frac{\sigma_s + \varepsilon \pm \sqrt{1 + \delta}}{\sigma_s - \varepsilon \mp \sqrt{1 + \delta}} \right)}{2}. \quad (6.3.6)
\]

Tunneling is possible for all energies satisfying

\[
-\sqrt{1 + \delta} + \sigma_s \geq \varepsilon \geq \sqrt{1 + \delta} - \sigma_s, \quad (6.3.7)
\]

for the strength parameter \( \sigma_s > \sqrt{1 + \delta} \).

One may invert Eq. (6.1.22) to find the relation between \( \zeta \) and \( z \):

\[
z = z(p_\perp, \varepsilon; \zeta) = \ell \arctanh \left( \frac{\varepsilon + \zeta \sqrt{1 + \delta}}{\sigma_s} \right) = \ell \frac{\ln \left( \frac{\sigma_s + \varepsilon + \zeta \sqrt{1 + \delta}}{\sigma_s - \zeta \sqrt{1 + \delta}} \right)}{2}. \quad (6.3.8)
\]

In terms of the function \( z(p_\perp, \varepsilon; \zeta) \), the Eq. (6.3.6) reads simply \( z_{\pm} = z(p_\perp, \varepsilon; \pm 1) \).

Inserting (6.3.8) into the equation for \( E(z) \) in Eq. (6.3.4), one obtains

\[
E(z) = E_0 \left[ 1 - \left( \frac{\zeta \sqrt{1 + \delta} - \varepsilon}{\sigma_s} \right)^2 \right] = E(p_\perp, \varepsilon; \zeta). \quad (6.3.9)
\]

\( G(0, \varepsilon) \) and \( G_\delta(0, \varepsilon) \) of Eqs. (6.1.25), (6.1.29) and (6.1.30) are calculated:

\[
G(0, \varepsilon) = 2\sigma_s^2 - \sigma_s \left[ (\sigma_s - \varepsilon)^2 - 1 \right]^{1/2} - \sigma_s \left[ (\sigma_s + \varepsilon)^2 - 1 \right]^{1/2}, \quad (6.3.10)
\]
and

\[
G(0, \mathcal{E}) + G_\delta(0, \mathcal{E}) = \frac{\sigma_s}{2} \left\{ \left[ (\sigma_s - \mathcal{E})^2 - 1 \right]^{-1/2} + \left[ (\sigma_s + \mathcal{E})^2 - 1 \right]^{-1/2} \right\}.
\tag{6.3.11}
\]

Substituting the functions \(G(0, \mathcal{E})\) and \(G_\delta(0, \mathcal{E})\) into Eqs. (6.1.35) and (6.1.36), one obtains the general expression for the pair production rate per volume slice at a given tunnel entrance point \(z_- (\mathcal{E})\) or the associated energy \(\mathcal{E}(z_-)\). The pair production rate per area is obtained by integrating over all slices permitted by the energy inequality (6.3.7).

In Fig. 6.2, the slice dependence of the integrand in the tunneling rate (6.1.35) for the Sauter potential (6.3.4) is shown and compared with the constant field expression (6.3.2) of Euler and Heisenberg, if it is evaluated at the \(z\)-dependent electric field \(E(z)\). This is done once as a function of the tunnel entrance point \(z\) and once as a function of the associated energy \(\mathcal{E}\). On each plot, the difference between the two curves illustrates the nonlocality of the tunneling process\(^2\). The integral is dominated by the region around \(\mathcal{E} \sim 0\),

\[\text{Figure 6.2: The slice dependence of the integrand in the tunneling rate (6.1.35) for the Sauter potential (6.3.4) is plotted: left, as a function of the tunnel entrance } z \text{ (compare with numeric results plotted in Fig. 1 of Ref. [229]); right, as a function of the associated energy } \mathcal{E}, \text{ which is normalized by the Euler-Heisenberg rate (6.3.2). The dashed curve in left figure shows the Euler-Heisenberg expression (6.3.2) evaluated for the } z\text{-dependent field } E(z) \text{ to illustrate the nonlocality of the production rate. The dashed curve in right figure shows the Euler-Heisenberg expression (6.3.2) which is independent of energy-level crossing } \mathcal{E}. \text{ The dimensionless parameters are } \sigma_s = 5, \ E_0 / E_c = 1.\]

where the tunneling length is shortest [see Fig. 6.1] and tunneling probability is largest. Both functions \(G(0, \mathcal{E})\) and \(G_\delta(0, \mathcal{E})\) have a symmetric peak at

\(^2\)Note that omitting the \(z\)-integral in the rate formula (6.1.35) does not justify calling the result a “local production rate”, as done in the abstract of Ref. [229]. The result is always nonlocal and depends on all gradients of the electric field.
\( \mathcal{E} = 0 \). Around the peak they can be expanded in powers of \( \mathcal{E} \) as

\[
G(0, \mathcal{E}) = 2[\sigma_s^2 - \sigma_s(\sigma_s^2 - 1)^{1/2}] + \frac{\sigma_s}{(\sigma_s^2 - 1)^{3/2}} \mathcal{E}^2 + \mathcal{O}(\mathcal{E}^4) = \]

\[
= G_0(\sigma_s) + \frac{1}{2} G_2(\sigma_s) \mathcal{E}^2 + \mathcal{O}(\mathcal{E}^4),
\]  
(6.3.12)

and

\[
G(0, \mathcal{E}) + G_0(0, \mathcal{E}) = \frac{\sigma_s}{(\sigma_s^2 - 1)^{1/2}} + \frac{1}{2} \frac{1}{(\sigma_s^2 - 1)^{5/2}} \mathcal{E}^2 + \mathcal{O}(\mathcal{E}^4) = \]

\[
= \mathcal{G}_0(\sigma_s) + \frac{1}{2} \mathcal{G}_2(\sigma_s) \mathcal{E}^2 + \mathcal{O}(\mathcal{E}^4). \quad \text{(6.3.13)}
\]

The exponential \( e^{-\pi G(0, \mathcal{E})E_c/E_0} \) has a Gaussian peak around \( \mathcal{E} = 0 \) whose width is of the order of \( 1/E_c \propto \hbar \). This implies that in the semi-classical limit, one may perform only a Gaussian integral and neglect the \( \mathcal{E} \)-dependence of the prefactor in (6.1.36). Recalling that \( \mathcal{E} \) in this section is in natural units with \( m_e c^2 = 1 \), one should replace \( \int d\mathcal{E} \) by \( m_e c^2 \int d\mathcal{E} \) and can perform the integral over \( \mathcal{E} \) approximately as follows

\[
\frac{\Gamma_{\text{JWKB}}}{V_\perp} \simeq D_s \frac{e E_0 m_e c^2}{4 \pi^2 \hbar c} \frac{1}{G_0} e^{-\pi (E_c/E_0) G_0} \int \frac{d\mathcal{E}}{2 \pi \hbar} e^{-\pi (E_c/E_0) G_0} \mathcal{E}^2/2 = \]

\[
= D_s \frac{e E_0}{4 \pi^2 \hbar c} \frac{1}{G_0} \frac{e^{-\pi (E_c/E_0) G_0}}{2 \pi \hbar \sqrt{G_0'' E_c / 2 E_0}}. \quad \text{(6.3.14)}
\]

For convenience, the limits of integration over \( E \) is extended from the interval \((-1 + \sigma_s, 1 - \sigma_s)\) to \((-\infty, \infty)\). This introduces exponentially small errors and can be ignored.

Using the relation (6.3.5) one may replace \( e E_0 m_e c^2 / \hbar c \) by \( e^2 E_0^2 2\ell / \sigma_s \), and obtain

\[
\frac{\Gamma_{\text{JWKB}}[\text{total}]}{V_\perp \ell} \simeq D_s \frac{e E_0^2}{2 \pi \hbar} \sqrt{\frac{E_0}{E_c}} \frac{(\sigma_s^2 - 1)^{5/4}}{\sigma_s^{5/2}} e^{-\pi G_0(\sigma_s) E_c/E_0}. \quad \text{(6.3.15)}
\]

This approximate result agrees \(^3\) with that obtained before with a different, somewhat more complicated technique proposed by Dunne and Schubert \cite{48} after the fluctuation determinant was calculated exactly in \cite{49} with the help of the Gelfand–Yaglom method, see Section 2.2 in Ref. \cite{202}. The advantage of knowing the exact fluctuation determinant could not, however, be fully exploited since the remaining integral was calculated only in the saddle

\(^3\)See Eq. (63) of Dunne and Schubert \cite{48}, and replace there \( \tilde{\gamma} \rightarrow 1/\sigma_s \). It agrees also with the later paper by Dunne et al. \cite{49} apart from a factor 2.
point approximation. The rate (6.3.15) agrees with the leading term of the expansion (42) of Kim and Page [53]. Note that the higher expansion terms calculated by the latter authors do not yet lead to proper higher order results since they are only of type II and III in the list after Eq. (6.1.39). The terms of equal order in $\bar{\hbar}$ in the expansion (6.1.8) of the solution of the Riccati equation are still missing.

Using the transformation rules (6.2.6), it is straightforward to obtain the pair production rate of the Sauter type of electric field depending on time rather than space. According to the transformation rules (6.2.6), one has to replace $\ell \to c\delta T$, where $\delta T$ is the characteristic time over which the electric field acts—the analog of $\ell$ in (6.3.4). Thus the field (6.3.4) becomes

$$E(t) = E_0 / \cosh^2 (t / \delta T), \quad V(t) = -\bar{s}_s m_e c^2 \tanh (t / \delta T).$$

(6.3.16)

According to the same rules, one must also replace $\sigma_s \to i\tilde{\sigma}_s$, where

$$\tilde{\sigma}_s \equiv eE_0 \delta T / m_e c.$$

(6.3.17)

This brings $G(0, \sigma_s)$ of Eq. (6.3.10) to the form

$$G(0, \sigma_s) \to G^I(\tilde{\sigma}_s) = 2[\tilde{\sigma}_s (\tilde{\sigma}_s^2 - 1)^{1/2} - \tilde{\sigma}_s^2],$$

(6.3.18)

and yields the pair production rate

$$\Gamma_{\text{JWKB[total]}} \simeq D_s \frac{\alpha E_0^2}{2\pi^2 \hbar} \sqrt{\frac{E_0}{E_c}} \left( \frac{\bar{s}_s^2 + 1}{\tilde{\sigma}_s^2} \right)^{5/4} e^{-\pi G^I(\sigma_s)E_c / E_0},$$

(6.3.19)

where $\Gamma_{\text{JWKB[total]}} = \partial N_{\text{JWKB}} / \partial z$ is the number of pairs produced per unit thickness in a spatial shell parallel to the $xy$-plane. This is in agreement with Ref. [49].

Note also that the constant field result (6.3.2) of Euler and Heisenberg cannot be deduced from (6.3.15) by simply taking the limit $\ell \to \infty$ as one might have expected. The reason is that the saddle point approximation (6.3.14) to the integral (6.1.36) becomes invalid in this limit. Indeed, if $\ell \propto \sigma_s$ is large in Eqs. (6.3.10) and (6.3.11), these become

$$G(0, \bar{\epsilon}) \to G(0, \epsilon) + G(0, \epsilon) \to \frac{1}{1 - \epsilon^2 / \sigma_s^2},$$

(6.3.20)

and the integral in (6.1.36) becomes approximately

$$e^{-\pi (E_c / E_0)} \int_{-\sigma_s}^{+\sigma_s} \frac{d\bar{\epsilon}}{2\pi \hbar} \left( 1 - \bar{\epsilon}^2 / \sigma_s^2 \right) e^{-\pi (E_c / E_0)(\bar{\epsilon}^2 / \sigma_s^2)}$$

(6.3.21)

For not too large $\ell \propto \sigma_s$, the integral can be evaluated in the leading Gaussian
corresponding to the previous result (6.3.15) for large-$\sigma_s$. For a constant field, however, where the integrands becomes flat, the Gaussian approximation is no longer applicable. Instead one must first set $\sigma_s \to \infty$ in the integrand of (6.3.21), making it constant. Then the integral (6.3.21) becomes
\begin{equation}
e^{-\pi(E_c/E_0)^2}e^{-\pi E_c/E_0}\frac{2\ell E_0}{m_e c^2}\frac{2\pi \hbar}{}.\tag{6.3.23}
\end{equation}
Inserting this into (6.1.36) one recovers the constant field result (6.3.2). One must replace $2\ell$ by $L$ to comply with the relation (6.1.27) from which one obtains
\begin{equation}\int d\mathcal{E} = \int dzeE(z) = eE_0 \int dz/ \cosh^2(z/\ell) = 2\ell eE_0 = LeE_0.\tag{6.3.24}
\end{equation}
In order to see the boundary effect on the pair production rate, this section is closed with a comparison between pair production rates in the constant field (6.3.2) and Sauter field (6.3.15) for the same field strength $E_0$ in the volume $V_{\perp \ell}$. The ratio $R_{\text{rate}}$ of pair production rates (6.3.15) and (6.3.2) in the volume $V_{\perp \ell}$ is defined as
\begin{equation}R_{\text{rate}} = \sqrt{\frac{E_0}{E_c}}e^{-\pi E_c/E_0}\frac{\left(\sigma_s^2 - 1\right)^{5/4}}{\sigma_s^{5/2}\sigma_G}\frac{E_c}{E_0}.\tag{6.3.24}
\end{equation}
The soft boundary of the Sauter field (6.3.4) reduces its pair production rate with respect to the pair production rate (6.3.2) computed in a constant field of width $L = 2\ell$. The reduction is shown quantitatively in Fig. 6.3, where curves are plotted for the rates (6.3.2) and (6.3.15), and and for their ratio (6.3.24) at $E_0 = E_c$ and $\sigma_s = \ell/\lambda_C$ [recall (6.3.5)]. One can see that the reduction is significant if the width of the field slab shrinks to the size of a Compton wavelength $\lambda_C$.

### 6.3.3 Constant electric field in half space

As a second application consider an electric field which is zero for $z < 0$ and goes to $-E_0$ over a distance $\ell$ as follows:
\begin{align}
E(z) &= -\frac{E_0}{2} \left[ \tanh \left( \frac{z}{\ell} \right) + 1 \right], \quad V(z) = -\frac{\sigma_s}{2} m_e c^2 \left\{ \ln \cosh \left( \frac{z}{\ell} \right) + \frac{z}{\ell} \right\}. \tag{6.3.25}
\end{align}
\footnote{This treatment is analogous to that of the translational degree of freedom in instanton calculations in Section 17.3.1 of [202] [see in particular Eq. (17.112)].}
Figure 6.3: Left: Ratio $R_{\text{rate}}$ defined in Eq. (6.3.24) is plotted as function of $\sigma_s$ in the left figure. Right: Number of pairs created in slab of Compton width per area and time as functions of $\sigma_s$. Upper curve is for the constant field (6.3.2), lower for the Sauter field (6.3.15)). Both plots are for $E_0 = E_c$ and $\sigma_s = \ell/\lambda_C$.

where $\sigma_s \equiv eE_0\ell/m_e c^2$. In Fig. 6.4, the positive and negative energy spectra $\mathcal{E}_\pm(z)$ defined by Eq. (6.1.1) for $p_z = p_\perp = 0$ are plotted to show energy gap and level crossing. From Eq. (6.1.3) one finds now the classical turning points

Figure 6.4: Energies (6.1.1) for a soft electric field step $E(z)$ of Eq. (6.3.25) and the potentials $V_\pm(z)$ (6.3.25) for $\sigma_s = 5$. Positive and negative-energies $\mathcal{E}_\pm(z)$ of Eq. (6.1.1) are plotted for $p_z = p_\perp = 0$ as functions of $\hat{z} = z/\ell$.

[instead of (6.3.6)]

$$z_\pm = \ell/2 \ln \left[ 2e^{(\mathcal{E}_\pm \sqrt{1 + \delta})/\sigma_s} - 1 \right].$$  

(6.3.26)
For tunneling to take place, the energy $\mathcal{E}$ has to satisfy

$$\mathcal{E} \leq \sqrt{1 + \delta} - \sigma_s \ln 2,$$

(6.3.27)

and $\sigma_s$ must be larger than $\sqrt{1 - \delta \zeta}$. Expressing $z/\ell$ in terms of $\zeta$ as

$$z = z(p_\perp, \mathcal{E}; \zeta) = \frac{\ell}{2} \ln \left[ \frac{2e^{(\mathcal{E} + \zeta \sqrt{1 + \delta})/\sigma_s}}{1 - 1/\sigma_s} \right],$$

(6.3.28)

so that $z_\pm = z(p_\perp, \mathcal{E}; \pm 1)$, one finds the electric field in the form

$$E(z) = E_0 \left[ 1 - \frac{1}{2} e^{(\zeta \sqrt{1 + \delta} - \mathcal{E})/\sigma_s} \right] \equiv E(p_\perp, \mathcal{E}; \zeta).$$

(6.3.29)

Inserting this into Eq. (6.1.25) and expanding $E_0 / E(p_\perp, \mathcal{E}; \zeta)$ in powers one obtains

$$G(p_\perp, \mathcal{E}) = 1 + \sum_{n=1}^{\infty} \frac{e^{-n\mathcal{E}/\sigma_s}}{2^n} \frac{2}{\pi} \int_{1}^{1} d\zeta \sqrt{1 - \zeta^2} e^{n\zeta/\sigma_s} =
= 1 + \sum_{n=1}^{\infty} e^{-n\mathcal{E}/\sigma_s} I_1(n\sqrt{1 + \delta}/\sigma_s),$$

(6.3.30)

where $I_1(x)$ is a modified Bessel function. Expanding $I_1(n\sqrt{1 + \delta}/\sigma_s)$ in powers of $\delta$:

$$I_1(n\sqrt{1 + \delta}/\sigma_s) = I_1(n/\sigma_s) + (n/4\sigma_s)[I_0(n/\sigma_s) + I_2(n/\sigma_s)]\delta + \ldots$$

(6.3.31)

one can identify

$$G(0, \mathcal{E}) = 1 + \sum_{n=1}^{\infty} e^{-n\mathcal{E}/\sigma_s} I_1(n/\sigma_s),$$

(6.3.32)

$$G(0, \mathcal{E}) + G_\delta(0, \mathcal{E}) = 1 + \frac{1}{2} \sum_{n=1}^{\infty} e^{-n\mathcal{E}/\sigma_s} [(n/\sigma_s)I_0(n/\sigma_s) - I_1(n/\sigma_s)].$$

(6.3.33)

The integral over $\mathcal{E}$ in Eq. (6.1.36) starts at $\mathcal{E}_< = 1 - \sigma_s \log 2$ where the integrand rises from 0 to 1 as $\mathcal{E}$ exceeds a few units of $\sigma_s$. The derivative of $e^{-\pi(E_\zeta/E_0)G(0, \mathcal{E})}$ drops from 1 to $e^{-\pi(E_\zeta/E_0)}$ over this interval. Hence the derivative $\partial_\mathcal{E} e^{-\pi(E_\zeta/E_0)G(0, \mathcal{E})}$ is peaked around some value $\tilde{\mathcal{E}}$. Thus the integral $\int d\mathcal{E} e^{-\pi(E_\zeta/E_0)G(0, \mathcal{E})}$ is performed by parts as

$$\int d\mathcal{E} e^{-\pi(E_\zeta/E_0)G(0, \mathcal{E})} = \mathcal{E} e^{-\pi(E_\zeta/E_0)G(0, \mathcal{E})} \bigg|_{\tilde{\mathcal{E}} < } - \int d\mathcal{E} \mathcal{E} \partial_\mathcal{E} e^{-\pi(E_\zeta/E_0)G(0, \mathcal{E})}.$$

(6.3.34)
The first term can be rewritten with the help of \( d\mathcal{E} = eE_0dz \) as \( e^{-\pi(E_c/E_0)|E_0|\ell/2} \), thus giving rise to the decay rate (6.3.2) in the volume \( V_\perp \ell/2 \), and the second term gives only a small correction to this. The second term in Eq. (6.3.34) shows that the boundary effects reduce the pair production rate compared with the pair production rate (6.3.2) in the constant field without any boundary.

### 6.4 Tunneling into bound States

We turn now to the case in which instead of an outgoing wave as given by (6.1.14) there is a bound state. A linearly rising electric field whose potential is harmonic is considered:

\[
E(z) = E_0 \left( \frac{z}{\lambda_C} \right), \quad V(z) = \frac{eE_0\lambda_C}{2} \left( \frac{z}{\lambda_C} \right)^2.
\]  

(6.4.1)

It will be convenient to parametrize the field strength \( E_0 \) in terms of a dimensionless reduced electric field \( \epsilon \) as \( E_0 = \epsilon \hbar c/e\lambda_c^2 = eE_\epsilon \). In Fig. 6.5, the positive and negative energy spectra \( \mathcal{E}_\pm(z) \) defined by Eq. (6.1.1) for \( p_z = p_\perp = 0 \) are plotted to show energy gap and level crossing for \( \epsilon > 0 \) (left) and \( \epsilon < 0 \) (right). If \( \epsilon \) is positive, Eq. (6.1.3) yields for \( z > 0 \),

\[
z_{\pm} = \lambda_C \sqrt{\frac{2}{\epsilon} \left( \mathcal{E} \mp \sqrt{1 + \delta} \right)^{1/2}}, \quad z_+ < z_-,
\]

(6.4.2)

and mirror-reflected turning points for \( z < 0 \), obtained by exchanging \( z_\pm \rightarrow -z_\pm \) in (6.4.2). Negative energy electrons tunnel into the potential well \(-z_+ < \)

---

**Figure 6.5:** Positive- and negative-energy spectra \( \mathcal{E}_\pm(z) \) of Eq. (6.1.1) for \( p_z = p_\perp = 0 \) as a function of \( \hat{z} \equiv z/\lambda_c \) for the linearly rising electric field \( E(z) \) with the harmonic potential (6.4.1). The reduced field strengths are \( \epsilon = 2 \) (left figure) and \( \epsilon = -2 \) (right figure). On the left, bound states are filled and positrons escape to \( z = \pm\infty \). On the right, bound electrons with negative energy tunnel out of the well and escape with increasing energy to \( z = \pm\infty \).
z < +z+, where \( E \geq E_+ \), forming bound states. The associated positrons run off to infinity.

### 6.4.1 JWKB transmission probability

Due to the physical application to be discussed in the next section, here only the tunneling process for \( \epsilon > 0 \) on the left-hand side of Fig. 6.5 will be studied. One can consider the regime \( z < 0 \) with the turning points \( -z_- < -z_+ \). The incident wave and flux for \( z < -z_- \) pointing in the positive \( z \)-direction are given by Eqs. (6.1.13) and (6.1.16). The wave function for \( -z_- < z < -z_+ \) has the form Eq. (6.1.12) with the replacement \( z \rightarrow -z_- \). The transmitted wave is now no longer freely propagating as in (6.1.14), but describes a bound state of a positive energy electron:

\[
\phi_{E_n}(z) = \frac{B}{(p_z)^{1/2}} \cos \left[ \frac{1}{\hbar} \int_{-z_+}^{-z_-} p_z dz - \frac{\pi}{4} \right]. \tag{6.4.3}
\]

The Sommerfeld quantization condition

\[
\frac{1}{\hbar} \int_{-z_+}^{+z_+} p_z dz = \pi(n + \frac{1}{2}), \quad n = 0, 1, 2, \ldots . \tag{6.4.4}
\]

fixes the energies \( E_n \). The connection rules for the wave functions (6.1.12) and (6.4.3) at the turning point \( -z_+ \) determine

\[
B = \sqrt{2} C_+ e^{-i\pi n} \exp \left[ -\frac{1}{\hbar} \int_{-z_-}^{-z_+} \kappa_z dz \right]. \tag{6.4.5}
\]

Assuming the states \( \phi_{E_n}(z) \) to be initially unoccupied, the transmitted flux to these states at the classical turning point \( -z_+ \) is

\[
\frac{\hbar}{m_e} \phi_{E_n}(z) \partial_z \phi_{E_n}(z) \bigg|_{z \rightarrow -z_+} = \frac{|B|^2}{(2m_e)} = \frac{|C_+|^2}{m_e} \exp \left[ -\frac{2}{\hbar} \int_{-z_-}^{-z_+} \kappa_z dz \right]. \tag{6.4.6}
\]

From Eqs. (6.1.16), (6.4.6), and (6.1.18) one then finds the JWKB transmission probability for positrons to fill these bound states leaving a positron outside:

\[
P_{\text{JWKB}}(p_\perp, E_n) = \exp \left[ -\frac{2}{\hbar} \int_{-z_-}^{-z_+} \kappa_z dz \right], \tag{6.4.7}
\]

which has the same form as Eq. (3.5.10). The same result is obtained once more for \( z > 0 \) with turning points \( z_+ < z_- \), which can be obtained from (6.4.7) by the mirror reflection \( -z_\pm \leftrightarrow z_\pm \).
6.4.2 Energy spectrum of bound states

From Eq. (6.1.7) for $p_z$ and Eq. (6.4.1) for the potential $V(z)$, the eikonal (6.4.4) is calculated to determine the energy spectrum $\epsilon_n$ of bound states

$$\frac{1}{\hbar} \int_{-z_+}^{+z_+} p_z dz = 2\frac{\epsilon}{\lambda^2_C} \int_{0}^{z_+} \left[ (z^2 - z^2_+) (z^2 - z^2_-) \right]^{1/2} dz$$

$$= 2\epsilon z_+ \frac{3}{\lambda^3_C} \left[ (z^2_+ + z^2_-) E(t) - (z^2_- - z^2_+) K(t) \right], \tag{6.4.8}$$

where $E(t), K(t)$ are complete elliptical integrals of the first and second kind, respectively, and $t \equiv z_+/z_-$. The Sommerfeld quantization rule (6.4.4) becomes

$$8 \left( \frac{2(\epsilon_n - \sqrt{1+\delta})}{\epsilon} \right)^{1/2} \left[ \epsilon_n E(t_n) - (\sqrt{1+\delta}) K(t_n) \right] = \pi(n + 1/2). \tag{6.4.9}$$

$t_n \equiv \left( \frac{\epsilon_n - \sqrt{1+\delta}}{\epsilon_n + \sqrt{1+\delta}} \right)^{1/2}$.

For any given transverse momentum $p_\perp = \sqrt{\delta}$, this determines the discrete energies $\epsilon_n$.

6.4.3 Rate of pair production

By analogy with Eqs. (6.1.26) and (6.1.36), the transmission probability (6.4.7) must now be integrated over all incident particles with the flux (6.1.21) to yield the rate of pair production:

$$\Gamma_{JWKB} \frac{V_\perp}{\bar{V}} = 2D_s \sum_n \frac{\omega_n}{2\pi} \int \frac{d^2 p_\perp}{(2\pi \hbar)^2} \mathcal{P}_{JWKB}(p_\perp, \epsilon_n), \tag{6.4.10}$$

$$\approx 2D_s \frac{|eE_0|}{4\pi^2 \hbar c} \sum_n \frac{\omega_n}{2\pi} \frac{1}{G(0, \epsilon_n) + G(0, \epsilon_n)} e^{-\pi(\epsilon/E_0)G(0, \epsilon_n)}. \tag{6.4.11}$$

In obtaining these expressions one has used the energy conservation law to perform the integral over $\epsilon$. This receives only contributions for $\epsilon = \epsilon_n$ where $\int d\epsilon = \omega_n \hbar \equiv \epsilon_n - \epsilon_{n-1}$. The factor 2 accounts for the equal contributions from the two regimes $z > 0$ and $z < 0$. The previous relation (6.1.27) is now replaced by

$$\omega_n \hbar = |eE(z^n_+) \Delta z^n_|. \tag{6.4.12}$$
Using Eq. (6.1.22) and expressing $z/\lambda_C > 0$ in terms of $\zeta$ as

$$z = z(p_\perp, \mathcal{E}_n; \zeta) = \lambda_C \sqrt{\frac{2}{\epsilon}} \left( \mathcal{E}_n - \zeta \sqrt{1 + \delta} \right)^{1/2}, \quad (6.4.13)$$

one calculates $z_\pm = z(p_\perp, \mathcal{E}_n; \pm 1)$, and find the electric field in the form

$$E(z) = E_0 \sqrt{\frac{2}{\epsilon}} \left( \mathcal{E}_n - \zeta \sqrt{1 + \delta} \right)^{1/2} \equiv E(p_\perp, \mathcal{E}_n; \zeta). \quad (6.4.14)$$

Inserting this into Eq. (6.1.25) one obtains

$$G(p_\perp, \mathcal{E}_n) = \frac{2}{\pi} \sqrt{\frac{\epsilon}{2}} \int_{-1}^{1} d\zeta \frac{\sqrt{1 - \zeta^2}}{[\mathcal{E}_n - \zeta \sqrt{1 + \delta}]^{1/2}}$$

$$= \frac{8}{3\pi} \sqrt{\frac{\epsilon}{2}} \left( \mathcal{E}_n^\delta + 1 \right)^{1/2} \left[ (1 - \mathcal{E}_n^\delta) K(q_n^\delta) + \mathcal{E}_n^\delta E(q_n^\delta) \right] \quad (6.4.15)$$

where $\mathcal{E}_n^\delta \equiv \mathcal{E}_n/(1 + \delta)^{1/2}$ and $q_n^\delta = \sqrt{2/(\mathcal{E}_n^\delta + 1)}$. Expanding $G(p_\perp, \mathcal{E}_n)$ in powers of $\delta$ one finds the zeroth order term

$$G(0, \mathcal{E}_n) = \frac{8}{3\pi} \sqrt{\frac{\epsilon}{2}} \left( \mathcal{E}_n + 1 \right)^{1/2} \left[ (1 - \mathcal{E}_n) K(q_n) + \mathcal{E}_n E(q_n) \right] \quad (6.4.16)$$

and the derivative

$$G_\delta(0, \mathcal{E}_n) = \frac{\sqrt{\epsilon}}{3\pi} q_n \left[ (4 - 5q_n + \mathcal{E}_n(7 - 6q_n)) E(q_n) + \right.$$

$$\left. + \ (1 - \mathcal{E}_n - 7\mathcal{E}_n^2)(q_n - 1) K(q_n) \right]. \quad (6.4.17)$$

where $q_n \equiv \sqrt{2/(\mathcal{E}_n + 1)}$. 

1018
7 Phenomenology of electron–positron pair creation and annihilation

7.1 $e^+e^-$ annihilation experiments in particle physics

The $e^+e^- \rightarrow \gamma + \gamma$ process predicted by Dirac was almost immediately observed [3]. The $e^+e^-$ annihilation experiments have since became possibly the most prolific field of research in the active domain of particle physics. The Dirac pair annihilation process (3.0.1) has no energy threshold and the energy release in the $e^+e^-$ collision is larger than $2m_e c^2$. This process is the only one in the limit of low energy. As the $e^+e^-$ energy increases the collision produces not only photons through the Dirac process (5.9.2) but also other particles. For early work in this direction, predicting resonances for pions, K-mesons etc., see [231]. Production of such particles are achievable and precisely conceived in experimental particle physics, but hardly possible with the vacuum polarization process. In particular when the energy in the center of mass is larger than twice muon mass $m_\mu$ about 100 times electron mass, the electron and positron electromagnetically annihilate into two muons $e^+e^- \rightarrow \mu^+\mu^-$ via the intermediation of a virtual photon. The cross-section in the center of mass frame is given by [232]

\[
\sigma_{e^+e^-\rightarrow \mu^+\mu^-} = \frac{16\pi^2\alpha^2(\hbar c)^2}{q^2_{\text{cm}}} \text{Im} \tilde{\omega}_{\mu^+\mu^-}(q^2_{\text{cm}}) = \frac{4\pi\alpha^2(\hbar c)^2}{3q^2_{\text{cm}}} = \frac{86.8\text{nb}}{q^2_{\text{cm}}(\text{GeV})^2}
\]

(7.1.1)

where $\tilde{\omega}_{\mu^+\mu^-}(q^2_{\text{cm}})$ is the muon part of the vacuum polarization tensor and $q^2_{\text{cm}} = c^2(p_+ + p_-)^2/4$ the square of energy of the center of mass frame, where $p_\pm$ are 4-momenta of positron and electron. At very high energy $m_\mu^2/q^2_{\text{cm}} \rightarrow 0$, $\text{Im} \tilde{\omega}_{\mu^+\mu^-}(q^2_{\text{cm}}) \rightarrow 1/(12\pi)$.

At very high energies of the order of several GeV, electron and positron electromagnetically annihilate into hadrons, whose cross-section has the same structure as the cross-section (7.1.1) with $\tilde{\omega}_{\mu^+\mu^-}(q^2_{\text{cm}})$ replaced by the hadron...
part of the vacuum polarization tensor $\bar{\omega}_{\text{hadrons}}(q^2_{\text{cm}})$,

$$
\sigma_{e^+e^-\rightarrow \text{hadron}} = \frac{16\pi^2\alpha^2(h\epsilon)^2}{q^2_{\text{cm}}} \text{Im} \bar{\omega}_{\text{hadrons}}(q^2_{\text{cm}}).
$$

(7.1.2)

The two cross-sections (7.1.1) and (7.1.2) are comparable, of the order of a few tens of nanobarns ($10^{-33}\text{cm}^2$). It is traditional to call $R$ the ratio of hadronic to electromagnetic annihilation cross-sections [232],

$$
R(q^2_{\text{cm}}) = \frac{\sigma_{e^+e^-\rightarrow \text{hadron}}}{\sigma_{e^+e^-\rightarrow \mu^+\mu^-}} = 12\pi\text{Im} \bar{\omega}_{\text{hadrons}}(q^2_{\text{cm}}).
$$

(7.1.3)

As the energy $q^2_{\text{cm}}$ of electron and positron collision increases and reaches the mass–energy thresholds of constituents of hadrons, i.e. “quarks”, narrow resonances occurs, see Fig. 7.1 for the ratio $R$ as a function of $\sqrt{s}$ measured at SLAC [233]. These resonances correspond to production of particles such as $J/\psi$, $Y$ etc. This provides a fruitful investigation of hadron physics. For a

![Graph showing cross-sections and ratio R in e^+e^- collisions](image)

**Figure 7.1:** The total cross section of $e^+e^- \rightarrow \text{hadron}$ (7.1.2) and the ratio $R = \sigma_{e^+e^-\rightarrow \text{hadron}}/\sigma_{e^+e^-\rightarrow \mu^+\mu^-}$ (7.1.3), where $s \equiv q^2_{\text{cm}}$. Reproduced from Ref. [C. Amsler et al. (Particle Data Group), Physics Letters B667, 1 (2008)].

review of this topic see [101].
As the center of mass energy $q^2_{\text{cm}}$ reaches the electroweak scale (several hundred GeVs), electron and positron annihilation process probes a rich domain of investigating electroweak physics, see for instance Refs. [234, 235]. Recent experiments on $e^+e^-$ collisions at LEP, SLAC and the Tevatron allowed precision tests of the electroweak Standard Model. In [236, 237] the results of these precision tests together with implications on parameters, in particular Higgs boson mass, as well as constraints for possible new physics effects are discussed.

Electron and positron collisions are used to produce many particles in the laboratory, which live too short to occur naturally. Several electron–positron colliders have been built and proposed for this purpose all over the world (CERN, SLAC, INP, DESY, KEK and IHEP), since the first electron–positron collider the “Anello d’Accumulazione” (AdA) was built by the theoretical proposal of Bruno Touschek in Frascati (Rome) in 1960 [55]. Following the success of AdA (luminosity $\sim 10^{25}/(\text{cm}^2\text{ sec})$, beam energy $\sim 0.25\text{ GeV}$), it was decided to build in the Frascati National Laboratory a storage ring of the same kind, Adone and then Daphne (luminosity $\sim 10^{33}/(\text{cm}^2\text{ sec})$, beam energy $\sim 0.51\text{ GeV}$), with the aim of exploring the new energy range in sub-nuclear physics opened by the possibility of observing particle-antiparticle interactions with center of mass at rest. The biggest of all is CERN’s Large Electron Positron (LEP) collider [148], which began operation in the summer of 1989 and have reached a maximal collision energy of 206.5 GeV. The detectors around the LEP ring have been able to perform precise experiments, testing and extending our knowledge of particles and their strong, electromagnetic and weak interactions, as described by the Standard Model for elementary particle physics.

All these clearly show how the study of $e^+e^-$ reaction introduced by Dirac have grown to be one of the most prolific field in particle physics and have received remarkable verification in energies up to TeV in a succession of machines increasing in energy.

### 7.2 The Breit–Wheeler process in laser physics

While the Dirac process (3.0.1) has been by far one of the most prolific in physics, the Breit–Wheeler process (3.0.2) has been one of the most elusive for direct observations. In Earth-bound experiments the major effort today is directed to evidence this phenomenon in very strong and coherent electromagnetic field in lasers. In this process collision of many photons may lead in the future to pair creation. This topic is discussed in the following Sections. Alternative evidence for the Breit–Wheeler process can come from optically thick electron–positron plasma which may be created in the future either in Earth-bound experiments, or currently observed in astrophysics, see Section 10. One additional way to probe the existence of the Breit–Wheeler process
is by establishing in astrophysics an upper limits to observable high-energy photons, as a function of distance, propagating in the Universe as pioneered by Nikishov [56], see Section 7.4.

We first briefly discuss the phenomenon of electron–positron pair production at the focus of an X-ray free electron laser, as given in the review articles [60, 238, 239]. In the early 1970’s, the question was raised whether intense laser beams could be used to produce a very strong electric field by focusing the laser beam onto a small spot of size of the laser wavelength \( \lambda \), so as to possibly study electron–positron pair production in vacuum [36, 240]. However, it was found that the power density of all available and conceivable optical lasers [241] is too small to have a sizable pair production rate for observations [30–32, 36, 37, 210, 240, 242–246], since the wavelength of optical lasers and the size of focusing spot are too large to have a strong enough electric field.

Definite projects for the construction of X-ray free electron lasers (XFEL) have been set up at both DESY and SLAC. Both are based on the principle pioneered by John Madey [57] of self-amplified spontaneous emission in an undulator, which results when charges interact with the synchrotron radiation they emit [58]. At DESY the project is called XFEL and is part of the design of the electron–positron collider TESLA [247–250] but is now being build as a stand-alone facility. At SLAC the project so-called Linac Coherent Light Source (LCLS) has been proposed [251–253]. It has been pointed out by several authors [254–257] that having at hand an X-ray free electron laser, the experimental study and application of strong field physics turn out to be possible. One will use not only the strong energy and transverse coherence of the X-ray laser beam, but also focus it onto a small spot hopefully with the size of the X-ray laser wavelength \( \lambda \approx O(0.1) \) nm [59], and obtain a very large electric field \( E \sim 1/\lambda \), much larger than those obtainable with any optical laser of the same power.

Using the X-ray laser, we can hopefully achieve a very strong electric field near to its critical value for observable electron–positron pair production in vacuum. Electron-positron pair production at the focus of an X-ray laser has been discussed in Ref. [254], and an estimate of the corresponding production rate has been presented in Ref. [255]. In Ref. [60, 238, 239], the critical laser parameters, such as the laser power and focus spot size, are determined in order for achieving an observable effect of pair production in vacuum.

The electric field produced by a single laser beam is the light-like static, spatially uniform electromagnetic field, and field invariants \( S \) and \( P \) (3.5.23) vanish [242]

\[
S = 0, \quad P = 0, \quad (7.2.1)
\]

leading to \( \epsilon = \beta = 0 \) and no pair production [25–27], this can be seen from Eqs. (3.5.23), (3.5.27) and (5.7.39). It is then required that two or more coherent laser beams form a standing wave at their intersection spot, where a strong
electric field can hypothetically be produced without magnetic field.

We assume that each X-ray laser pulse is split into two equal parts and recombined to form a standing wave with a frequency $\omega$, whose electromagnetic fields are then given by

$$\mathbf{E}(t) = [0, 0, E_{\text{peak}} \cos(\omega t)], \quad \mathbf{B}(t) = (0, 0, 0),$$

(7.2.2)

where the peak electric field is \[60, 238, 239\],

$$E_{\text{peak}} = \sqrt{\frac{P_{\text{laser}}}{\pi \sigma_{\text{laser}}^2}} \simeq 1.1 \cdot 10^{17} \left( \frac{P_{\text{laser}}}{1 \text{TW}} \right)^{1/2} \left( \frac{0.1 \text{nm}}{\sigma_{\text{laser}}} \right) \frac{V}{\text{m}},$$

(7.2.3)

as expressed in terms of the laser power $P_{\text{laser}}$ (1 TW=10^{12}W), with the focus spot radius $\sigma_{\text{laser}}$. Eq. (7.2.3) shows that the peak electric energy density $E_{\text{peak}}^2/2$ is created in a spot of area $\pi \sigma_{\text{laser}}^2$ by an X-ray laser of power $P_{\text{laser}}$. The laser beam intensity on the focused spot is then given by

$$I_{\text{laser}} = \frac{P_{\text{laser}}}{\pi \sigma_{\text{laser}}^2} \simeq \frac{c}{4\pi} E_{\text{peak}}^2.$$

For a laser pulse with wavelength $\lambda$ about 1$\mu$m and the theoretical diffraction limit $\sigma_{\text{laser}} \simeq \lambda$ being reached, the critical intensity laser beam can be defined as

$$I_{\text{laser}}^c = \frac{c}{4\pi} E_{\text{c}}^2 \simeq 4.6 \cdot 10^{29} \text{W/cm}^2,$$

(7.2.4)

which corresponds to the peak electric field approximately equal to the critical value $E_{\text{c}}$ in (2.0.1).

### 7.2.1 Phenomenology of pair production in alternating fields

To compute pair production rate in an alternating electric field (7.2.2) of laser wave in a semi-classical manner, one assumes the conditions that the peak electric field $E_{\text{peak}}$ is much smaller than the critical field $E_{\text{c}}$ (2.0.1) and the energy $\hbar \omega$ of the laser photons is much smaller than the rest energy of electron $m_e c^2$,

$$E_{\text{peak}} \ll E_{\text{c}}, \quad \hbar \omega \ll m_e c^2.$$  

(7.2.5)

These conditions are well satisfied at realistic optical as well as X-ray lasers [60,238,239].

The phenomenon of electron–positron pair production in alternating electric fields was studied in Refs. [30–32,36,37,209,210,243–246,258]. By using generalized JWKB method [36] and imaginary time method [30–32,244,246] the rate of pair production was computed. In Ref. [36], the rate of pair pro-
phenomenology of electron–positron pair creation and annihilation was estimated to be (see Section 5.8),

$$\tilde{\mathcal{P}} = \frac{c}{4\pi^3 \lambda_C^4} \left( \frac{E_{\text{peak}}}{E_c} \right)^2 \frac{\pi}{\tilde{g}(\eta)} + \frac{1}{2\pi \eta \tilde{g}'(\eta)} \exp \left[ -\pi \frac{E_{\text{peak}}}{E_c} \tilde{g}(\eta) \right], \quad (7.2.6)$$

where the complex function $\tilde{g}(\eta)$ is given in Refs. [30–32, 36] (see Eq. (5.8.24)),

$$\tilde{g}(\eta) = \frac{4}{\pi} \int_0^1 du \left[ \frac{1 - u^2}{1 + \eta^{-2} u^2} \right]^{1/2} = \left\{ \begin{array}{ll}
1 - \frac{1}{8\eta^3} + O(\eta^{-4}), & \eta \gg 1 \\
\frac{4\eta}{\pi} \ln \left( \frac{4}{\eta \gamma} \right) + O(\eta^3), & \eta \ll 1
\end{array} \right. \quad (7.2.7)$$

and the parameter $\eta$ is defined as the work done by the electric force $eE_{\text{peak}}$ in the Compton wavelength $\lambda_C$ in unit of laser photon energy $\hbar \omega$,

$$\eta = \frac{e E_{\text{peak}} \lambda_C}{h \omega} = \frac{m_e c^2 E_{\text{peak}}}{h \omega E_c}, \quad (7.2.8)$$

which is the same as $\eta$ in (5.8.23) in Section 5.8 and agrees with its time-average (5.9.4) and (5.9.5), over one period of laser wave. The exponential factor in Eq. (7.2.6) has been confirmed by later works [30–32, 244, 246], which determine more accurately the pre-exponential factor by taking into account also interference effects. The parameter $\eta$ is related to the adiabaticity parameter $\gamma = 1/\eta$.

In the strong field and low frequency limit ($\eta \gg 1$), formula (7.2.6) agrees to the Schwinger non-perturbative result (5.7.25) for a static and spatially uniform field, apart from an “inessential” (see Ref. [36]) pre-exponential factor of $\pi$. This is similar to the adiabatic approximation of a slowly varying electric field that we discuss in Section 7.5.5. On the other hand, for $\eta \ll 1$, i.e. in high frequency and weak field limit, Eq. (7.2.6) yields [36],

$$\tilde{\mathcal{P}} \approx \frac{c}{4\pi^3 \lambda_C^4} \left( \frac{\hbar \omega}{m_e c^2} \right)^2 \frac{\pi \eta}{2 \ln(4/\eta)} \left( \frac{e \eta}{4} \right)^{2m_e c^2} [1 + O(\eta^2)], \quad (7.2.9)$$

corresponding to the $n$th order perturbative result, where $n$ is the minimum number of quanta of laser field required to create an $e^+ e^-$ pair:

$$n \gtrsim \frac{2m_e c^2}{h \omega} \gg 1. \quad (7.2.10)$$

The pair production rate (7.2.6) interpolates analytically between the adiabatic, non-perturbative tunneling mechanism (5.7.25) ($\eta \gg 1, \gamma \ll 1$) and the anti-adiabatic, perturbative multi-photon production mechanism (7.2.9) ($\eta \ll 1, \gamma \gg 1$).

In Refs. [244, 246], it was found that the pair production rate, under the condition (7.2.5), can be expressed as a sum of probabilities $w_n$ of many photon
7.2 The Breit–Wheeler process in laser physics

processes,

\[ \tilde{P}_p = \sum_{n>n_0} w_n, \quad \text{with} \quad n_0 = \frac{m_e c^2}{\hbar \omega} \Delta_m, \quad (7.2.11) \]

where \( \Delta_m \) indicates an effective energy gap \( m_e c^2 \Delta_m \), due to the transverse oscillations of the electron propagating in a laser wave (see Section 5.10 and Eq. (5.10.11)). In the limiting cases of small and large \( \eta \), the result is given by [246],

\[ \tilde{P}_p \simeq \frac{c}{4 \pi^3 \lambda_C^4} \left[ \sqrt{2} \left( \frac{E_{\text{peak}}}{E_c} \right)^{5/2} \exp \left[ -\pi \left( \frac{E_c}{E_{\text{peak}}} \right) (1 - \frac{1}{8 \eta^2} + O(\eta^{-4})) \right], \quad \eta \gg 1 \right. \]

\[ \left. \frac{\sqrt{2}}{2} \left( \frac{\hbar \omega}{m_e c^2} \right)^{5/2} \sum_n \left( \frac{e \eta}{4} \right)^{2n} e^{-\phi \text{Erfi} \left( \phi^{1/2} \right)}, \quad \eta \ll 1, \right. \]

(7.2.12)

where \( n > \left( 2m_e c^2 / \hbar \omega \right) \), \( \phi = 2(n - 2m_e c^2 / \hbar \omega) \) and \( \text{Erfi}(x) \) is the imaginary error function [259]. The range of validity of results (7.2.6), (7.2.9), (7.2.11) and (7.2.12) is indicated by the conditions (7.2.5).

### 7.2.2 Pair production in X-ray free electron lasers

According to Eq. (7.2.3) for the electric field \( E \) of an X-ray laser, in order to obtain an observable effect of pair production we need to have a large power \( P \), a small laser focusing spot radius \( \sigma_{\text{laser}} \) and a long duration time \( \Delta t \) of the coherent laser pulse. The power of an X-ray free electron laser is limited by the current laser technology. The focusing spot radii \( \sigma_{\text{laser}} \) are limited by the diffraction to the order of the X-ray laser beam wavelength. In Ref. [60, 238, 239], it was estimated that to produce at least one pair of electron and positron, we need the minimum power of the X-ray laser to be \( \sim 2.5 \) to \( 4.5 \text{TW} \) corresponding to an electric field of \( \sim 1.7 \) to \( 2.3 \cdot 10^{15} \text{V/cm} \) to \( 0.1E_c \), provided the laser wavelength is \( \lambda \sim 0.1 \text{nm} \) and the theoretical diffraction limit \( \sigma_{\text{laser}} \sim \lambda \) is actually reached and the laser coherent duration time \( \Delta t \sim 10^{-13} \) to \( 10^{-16} \) second. Based on these estimations, Ringwald concluded [60, 238, 239] that with present available techniques, the power density and electric fields of X-ray laser are far too small to produce a sizable Sauter-Euler-Heisenberg-Schwinger effect. If the techniques for X-ray free electron laser are considerably improved, so that the XFEL power can reach the terawatt regime and the focusing spot radii can reach the theoretical diffraction limit, we will still have the possibility of investigating the Sauter-Euler-Heisenberg-Schwinger phenomenon by a future XFEL.

### 7.2.3 Pair production by a circularly polarized laser beam

Instead of a time varying electric field (7.2.2) that is created by an intersection of more than two coherent laser beams, it was suggested [260, 261] to use
a focused circularly polarized laser beam having nonvanishing field invariants $S, P$ (3.5.23) and strong electromagnetic fields $E, B$ for pair production. It is well known that the electromagnetic field of a focused light beam is not transverse, however, one can always represent the field of a focused beam as a superposition of fields with transverse either electric ($e$-polarized) field or magnetic ($h$-polarized) field only, see e.g., [262].

In Ref. [261], the $e$-polarized electric and magnetic fields $(E^e, B^e)$ propagating in the $\hat{z}$-direction is described by the following exact solution of Maxwell equations [263],

\[
E^e = i E_{\text{peak}} e^{-i\psi} \left[ F_1(\mathbf{e}_x \pm \mathbf{e}_y) - F_2 e^{\pm 2i\phi} (\mathbf{e}_x \mp i \mathbf{e}_y) \right]; \quad (7.2.13)
\]

\[
B^e = \pm E_{\text{peak}} e^{-i\psi} \left\{ \left( 1 - i\delta^2 \frac{\partial}{\partial \chi} \right) \right. \\
\left[ F_1(\mathbf{e}_x \pm \mathbf{e}_y) + F_2 e^{\pm 2i\phi} (\mathbf{e}_x \mp i \mathbf{e}_y) \right] + 2i \Delta e^{\pm 2i\phi} \frac{\partial F_1}{\partial \xi} \mathbf{e}_z \left\}, \quad (7.2.14)
\]

where $\psi = \omega(t - z/c)$, $e^{\pm 2i\phi} = (x + iy)/\rho$, $\chi = z\Delta/R$, $\xi = \rho/R$ and $\rho = \sqrt{x^2 + y^2}$. The focusing parameter $\Delta = \lambda/(2\pi R)$ is expressed in terms of laser’s wavelength $\lambda$ and the focal spot radius $R$. The functions $F_{1,2}(\xi, \chi, \Delta)$ obey differential equations [263], go to zero sufficiently fast when $\xi, |\chi| \to \infty$ and conditions $F_1(0,0,\Delta) = 1; F_2(0,0,\Delta) = 0$ are satisfied for $\Delta \to 0$ [261]. The $h$-polarized electric and magnetic fields $E^h = \pm i B^e$ and $B^h = \mp i E^e$ [263].

Corresponding to electromagnetic fields (7.2.13), (7.2.14), field invariants $S^e, P^e$ are given by Eq. (3.5.23) and $e, \beta$ by Eq. (3.5.25) in Section 3.5. The total number of electron and positron pairs is given by Eq. (5.7.39) for $n = 1$ (see also Eq. (5.7.41),

\[
N_{e^+ e^-} \simeq \frac{a}{\pi} \int_V dV \int_0^\tau dt e^\beta \coth \frac{\pi \beta}{\varepsilon} \exp \left( - \frac{\pi E_e}{\varepsilon} \right), \quad (7.2.15)
\]

where the integral is taken over the volume $V$ and duration $\tau$ of the laser pulse. The qualitative estimations and numerical calculations of total number $N_{e^+ e^-}$ of electron and positron pairs in terms of laser intensity $I_{\text{laser}}$ and focusing parameter $\Delta$ are presented in Ref. [261]. Two examples for $e$-polarized electromagnetic fields are as follows:

1. for $E_{\text{peak}} = E_c$, $\lambda = 1 \mu m$, $\tau = 10 fs$ and $\Delta = 0.1$ (the theoretical diffraction limit), the laser beam intensity $I_{\text{laser}} \simeq 1.5 \cdot 10^{29} W/cm^2 \simeq 0.31 I_{\text{laser}}^c$ the critical intensity (7.2.4). The total number of pairs created $N_{e^+ e^-} \simeq 5 \cdot 10^{20}$ according to the Schwinger formula Eq. (7.2.15) for pair production rate;

2. with the same values of laser parameters $\Delta$, $\lambda$ and $\tau$, while the laser pulse intensity $I_{\text{laser}} \simeq 5 \cdot 10^{27} W/cm^2 \sim 10^{-2} I_{\text{laser}}^c$ corresponding to
$E_{\text{peak}} \sim 0.18E_c$, Eq. (7.2.15) gives $N_{e^+e^-} \sim 20$.

Because the volume $V$ and time occupied by one pair, the average number of pairs $N_{e^+e^-} \sim 20$ is large and possibly observable even if the peak value of electric field is only 18% of the critical value. In addition, pair production is much more effective by the $e$-polarized electric and magnetic fields $E^e, B^e$ than by the $h$-polarized fields $E^h, B^h$. The detailed analysis of the dependence of the number of pairs $N_{e^+e^-}$ on the laser intensity $I_{\text{laser}}$ and focusing parameter $\Delta$ is given in [264], and results are presented in Fig. (7.2). In particular, it is shown that for the case of two counter-propagating focused laser pulses with circular polarizations, pair production becomes experimentally observable when the laser intensity $I_{\text{laser}} \sim 10^{26}$ W/cm$^2$ for each beam, which is about 1 $\sim$ 2 orders of magnitude lower than for a single focused laser pulse, and more than 3 orders of magnitude lower than the critical intensity (7.2.4). In these calculations the “imaginary time” method is useful [246, 264, 265], which gives a clear description of tunneling of a quantum particle through a potential barrier. Recently the process of electron–positron pair creation in the superposition of a nuclear Coulomb and a strong laser field was studied in [266].

It was pointed [261, 264] that the exploited method becomes inconsistent and one should take into account back reaction of the pair production effect on the process of laser pulse focusing at such high laser intensity and $E_{\text{peak}} \sim E_c$. It has already been argued in Refs. [209, 210, 244, 246] that for the superstrong field regime $E \gtrsim 0.1E_c$, such back reaction of the produced electron–positron pairs on the external field and the mutual interactions between these particles have to be considered. These back reaction effects on pair production by laser beams leading to the formation of plasma oscillation have been studied in Refs. [267–269]. Our studies [72, 73] show that the plasma oscillation and electron–positron–photon collision are important for electric fields $E \gtrsim 0.1E_c$, see Section 9.7.

### 7.2.4 Availability of laser technology for pair production

There are several ways to increase the electromagnetic fields of a laser beam. One way is to increase the frequency of the laser radiation and then focus it onto a tiny region. X-ray lasers can be used [60, 238, 239, 257, 268]. Another way is, clearly, to increase the intensities of laser beams. The recent development of laser technology and the invention of the chirped pulse amplification (CPA) method have led to a stunning increase of the light intensity ($10^{22}$ W/cm$^2$) in a laser focal spot [270, 271]. To achieve intensities of the order $10^{24} - 25$ W/cm$^2$, a scheme was suggested in Ref. [272], where a quasi-soliton wave between two foils is pumped by the external laser field up to an ultrahigh magnitude. Using the method based on the simultaneous laser fre-
Figure 7.2: Logarithm of the number of pairs $N_{e^+e^-}$ produced by the field of two counter-propagating laser-pulses (circular polarization) is shown as functions of: (a) the beam intensity $I_{\text{laser}}$ for the focusing parameters $\Delta = 0.1, 0.075, 0.05$ and 0.01 (the curves 1, 2, 3 and 4 correspondingly); (b) the focusing parameter $\Delta$ for the beam intensity $I_{\text{laser}} = 4 \cdot 10^{26} \text{W/cm}^2$ and laser-pulse duration $\tau = 10^{-14}$ sec. This figure is reproduced from Fig. 6 in Ref. [264].
frequency upshifting and pulse compression, another scheme for reaching critical intensities has also been suggested in Ref. [273,274], where the interaction of the laser pulse with electron density modulations in a plasma produced by a counter-propagating breaking wake plasma wave, results in the frequency upshift and pulse focusing. In addition, it has been suggested [270] a path to reach the extremely high intensity level of $10^{26-28}\text{W/cm}^2$ already in the coming decade. Such field intensities are very close to the value of critical intensity $I_{\text{laser}}^c$ (7.2.4). For a recent review, see Ref. [275]. This technological situation has attracted the attention of the theorists who involved in physics in strong electromagnetic fields.

Currently available technologies allow for intensities of the order of $10^{20}\text{W/cm}^2$ leading to abundant positron production essentially through the Bethe-Heitler process (3.3.1) with number densities of the order of $10^{16}\text{cm}^{-3}$ [276].

7.3 Phenomenology of pair production in electron beam-laser collisions

7.3.1 Experiment of electron beam-laser collisions

After the availability of high dense and powerful laser beams, the Breit–Wheeler process (3.0.2) has been reconsidered in Refs. [40–45,61,62] for high-energy multiple photon collisions. The phenomenon of $e^+e^-$ pair production in multi-photon light-by-light scattering has been reported in [38,39,211,254] on the experiment SLAC-E-144 [63, 277].

As described in Ref. [39], such a large center of mass energy ($2m_e c^2 = 1.02\text{MeV}$) can be possibly achieved in the collision of a laser beam against another high-energy photon beam. With a laser beam of energy 2.35 eV, a high-energy photon beam of energy 111 GeV is required for the Breit–Wheeler reaction (3.0.2) to be feasible. Such a high-energy photon beam can be created for instance by backscattering the laser beam off a high-energy electron beam, i.e., by inverse Compton scattering. With a laser beam of energy 2.35 eV (wavelength 527 nm) backscattering off a high-energy electron beam of energy 46.6 GeV, as available at SLAC [63], the maximum energy acquired by Compton-backscattered photon beam is only 29.2 GeV. This is still not enough for the Breit–Wheeler reaction (3.0.2) to occur, since such photon energy is four times smaller than the needed energetic threshold.

Nevertheless in strong electromagnetic fields and a long coherent time-duration $\Delta t = \frac{2\pi}{\omega}$ of the laser pulse, the number $n$ of laser photons interacting with scattered electron becomes large, when the intensity parameter of laser fields $\eta$ (5.9.4) approaches or even exceeds unity. Once this number $n$ is larger than the critical number $n_0$ defined after Eq. (5.10.11) in Section 7.3, pair production by the nonlinear Breit–Wheeler reaction (5.9.2) for
high-energy multiple photon collisions becomes feasible.

The probability of pair production by the processes (5.9.1) and (5.9.2) is given by Eqs. (5.10.9) and (5.10.11) for any values of \( \eta \) in Section 5.10. In high frequency and weak field limit \( \eta \ll 1 \), the probability \( \mathcal{P}_{e\gamma} \) (5.10.9) and \( \mathcal{P}_{\gamma\gamma} \) (5.10.11) for fairly small \( n \) are proportional to \( \eta^{2n} \), i.e.

\[
\mathcal{P}_{e\gamma} \propto \eta^{2n}, \quad \mathcal{P}_{\gamma\gamma} \propto \eta^{2n} \tag{7.3.1}
\]

(see Eqs. (7.2.9), (7.2.12)). This corresponds to the anti-adiabatic, perturbative multi-photon production mechanism (7.2.9), (7.2.12) for \( (\eta \ll 1) \). In low frequency and strong field limit \( \eta \gg 1 \), it essentially refers to process in a constant and uniform field where \( E \) and \( B \) are orthogonal and equal in magnitude. This corresponds to the adiabatic limit of a slowly varying electromagnetic field discussed in Section 7.2.1.

For \( n \geq 5 \) laser photons of energy 2.35eV colliding with a photon of energy 29 GeV, the process of nonlinear Breit–Wheeler pair production becomes energetically accessible. In Refs. [38,39,211], it is reported that nonlinear Compton scattering (5.9.1) and nonlinear Breit–Wheeler electron–positron pair production (5.9.2) have been observed in the collision of 46.6 GeV and 49.1 GeV electrons of the Final Focus Test Beam at SLAC with terawatt pulse of 1053 nm (1.18 eV) and 527 nm (2.35 eV) wavelengths from a Nd:glass laser. The rate of pair production, i.e., \( R_{e^+} \) of positrons/(laser shot) is measured in terms of the parameter \( \eta \) (\( \eta \ll 1 \)), as shown in Fig. 7.3, where line represents a power law fit to the data which gives [39],

\[
R_{e^+} \propto \eta^{2n}, \quad \text{with} \quad n = 5.1 \pm 0.2(\text{stat})^{+0.5}_{-0.8}(\text{syst}). \tag{7.3.2}
\]

These experimental results are found to be in agreement with theoretical predictions (7.3.1), i.e., (5.10.9), (5.10.11) for small \( \eta \); as well as with (7.2.12) and (7.2.9) for \( \omega \rightarrow \gamma \omega \) in the frame of reference where the electron beam is at rest. This shows that the pair production of Breit–Wheeler type by the anti-adiabatic, perturbative multi-photon production mechanism, described by Eqs. (5.10.9), (5.10.11) or (7.2.9), (7.2.12) for small \( \eta \ll 1 \), has been experimentally confirmed. However, one has not yet experimentally observed the pair production by the adiabatic, non-perturbative tunneling mechanism, described by Eqs. (5.10.9), (5.10.11) or (7.2.9), (7.2.12) for large \( \eta \gg 1 \), i.e. for static and constant electromagnetic fields. Nevertheless, pair production probabilities Eqs. (7.2.6), (7.2.12) and Eqs. (5.10.9), (5.10.11) interpolates between both \( \eta \ll 1 \) and \( \eta \gg 1 \) regimes. Based on such analyticity of these probability functions in terms of the laser intensity parameter \( \eta \), we expect the pair production to be observed in \( \eta \gg 1 \) regime.
7.3 Phenomenology of pair production in electron beam-laser collisions

Figure 7.3: Dependence of the positron rate per laser shot on the laser field-strength parameter $\eta$. The line shows a power law fit to the data. The shaded distribution is the 95% confidence limit on the residual background from showers of lost beam particles after subtracting the laser-off positron rate. This figure is reproduced from Fig. 4 in Ref. [39]
7.3.2 Pair production viewed in the rest frame of electron beam

In the reference frame where the electron beam is at rest, one can discuss [254] pair production in the processes (5.9.1) and (5.9.2) by using pair production rate Eqs. (7.2.6), (7.2.9), (7.2.12) in Section 7.2. In the experiment of colliding 46.6GeV electron beam with 2.35eV (527nm) laser wave, the field strength in the laboratory is $E_{\text{lab}} \simeq 6 \times 10^{10} \text{V/cm}$ and intensity $I \simeq 10^{19} \text{W/cm}^2$ for $\eta = 1$ [39]. The Lorentz gamma factor of the electron beam $\gamma = \frac{E_e}{m_e c^2} \simeq 9.32 \times 10^4$ for $E_e \simeq 46.6 \text{GeV}$. In the rest frame of the electron beam, the electric field is given by

$$E_{\text{rest}} = \gamma (E_{\text{lab}} + \mathbf{v} \times B_{\text{lab}}) = \gamma E_{\text{lab}} (1 + |\mathbf{v}|) \simeq 2 \gamma E_{\text{lab}}$$ \hspace{1cm} (7.3.3)

where laser’s electromagnetic fields $\mathbf{E} \cdot \mathbf{B} = 0$, $|\mathbf{E}| = |\mathbf{B}|$, $\mathbf{B} \times \hat{k} = \mathbf{E}$, and laser’s wave vector $\hat{k} = -\mathbf{v}$, thus one has

$$E_{\text{rest}} \simeq 2 \gamma E_{\text{lab}} \simeq 2 \cdot 10^5 E_{\text{lab}} \sim 0.86 E_c.$$ \hspace{1cm} (7.3.4)

The field of 2.35eV laser wave is well defined coherent wave field with wavelength $\lambda_{\text{lab}} = 5.27 \times 10^{-5} \text{cm}$ and frequency $\omega_{\text{lab}} = 3.57 \times 10^{15} \text{sec}$ (the period $\Delta t_{\text{lab}} = 2\pi / \omega = 1.76 \times 10^{-15} \text{sec}$). In the rest frame of electron beam, $\lambda_{\text{rest}} = \gamma \lambda_{\text{lab}} = 4.91 \text{cm}$ and $\Delta t_{\text{rest}} = \Delta t_{\text{lab}} / \gamma = 1.9 \times 10^{-20} \text{sec}$. Comparing these wavelength and frequency of laser wave field with the spatial length $\hbar / m_e c = 3.86 \times 10^{-11} \text{cm}$ and timescale $\hbar / m_e c^2 = 1.29 \times 10^{-21} \text{sec}$ of spontaneous pair production in vacuum, we are allowed to apply the homogeneous and adiabatic approximation discussed in Section 7.5.5, and use the rate of pair production (7.2.6), (7.2.9), (7.2.12) in Section 7.2.

7.4 The Breit–Wheeler cutoff in high-energy $\gamma$-rays

Having determined the theoretical basis as well as attempts to verify experimentally the Breit–Wheeler formula we turn to a most important application of the Breit–Wheeler process in the framework of cosmology. As pointed out by Nikishov [56] the existence of background photons in cosmology puts a stringent cutoff on the maximum trajectory of the high energy photons.

The Breit–Wheeler process for the photon-photon pair production is one of most relevant elementary processes in high-energy astrophysics. In addition to the importance of this process in dense radiation fields of compact objects [278], the essential role of this process in the context of intergalactic absorption of high-energy $\gamma$-rays was first pointed out by Nikishov [56, 279]. The spectra of TeV radiation observed from distant ($d > 100 \text{ Mpc}$) extra-
galactic objects suffer essential deformation during the passage through the intergalactic medium, caused by energy-dependent absorption of primary \( \gamma \)-rays interacting with the diffuse extragalactic background radiation, for the optical depth \( \tau_{\gamma\gamma} \) most likely significantly exceeding one [279–282]. A relevant broad-band information about the cosmic background radiation (CBR) is important for the interpretation of the observed high-energy \( \gamma \) spectra [283–286]. For details, readers are referred to Refs. [287, 288]. In this section, we are particularly interested in such absorption effect of high-energy \( \gamma \)-ray, originated from cosmological sources, interacting with the Cosmic Microwave Background (CMB) photons. Fazio and Stecker [289, 290] were the first who calculated the cutoff energy versus redshift for cosmological \( \gamma \)-rays. This calculation was applied to further study of the optical depth of the Universe to high-energy \( \gamma \)-rays [291–293]. With the Fermi telescope, such study turns out to be important to understand the spectrum of high-energy \( \gamma \)-ray originated from sources at cosmological distance, we therefore offer the details of theoretical analysis as follow [294].

We study the Breit–Wheeler process (3.0.2) in the case that high-energy photons \( \omega_1 \), originated from sources at cosmological distance \( z \), on their way to us, collide with CMB photons \( \omega_2 \) in the rest frame of CMB photons, leading to electron–positron pair production. We calculate the opacity and mean free path of these high-energy photons, find the energy range of absorption as a function of the cosmological redshift \( z \).

In general, a high-energy photon with a given energy \( \omega_1 \), collides with background photons having all possible energies \( \omega_2 \). We assume that \( i \)-type background photons have the spectrum distribution \( f_i(\omega_2) \), the opacity is then given by

\[
\tau^i_{\gamma\gamma}(\omega_1, z) = \int_0^z \frac{dz'}{H(z')} \int_{m_\gamma^2 c^4/\omega_1}^{\infty} \frac{\omega_2^2 d\omega_2}{\pi^2} f_i(\omega_2) \sigma_{\gamma\gamma} \left( \frac{\omega_1 \omega_2}{m_\gamma^2 c^4} \right),
\]

(7.4.1)

where \( m_\gamma^2 c^4/\omega_1 \) is the energy threshold (3.2.8) above which the Breit–Wheeler process (3.0.2) can occur and the cross-section \( \sigma_{\gamma\gamma} \) is given by Eqs. (3.2.9); \( H(z) \) is the Hubble function, obeyed the Friedmann equation

\[
H(z) = H_0 \left[ \Omega_M (z + 1)^3 + \Omega_\Lambda \right]^{1/2}.
\]

(7.4.2)

We will assume \( \Omega_M \simeq 0.3 \) and \( \Omega_M \simeq 0.7 \) and \( H_0 = 75 \text{Km/s/Mpc} \). The total opacity is then given by

\[
\tau^{\text{total}}_{\gamma\gamma}(\omega_1, z) = \sum_i \tau^i_{\gamma\gamma}(\omega_1, z),
\]

(7.4.3)

which the sum is over all types of photon backgrounds in the Universe.

In the case of CMB photons the ir distribution is black-body one \( f_{\text{CMB}}(\omega_2/T) = \)
$1/(e^{\omega_2/T} - 1)$ with the CMB temperature $T$, the opacity is given by

$$
\tau_{\gamma\gamma}(\omega_1, z) = \int_0^z \frac{dz'}{H(z')} \int_{m_e^2c^4/\omega_1}^{\infty} \frac{d\omega_2}{\pi^2} \frac{\omega_2^2}{e^{\omega_2/T} - 1} \sigma_{\gamma\gamma}(\omega_1\omega_2/m_e^2c^4),
$$

(7.4.4)

where the Boltzmann constant $k_B = 1$. To simply Eq. (7.4.4), we set $x = \omega_1\omega_2/m_e^2c^4$.

In terms of CMB temperature and high-energy photons energy at the present time,

$$
T = (z + 1)T_0; \quad \omega_{1,2} = (z + 1)\omega^0_{1,2},
$$

(7.4.5)

we obtain,

$$
\tau_{\gamma\gamma}(\omega^0_1, z) = \frac{1}{R_0} \int_0^z \frac{dz'}{H(z')} (z + 1)^3 \left( \frac{m_e^2c^4}{\omega^0_1} \right)^3 \int_1^\infty \frac{dx}{\pi^2} \frac{x^2}{\exp(x/\theta) - 1} \sigma_{\gamma\gamma}(x),
$$

(7.4.6)

where

$$
\theta = x_0(z + 1)^2; \quad x_0 = \frac{\omega^0_1T_0}{m_e^2c^4},
$$

(7.4.7)

and $x_0$ is the energy $\omega^0_1$ in unit of $m_e^2c^2(m_e^2c^2/T_0) = 1.11 \cdot 10^{15}$eV.

The $\tau_{\gamma\gamma}(\omega^0_1, z) = 1$ gives the relationship $\omega^0_1 = \omega^0_1(z)$ that separates the optically thick $\tau_{\gamma\gamma}(\omega^0_1, z) > 1$ and optically thin $\tau_{\gamma\gamma}(\omega^0_1, z) < 1$ regimes in the $\omega^0_1 - z$ plane.

The integral (7.4.6) is evaluated numerically and the result is presented in Fig. 7.4. It clearly shows the following properties:

1. for the redshift $z$ smaller than a critical value $z_c \simeq 0.1 (z < z_c)$, the CMB is transparent to photons with arbitrary energy, this indicates a minimal free path of high-energy photons;

2. for the redshift $z$ larger than the value ($z > z_c$), there are two branches of solutions for $\tau_{\gamma\gamma}(\omega^0_1, z) = 1$, respectively corresponding to the different energy dependence of the cross-section (3.2.9): the cross-section $\sigma_{\gamma\gamma}(x)$ increases with the center of mass energy $x = \omega_1\omega_2^2/(m_e^2c^2)^2$ from the energy threshold $x = 1$ to $x \simeq 1.97$, and decreases (3.2.10) from $x \simeq 1.97$ to $x \to \infty$. The energy of the CMB photon corresponding to the critical redshift $z \simeq 0.1, \omega^0_1$ is $\simeq 1.15 \cdot 10^{15}$eV which separates two branches of the solution. The position of this point in Fig. 7.4 is determined by the maximal cross-section at $x \simeq 1.97$. Due to existence of these two solutions for a given redshift $z$, photons having energies in the grey region of Fig. 7.4 cannot reach the observer, while photons from the white region of Fig. 7.4 are observable.

3. above the critical redshift $z_c$, low-energy photons can reach us since their energies are smaller than the energetic threshold for the Breit–Wheeler
Figure 7.4: This is a Log-Log plot for high energy photon energy $x_0$ in units of $1.11 \cdot 10^{15}$ eV versus redshift $z$. The grey region represents optically thick case, while the white one is for optically thin case. The boundary between the two is the two-branch solution of Eq. 7.4.6 for $\tau_{\gamma\gamma} = 1$. There is a critical redshift $z_c \approx 0.1$ for a photon with arbitrary energy, which can reach the observer. The value of the photon energy corresponding to this critical redshift is $\omega_1^0 \approx 1.11 \cdot 10^{15}$ eV.
process (3.0.2). In addition, high-energy photons are also observable due to the fact that the cross-section of Breit–Wheeler process (3.0.2) decreases with increasing energy of photons. For large redshifts \( z \sim 10^3 \), the Universe is opaque and we disregard this regime.

In Section 5.4 we considered another relevant process which is double pair production (5.4.1). This process contributes to the opacity at very high energies and its effect has been computed in [295]. We also computed the effect of this process on our diagram in Fig. 7.4. This process becomes relevant at very high redshift \( z \sim 10^3 \).

Due to the fact that there are other radiation backgrounds contributing into (7.4.3), the background of CMB photons gives the lower limit for opacity for high-energy photons with respect to the Breit–Wheeler process (3.0.2). Finally, we point out that the small-energy solution for large redshift in Fig. (7.4) agrees with the one found by Fazio and Stecker [289, 290].

### 7.5 Theory of pair production in Coulomb potential

By far the major attention to build a critical electric field has occurred in the physics of heavy nuclei and in heavy-ion collisions. We recall in the following some of the basic ideas, calculations, as well as experimental attempts to obtain the pair creation process in nuclear physics.

#### 7.5.1 The \( Z = 137 \) catastrophe

Soon after the Dirac equation for a relativistic electron was discovered [80, 296], Gordon [297, 298] (for all \( Z < 137 \)) and Darwin [299] (for \( Z = 1 \)) found its solution in the point-like Coulomb potential \( V(r) = -Z \alpha / r, \ 0 < r < \infty \). Solving the differential equations for the Dirac wave function, they obtained the well-known Sommerfeld’s formula [300] for the energy spectrum,

\[
E(n,j) = m_e c^2 \left[ 1 + \left( \frac{Z \alpha}{n - |K| + (K^2 - Z^2 \alpha^2)^{1/2}} \right)^2 \right]^{-1/2}.
\]

(7.5.1)

Here the principle quantum number \( n = 1, 2, 3, \ldots \) and

\[
K = \begin{cases} 
-(j + 1/2) = -(l + 1), & \text{if } j = l + \frac{1}{2}, \ l \geq 0 \\
(j + 1/2) = l, & \text{if } j = l - \frac{1}{2}, \ l \geq 1 
\end{cases}
\]

(7.5.2)

where \( l = 0, 1, 2, \ldots \) is the orbital angular momentum corresponding to the upper component of Dirac bi-spinor, \( j \) is the total angular momentum, and
the states with $K = \mp 1, \mp 2, \mp 3, \cdots, \mp (n - 1)$ are doubly degenerate\(^1\), while the state $K = -n$ is a singlet [297–299]. The integer values $n$ and $K$ label bound states whose energies are $\mathcal{E}(n,j) \in (0,mc^2)$. For the example, in the case of the lowest-energy states, one has

\begin{align*}
\mathcal{E}(1S_\frac{1}{2}) &= mc^2 \sqrt{1 - (Z\alpha)^2}, \\
\mathcal{E}(2S_\frac{1}{2}) &= \mathcal{E}(2P_\frac{1}{2}) = mc^2 \sqrt{\frac{1 + \sqrt{1 - (Z\alpha)^2}}{2}}, \\
\mathcal{E}(2P_\frac{3}{2}) &= mc^2 \sqrt{1 - \frac{1}{4}(Z\alpha)^2}.
\end{align*}

For all states of the discrete spectrum, the binding energy $mc^2 - \mathcal{E}(n,j)$ increases as the nuclear charge $Z$ increases, as shown in Fig. 7.5. When $Z = 137$, $\mathcal{E}(1S_{1/2}) = 0$, $\mathcal{E}(2S_{1/2}) = \mathcal{E}(2P_{1/2}) = (mc^2)/\sqrt{2}$ and $\mathcal{E}(2S_{3/2}) = mc^2\sqrt{3/2}$. Gordon noticed in his pioneer paper [297, 298] that no regular solutions with $n = 1, j = 1/2, l = 0$, and $K = -1$ (the $1S_{1/2}$ ground state) are found beyond $Z = 137$. This phenomenon is the so-called “$Z = 137$ catastrophe” and it is associated with the assumption that the nucleus is point-like in calculating the electronic energy spectrum.

In fact, it was shown since the pioneering work of Pomeranchuk [64] that in nature there cannot be a point-like charged object with effective coupling constant $Z\alpha > 1$ since the entire electron shell will collapse to the center $r = 0$.

### 7.5.2 Semi-Classical description

In order to have further understanding of this phenomenon, we study it in the semi-classical scenario. For simplicity we treat relativistic electron as a scalar particle fulfilling the Klein–Gordon equation, but still obeying Fermi-Dirac statistics. Setting the origin of spherical coordinates $(r, \theta, \phi)$ at the point-like charge, we introduce the vector potential $A_\mu = (A, A_0)$, where $A = 0$ and $A_0$ is the Coulomb potential. The motion of a relativistic “electron” with mass $m$ and charge $e$ is described by its radial momentum $p_r$, angular momenta $p_\phi$, and the Hamiltonian

\begin{equation}
H_\pm = \pm mc^2 \sqrt{1 + \left(\frac{p_r}{mc}\right)^2 + \left(\frac{p_\phi}{mc^2r}\right)^2 + V(r),}
\end{equation}

where the potential energy $V(r) = eA_0$, and $\pm$ corresponds for positive and negative solutions. The states corresponding to negative energy solutions are

\(^{1}\)This degeneracy is removed by radiative corrections [90, 133]. The shift of the level $2S_{1/2}$ up, compared to the level $2P_{1/2}$ (the famous Lamb shift) was discovered out of the study of fine structure of the hydrogen spectrum.
Figure 7.5: Atomic binding energies as function of nuclear charge $Z$. This figure is reproduced from Fig. 1 in Ref. [66].

fully occupied. The angular momentum $p_\phi$ is conserved, when the Hamiltonian is spherically symmetric. For a given angular momentum $p_\phi$, the Hamiltonian (7.5.6) describes electron’s radial motion in the following effective potential

$$E_{\pm} = \pm m_e c^2 \sqrt{1 + \left( \frac{p_\phi}{m_e c \alpha} \right)^2 + V(r)}. \quad (7.5.7)$$

The Coulomb potential energy $V(r)$ is given by

$$V(r) = -\frac{Ze^2}{r}. \quad (7.5.8)$$

In the classical scenario, given different values of angular momenta $p_\phi$, the stable circulating orbits (states) are determined by the minimum of the effective potential $E_+(r)$ (7.5.7). Using $dE_+(r)/dr = 0$, we obtain the stable orbit location at the radius $R_L$ in the unit of the Compton length $\lambda_C$,

$$R_L(p_\phi) = Z\alpha \lambda_C \sqrt{1 - \left( \frac{Z\alpha}{p_\phi/\hbar} \right)^2}, \quad (7.5.9)$$
where $\alpha = e^2/\hbar c$ and $p_\phi > Z\alpha$. Substituting Eq. (7.5.9) into Eq. (7.5.7), we find the energy of the electron at each stable orbit,

$$\mathcal{E}(p_\phi) \equiv \min(E_+) = m_e c^2 \sqrt{1 - \left(\frac{Z\alpha}{p_\phi/\hbar}\right)^2}. \quad (7.5.10)$$

The last stable orbits (minimal energy) are given by

$$p_\phi \rightarrow Z\alpha \hbar + 0^+, \quad R_L(p_\phi) \rightarrow 0^+, \quad \mathcal{E}(p_\phi) \rightarrow 0^+.$$ \quad (7.5.11)

For stable orbits with $p_\phi/\hbar \gg 1$, the radii $R_L/\lambda_C \gg 1$ and energies $\mathcal{E} \rightarrow m_e c^2 + 0^-$; electrons in these orbits are critically bound since their binding energy goes to zero. As the energy spectrum (7.5.1), see Eqs. (7.5.3,7.5.4,7.5.5), Eq. (7.5.10) shows, only positive or null energy solutions (states) exist in the presence of a point-like nucleus.

In the semi-classical scenario, the discrete values of angular momentum $p_\phi$ are selected by the Bohr-Sommerfeld quantization rule

$$\int p_\phi d\phi \simeq \hbar (l + \frac{1}{2}), \quad \Rightarrow \quad p_\phi(l) \simeq \hbar (l + \frac{1}{2}), \quad l = 0, 1, 2, 3, \ldots \quad (7.5.12)$$

describing the semi-classical states of radius and energy

$$R_L(l) \simeq (Z\alpha)^{-1} \lambda_C \sqrt{1 - \left(\frac{2Z\alpha}{2l + 1}\right)^2}, \quad (7.5.13)$$

$$\mathcal{E}(l) \simeq m_e c^2 \sqrt{1 - \left(\frac{2Z\alpha}{2l + 1}\right)^2}. \quad (7.5.14)$$

Other values of angular momentum $p_\phi$, radius $R_L$ and energy $\mathcal{E}$ given by Eqs. (7.5.9,7.5.10) in the classical scenario are not allowed. When these semi-classical states are not occupied as required by the Pauli Principle, the transition from one state to another with different discrete values ($l_1, l_2$ and $\Delta l = l_2 - l_1 = \pm 1$) is possible by emission or absorption of a spin-1 ($\hbar$) photon. Following the energy and angular-momentum conservations, photons emitted or absorbed in the transition have angular momentum $p_\phi(l_2) - p_\phi(l_1) = \hbar(l_2 - l_1) = \pm \hbar$ and energy $\mathcal{E}(l_2) - \mathcal{E}(l_1)$. As required by the Heisenberg uncertainty principle $\Delta \phi \Delta p_\phi \simeq 4\pi p_\phi(l) \gtrsim \hbar$, the absolute ground state for minimal energy and angular momentum is given by the $l = 0$ state, $p_\phi \sim \hbar/2$, $R_L \sim Z\alpha \lambda_C \sqrt{1 - (2Z\alpha)^2} > 0$ and $\mathcal{E} \sim m_e c^2 \sqrt{1 - (2Z\alpha)^2} > 0$ for $Z\alpha \leq 1/2$. Thus the stability of all semi-classical states $l > 0$ is guaranteed by the Pauli principle. In contrast for $Z\alpha > 1/2$, there is not an absolute ground state in the semi-classical scenario.

We see now how the lowest-energy states are selected by the quantization...
rule in the semi-classical scenario out of the last stable orbits (7.5.11) in the classical scenario. For the case of $Z\alpha \leq 1/2$, equating (7.5.11) to (7.5.12), we find the selected state $l = 0$ is only possible solution so that the ground state $l = 0$ in the semi-classical scenario corresponds to the last stable orbits (7.5.11) in the classical scenario. On the other hand for the case $Z\alpha > 1/2$, equating (7.5.11) to (7.5.12), we find the selected state $l = \tilde{l} \equiv (Z\alpha - 1)/2 > 0$ in the semi-classical scenario corresponds to the last stable orbits (7.5.11) in the classical scenario. This state $l = \tilde{l} > 0$ is not protected by the Heisenberg uncertainty principle from quantum mechanically decaying in $\hbar$-steps to the states with lower angular momentum and energy (correspondingly smaller radius $R_L$ (7.5.13)) via photon emissions. This clearly shows that the “$Z = 137$-catastrophe” corresponds to $R_L \rightarrow 0$, falling to the center of the Coulomb potential and all semi-classical states ($l$) are unstable.

### 7.5.3 The critical value of the nuclear charge $Z_{cr} = 173$

A very different situation is encountered when considering the fact that the nucleus is not point-like and has an extended charge distribution [64,65,301–307]. In that case the $Z = 137$ catastrophe disappears and the energy levels $E(n,j)$ of the bound states $1S$, $2P$ and $2S$, $\cdots$ smoothly continue to drop toward the negative energy continuum as $Z$ increases to values larger than 137, as shown in Fig. 7.5. The reason is that the finite size $R$ of the nucleus charge distribution provides a cutoff for the boundary condition at the origin $r \rightarrow 0$ and the energy levels $E(n,j)$ of the Dirac equation are shifted due to this cutoff. In order to determine the critical value $Z_{cr}$ when the negative energy continuum ($E < -m_e c^2$) is encountered (see Fig. 7.5), Zeldovich and Popov [65,304–307] solved the Dirac equation corresponding to a nucleus of finite extended charge distribution, i.e., the Coulomb potential is modified as

$$V(r) = \begin{cases} -\frac{Ze^2}{r}, & r > R, \\ -\frac{Ze^2}{R} f\left(\frac{r}{R}\right), & r < R, \end{cases} \tag{7.5.15}$$

where $R \sim 10^{-12}$ cm is the size of the nucleus. The form of the cutoff function $f(x)$ depends on the distribution of the electric charge over the volume of the nucleus ($x = r/R, 0 < x < 1$, with $f(1) = 1$). Thus, $f(x) = (3 - x^2)/2$ corresponds to a constant volume density of charge.

Solving the Dirac equation with the modified Coulomb potential (7.5.15) and calculating the corresponding perturbative shift $\Delta E_R$ of the lowest-energy level (7.5.3) Popov obtains [65,304]

$$\Delta E_R = m_e c^2 \left(\frac{2}{\gamma_z} \right)^2 \frac{2\gamma_z}{1 + 2\gamma_z} \left[ 1 - 2\gamma_z \int_0^1 f(x)x^{2\gamma_z}dx \right], \tag{7.5.16}$$
where \( \xi = Z\alpha, \gamma_z = \sqrt{1 - \xi^2} \) and \( \Lambda = \ln(h/m_e c R) \gg 1 \) is a logarithmic parameter in the problem under consideration. The asymptotic expressions for the \( 1S_{1/2} \) energy that were obtained are [65, 307]

\[
\mathcal{E}(1S_{1/2}) = m_e c^2 \begin{cases} 
\sqrt{1 - \xi^2} \coth(\Lambda \sqrt{1 - \xi^2}), & 0 < \xi < 1, \\
\Lambda^{-1}, & \xi = 1, \\
\sqrt{\xi^2 - 1} \cot(\Lambda \sqrt{\xi^2 - 1}), & \xi > 1.
\end{cases}
\quad (7.5.17)
\]

As a result, the “\( Z = 137 \) catastrophe” in Eq. (7.5.1) disappears and \( \mathcal{E}(1S_{1/2}) = 0 \) gives

\[
\xi_0 = 1 + \frac{\pi^2}{8\Lambda} + \mathcal{O}(\Lambda^{-4});
\quad (7.5.18)
\]

the state \( 1S_{1/2} \) energy continuously goes down to the negative energy continuum since \( Z\alpha > 1 \), and \( \mathcal{E}(1S_{1/2}) = -1 \) gives

\[
\xi_{cr} = 1 + \frac{\pi^2}{2\Lambda(\Lambda + 2)} + \mathcal{O}(\Lambda^{-4})
\quad (7.5.19)
\]
as shown in Fig. 7.5. In Ref. [65, 304] Popov and Zeldovich found that the critical value \( \xi_c^{(n)} = Z_c \alpha \) for the energy levels \( nS_{1/2} \) and \( nP_{1/2} \) reaching the negative energy continuum is equal to

\[
\xi_c^{(n)} = 1 + \frac{n^2\pi^2}{2\Lambda^2} + \mathcal{O}(\Lambda^{-3}).
\quad (7.5.20)
\]

The critical value increases rapidly with increasing \( n \). As a result, it is found that \( Z_{cr} \approx 173 \) is a critical value at which the lowest-energy level of the bound state \( 1S_{1/2} \) encounters the negative energy continuum, while other bound states encounter the negative energy continuum at \( Z_{cr} > 173 \) (see also Ref. [302] for a numerical estimation of the same spectrum). The change in the vacuum polarization near a high-Z nucleus arising from the finite extent of the nuclear charge density was computed in [308–310] with all calculations done analytically, and to all orders in \( Z\alpha \). Note that for two nuclei with charges \( Z_1 \) and \( Z_2 \) respectively, if \( Z_1 > Z_2 \) and \( K \)-shell of the \( Z_1 \)-nucleus is empty, then \( Z_2 \) may be a neutral atom. In this case two nuclei make a quasi-molecular state for which the ground term (1s\( \sigma \)) is unoccupied by electrons: so spontaneous production of positrons is also possible [311, 312]. We refer the readers to Ref. [65, 304–307, 313] for mathematical and numerical details.

When \( Z > Z_{cr} = 173 \), the lowest-energy level of the bound state \( 1S_{1/2} \) enters the negative energy continuum. Its energy level can be estimated as follows

\[
\mathcal{E}(1S_{1/2}) = m_e c^2 - \frac{Z\alpha}{\bar{r}} < -m_e c^2,
\quad (7.5.21)
\]
where \( \bar{r} \) is the average radius of the \( 1S_{1/2} \) state’s orbit, and the binding energy
of this state satisfies $Z\alpha/\bar{r} > 2m_e c^2$. If this bound state is unoccupied, the bare nucleus gains a binding energy $Z\alpha/\bar{r}$ larger than $2m_e c^2$, and becomes unstable against the production of an electron–positron pair. Assuming this pair production occurs around the radius $\bar{r}$, we have energies for the electron ($\epsilon_-$) and positron ($\epsilon_+$) given by

$$\epsilon_- = \sqrt{|cp_-|^2 + m_e^2 c^4 - \frac{Z\alpha}{\bar{r}}}; \quad \epsilon_+ = \sqrt{|cp_+|^2 + m_e^2 c^4 + \frac{Z\alpha}{\bar{r}}}, \quad (7.5.22)$$

where $p_\pm$ are electron and positron momenta, and $p_- = -p_+$. The total energy required for the production of a pair is

$$\epsilon_{-+} = \epsilon_- + \epsilon_+ = 2\sqrt{|cp_-|^2 + m_e^2 c^4}, \quad (7.5.23)$$

which is independent of the potential $V(\bar{r})$. The potential energies $\pm eV(\bar{r})$ of the electron and positron cancel out each other and do not contribute to the total energy (7.5.23) required for pair production. This energy (7.5.23) is acquired from the binding energy ($Z\alpha/\bar{r} > 2m_e c^2$) by the electron filling into the bound state $1S_{1/2}$. A part of the binding energy becomes the kinetic energy of positron that goes out. This is analogous to the familiar case when a proton ($Z = 1$) catches an electron into the ground state $1S_{1/2}$, and a photon is emitted with the energy not less than 13.6 eV. In the same way, more electron–positron pairs are produced, when $Z \gg Z_{cr}$ and the energy levels of the next bound states $2P_{1/2}, 2S_{3/2}, \ldots$ enter the negative energy continuum, provided these bound states of the bare nucleus are unoccupied.

### 7.5.4 Positron production

Gershtein and Zeldovich [314, 315] proposed that when $Z > Z_{cr}$ the bare nucleus produces spontaneously pairs of electrons and positrons: the two positrons$^2$ run off to infinity and the effective charge of the bare nucleus decreases by two electrons, which corresponds exactly to filling the K-shell$^3$ A more detailed investigation was made for the solution of the Dirac equation at $Z \sim Z_{cr}$, when the lowest electron level $1S_{1/2}$ merges with the negative energy continuum, in Refs. [304–307, 316]. It was there further clarified the situation, showing that at $Z \gtrsim Z_{cr}$, an imaginary resonance energy of Dirac equation appears,

$$\epsilon = \epsilon_0 - i\frac{\Gamma_{\text{nucl}}}{2}, \quad (7.5.24)$$

$^2$Hyperfine structure of $1S_{1/2}$ state: single and triplet.

$^3$An assumption was made in Ref. [314, 315] that the electron density of $1S_{1/2}$ state, as well as the vacuum polarization density, is delocalized at $Z \rightarrow Z_{cr}$. Later it was proved to be incorrect [65, 305, 306].
where

$$\epsilon_0 = -m_e - a(Z - Z_{cr}),$$

(7.5.25)

$$\Gamma_{\text{nucl}} \sim \theta(Z - Z_{cr}) \exp \left( -b \sqrt{\frac{Z_{cr}}{Z - Z_{cr}}} \right),$$

(7.5.26)

and $a, b$ are constants, depending on the cutoff $\Lambda$ (for example, $b = 1.73$ for $Z = Z_{cr} = 173$ [65,305,306]). The energy and momentum of emitted positrons are $|\epsilon_0|$ and $|p| = \sqrt{|\epsilon_0| - m_e c^2}$.

The kinetic energy of the two positrons at infinity is given by

$$\epsilon_p = |\epsilon_0| - m_e c^2 = a(Z - Z_{cr}) + \cdots,$$

(7.5.27)

which is proportional to $Z - Z_{cr}$ (as long as $(Z - Z_{cr}) \ll Z_{cr}$) and tends to zero as $Z \to Z_{cr}$. The pair production resonance at the energy (7.5.24) is extremely narrow and practically all positrons are emitted with almost same kinetic energy for $Z \sim Z_{cr}$, i.e. nearly mono-energetic spectra (sharp line structure). Apart from a pre-exponential factor, $\Gamma_{\text{nucl}}$ in Eq. (7.5.26) coincides with the probability of positron production, i.e., the penetrability of the Coulomb barrier (see Section 3.5). The related problems of vacuum charge density due to electrons filling into the K-shell and charge renormalization due to the change of wave function of electron states are discussed in Refs. [317–321]. An extensive and detailed review on this theoretical issue can be found in Refs. [65,66,313,322].

On the other hand, some theoretical work has been done studying the possibility that pair production, due to bound states encountering the negative energy continuum, is prevented from occurring by higher order processes of quantum field theory, such as charge renormalization, electron self-energy and nonlinearities in electrodynamics and even Dirac field itself [222,323–328]. However, these studies show that various effects modify $Z_{cr}$ by a few percent, but have no way to prevent the binding energy from increasing to $2m_e c^2$ as $Z$ increases, without simultaneously contradicting the existing precise experimental data on stable atoms [329]. Contrary claim [330] according to which bound states are repelled by the lower continuum through some kind of self-screening appear to be unfounded [66].

It is worth noting that an overcritical nucleus ($Z \geq Z_{cr}$) can be formed for example in the collision of two heavy nuclei [303,314–316,331–334]. To observe the emission of positrons originated from pair production occurring near to an overcritical nucleus temporally formed by two nuclei, the following necessary conditions have to be full filled: (i) the atomic number of an overcritical nucleus is larger than $Z_{cr} = 173$; (ii) the lifetime of the overcritical nucleus must be much longer than the characteristic time $(\hbar / m_e c^2)$ of pair production; (iii) the inner shells (K-shell) of the overcritical nucleus should be
unoccupied.

The collision of two Uranium nuclei with \( Z = 92 \) was considered by Zeldovich, Popov and Gershtein [65, 311]. The conservation of energy in the collision reads

\[
M_n v_0^2 = (Ze)^2 / R_{\text{min}},
\]

where \( v_0 \) is the relative velocity of the nuclei at infinity, \( R_{\text{min}} \) is the smallest distance, and \( M_n \) is the Uranium atomic mass. In order to have \( R_{\text{min}} \simeq 30 \text{fm} \) a fine tuning of the initial velocity narrowly peaked around \( v_0 \simeq 0.034c \) is needed. The characteristic collision time would be then \( \Delta t_c = R_{\text{min}} / v_0 \simeq 10^{-20} \text{s} \). The interesting possibility then occurs, that the typical velocity of an electron in the inner shell (\( r \sim 115.8 \text{fm} \)) is \( v \sim c \) and therefore its characteristic time \( \Delta \tau_0 \sim r / v \sim 4 \cdot 10^{-22} \text{s} \). This means that the characteristic collision time \( \Delta t_c \) in which the two colliding nuclei are brought into contact and separated again can be in principle much larger than the timescale \( \Delta t_0 \) of electron evolution. This would give justification for an adiabatic description of the collision in terms of quasimolecules. The formation of “quasimolecules” could also be verified by the characteristic molecular-orbital X-rays radiation due to the electron transitions between “quasimolecules” orbits [329, 335–341]. However, this requires the above mentioned fine tuning in the bombarding energies \( (M_n v_0^2 / 2) \) close to the nuclear Coulomb barrier.

However, we notice that the above mentioned condition (ii) has never been fulfilled in heavy-ion collisions. There has been up to now various unsuccessful attempts to broaden this time of encounter by ‘sticking’ phenomena. Similarly, the condition (iii) is not sufficient for pair production, since electrons that occupied outer shells of high energies must undergo a rapid transition to occupy inner shells of lower energies, which is supposed to be vacant and encountering the negative energy continuum. If such transition and occupation take place faster than pair production, the pair production process is blocked. As a consequence, it needs a larger value of \( Z > Z_{\text{cr}} = 173 \) to have stronger electric field for vacant out shells encountering the negative energy continuum (see Eq. (7.5.20)) so that electrons produced can occupy outer shells. This makes pair production even less probable to be observed, unless the overcritical charged nucleus is bare, i.e. all shells are vacant.

### 7.5.5 Homogeneous and adiabatic approximation

There is a certain analogy between positron production by a nucleus with \( Z > Z_{\text{cr}} \) and pair production in a homogeneous electrostatic field. We note that in a Coulomb potential of a nucleus with \( Z = Z_{\text{cr}} \) the corresponding electric field \( E_{\text{cr}} = Z_{\text{cr}} e / r^2 \) is comparable with the critical electric field \( E_c \), (2.0.1), when \( r \sim \lambda_c \). However, the condition \( E > E_c \) is certainly the necessary condition in order to have the pair creation but not a sufficient one: the spatial extent of the region where \( E > E_c \) occurs must be larger than the
de Broglie wavelength of the created electron—positron pair. If a pair production takes place, electrons should be bound into the K-shell nucleus and positrons should run off to infinity. This intuitive reasoning builds the connection between the phenomena of pair production in the Coulomb potential at charge $Z > Z_{cr}$ and the one in an external constant electric field which was treated in Section 5.7. The exact formula for pair production probability $W$ in an overcritical Coulomb potential has not yet been obtained in the framework of QED. We cannot expect a literal coincidence of formulas for the probability $W$ of pair production in these different cases, since the Schwinger formula (5.7.25) is exactly derived for a homogeneous field, while the Coulomb potential is strongly inhomogeneous at small distances. Some progress in the treatment of this problem is presented in Section 6.

All the discussions dealing with pair production in an external homogeneous electric field or a Coulomb potential assume that the electric field be static. Without the feedback of the particles created on the field this will clearly lead to a divergence of the number of pairs created. In the real description of the phenomenon at $t \to -\infty$ we have an initial empty vacuum state. We then have the turned on of an overcritical electric field and ongoing process of pair creation with their feedback on a time on the electric field and a final state at remote future $t \to +\infty$ with the electron and positron created and the remaining subcritical electric field. To describe this very different regimes a simplified “adiabatic approximation” can be adopted by assuming the existence of a homogeneous field only during a finite time interval $[-\mathcal{T}, \mathcal{T}]$. That time $\mathcal{T}$ should of course shorter than the feedback time. During that time interval the Schwinger formula (5.7.25) is assumed to be applicable and it is appropriate to remark that the overcritical electric fields are related to very high energy densities: $E_c^2/2 = 9.53 \cdot 10^{26}$ ergs/cm$^3$. In the adiabatic approximation an effective spatial limitation to the electric field is also imposed. Therefore the constant overcritical electric field and the application of the Schwinger formula is limited both in space and time. Progress in this direction has been presented in [73], see Section 9.8. A significant amount of pairs is only produced if the finite lifetime of the overcritical electric field is larger than the characteristic time of pair production ($\hbar/m_e c^2$) and the spatial extent of the electric field is larger than the tunneling length $a$ (3.5.6).

We have already discussed in Secs. 7.2 and 7.3 the experimental status of electron–positron pair creation in X-ray free electron laser and an electron-beam–laser collision, respectively. We now turn in Section 7.6 to the multiyear attempts in creating electron–positron pairs in heavy-ion collisions.
7.6 Pair production in heavy-ion collisions

7.6.1 A transient super heavy “quasimolecules”

There has been a multiyear effort to observe positrons from pair production associated with the overcritical field of two colliding nuclei, in heavy-ion collisions [314–316, 329, 333, 334, 342]. The hope was to use heavy-ion collisions to form transient superheavy “quasimolecules”: a long-lived metastable nuclear complex with $Z > Z_{cr}$. It was expected that the two heavy ions of charges respectively $Z_1$ and $Z_2$ with $Z_1 + Z_2 > Z_{cr}$ would reach small inter-nuclear distances well within the electron’s orbiting radii. The electrons would not distinguish between the two nuclear centers and they would evolve as if they were bounded by nuclear “quasimolecules” with nuclear charge $Z_1 + Z_2$. Therefore, it was expected that electrons would evolve quasi-statically through a series of well defined nuclear “quasimolecules” states in the two-center field of the nuclei as the inter-nuclear separation decreases and then increases again.

When heavy-ion collision occurs the two nuclei come into contact and some deep inelastic reaction occurs determining the duration $\Delta t_s$ of this contact. Such “sticking time” is expected to depend on the nuclei involved in the reaction and on the beam energy. Theoretical attempts have been proposed to study the nuclear aspects of heavy-ion collisions at energies very close to the Coulomb barrier and search for conditions, which would serve as a trigger for prolonged nuclear reaction times, to enhance the amplitude of pair production. The sticking time $\Delta t_s$ should be larger than $1 \sim 2 \cdot 10^{-21}$ sec [66] in order to have significant pair production, see Fig. 7.6. Up to now no success has been achieved in justifying theoretically such a long sticking time. In reality the characteristic sticking time has been found of the order of $\Delta t \sim 10^{-23}$ sec, hundred times shorter than the one needed to activate the pair creation process. Moreover, it is recognized that several other dynamical processes can make the existence of a sharp line corresponding to an electron–positron annihilation very unlikely [66, 329, 343–345].

It is worth noting that several other dynamical processes contribute to the production of positrons in undercritical as well as in overcritical collision systems [222, 323–325]. Due to the time-energy uncertainty relation (collision broadening), the energy spectrum of such positrons has a rather broad and oscillating structure, considerably different from a sharp line structure that we would expect from pair production positron emission alone.

7.6.2 Experiments

As remarked above, if the sticking time $\Delta t_s$ could be prolonged, the probability of pair production in vacuum around the superheavy nucleus would be enhanced. As a consequence, the spectrum of emitted positrons is expected
to develop a sharp line structure, indicating the spontaneous vacuum decay caused by the overcritical electric field of a forming superheavy nuclear system with $Z \geq Z_{cr}$. If the sticking time $\Delta t_s$ is not long enough and the sharp line of pair production positrons has not yet well developed, in the observed positron spectrum it is difficult to distinguish the pair production positrons from positrons created through other different mechanisms. Prolonging the “sticking time” and identifying pair production positrons among all other particles [335,346] created in the collision process has been an object of a very large experimental campaign [347–354].

For nearly 20 years the study of atomic excitation processes and in particular of positron creation in heavy-ion collisions has been a major research topic at GSI (Darmstadt) [355–358]. The Orange and Epos groups at GSI (Darmstadt) discovered narrow line structures (see Fig. 7.7) of unexplained origin, first in the single positron energy spectra and later in coincident electron–positron pair emission. Studying more collision systems with a wider range of the combined nuclear charge $Z = Z_1 + Z_2$ they found that narrow line structures were essentially independent of $Z$. This has ruled out the explanation as a pair production positron, since the line was expected to be at the position of the $1s\sigma$ resonance, i.e., at a kinetic energy given by Eq. (7.5.27), which is strongly $Z$ dependent. Attempts to link this positron line to spon-
taneous pair production have failed. Other attempts to explain this positron line in term of atomic physics and new particle scenario were not successful as well [66].

Figure 7.7: Two typical example of coincident electron–positron spectra measured by the Epose group in the system U+Th (left) and by the Orange group in U+Pb collisions (right). When plotted as a function of the total energy of the electron and positron, very narrow line structures were observed. This figure is reproduced from Fig. 7 in Ref. [66].

The anomalous positron line problem has perplexed experimentalists and theorists alike for more than a decade. Moreover, later results obtained by the Apex collaboration at Argonne National Laboratory showed no statistically significant positron line structures [359, 360]. This is in strong contradiction with the former results obtained by the Orange and Epos groups. However, the analysis of Apex data was challenged in the comment by Ref. [361, 362] pointing out that the Apex measurement would have been less sensitive to extremely narrow positron lines. A new generation of experiments (Apex at Argonne and the new Epos and Orange setups at GSI) with much improved counting statistics has failed to reproduce the earlier results [66].

To overcome the problem posed by the short timescale of pair production ($10^{-21}$ sec), hopes rest on the idea to select collision systems in which a nuclear reaction with sufficient sticking time occurs. Whether such a situation can be realized still is an open question [66]. In addition, the anomalous positron line problem and its experimental contradiction overshadow the field of the pair production in heavy-ion collisions.

In summary, clear experimental signals for electron–positron pair production in heavy-ion collisions are still missing [66] at the present time. For more recent information on the pair production in the heavy-ion collisions see [363–365] and for complete references the resource letter [366].
Having reviewed the situation of electron–positron pair creation by vacuum polarization in Earth-bound experiments we turn now to the corresponding problems in the realm of astrophysics. The obvious case is the one of black holes where the existence of critical field is clearly predicted by the analytic solutions of the Einstein-Maxwell field equations.
8 The extraction of blackholic energy from a black hole by vacuum polarization processes

It is becoming more and more clear that the theoretical description of the gravitational collapse process to a Kerr–Newman black hole, with all the aspects of nuclear physics and electrodynamics involved, is likely the most complex problem in physics and astrophysics. Specific to this report is the opportunity given by the process of gravitational collapse to study for the first time the above mentioned three quantum processes simultaneously at work under ultrarelativistic special and general relativistic regimes. The process of gravitational collapse is characterized by the gravitational timescale $\Delta t_g = GM/c^3 \simeq 5 \cdot 10^{-6}(M/M_\odot)$ sec, where $G$ is the gravitational constant, $M$ is the mass of a collapsing object, and the energy involved is of the order of $\Delta E = 10^{54}M/M_\odot$ ergs. This is one of the most energetic and most transient phenomena in physics and astrophysics and needs for its correct description the identification of the basic constitutive processes occurring in a highly time varying regime. Our approach in this Section is to proceed with an idealized model which can give us estimates of the basic energetics and some leading features of the real phenomenon. We shall describe: (1) the basic energetic process of an already formed black hole; (2) the vacuum polarization process a la Schwinger of an already formed Kerr–Newman black hole; (3) the basic formula of the dynamics of the gravitational collapse. We shall in particular recover the Tolman-Oppenheimer-Snyder solutions in a more explicit form and give exact analytic solution for the description of the gravitational collapse of charged and uncharged shells. This will allow, among others, to recall the mass formula of the black hole, to clarify the special role of the irreducible mass in that formula, and to have a general derivation of the maximum extractable energy in the process of gravitational collapse. We will as well address some conceptual issues between general relativity and thermodynamics which have been of interest to theoretical physicists in the last forty years. Of course in this brief Section we will be only recalling some of these essential themes and refer to the literature where in-depth analysis can be found. Since we are interested in the gravitational collapse we are going to examine only processes involving masses larger than the critical mass of neutron stars, for convenience established in $3.2 M_\odot$ [367]. We are consequently
8 The extraction of blackholic energy from a black hole by vacuum polarization processes

not addressing the research on mini-black-holes [368] which involves energetics $10^{41}$ times smaller than the ones involved in gravitational collapse and discussed in this report. This problematics implies a yet unknown physics of applying quantum mechanics in conditions where the curvature of space-time is comparable to the wavelength of the particle.

We recall here the basic steps leading to the study of the electrodynamics of a Kerr–Newman black hole, indicating the relevant references. In this Section we use the system of units $c = G = \hbar = 1$.

8.1 Test particles in Kerr–Newman geometries

According to the uniqueness theorem for stationary, regular black holes (see Ref. [369]), the process of gravitational collapse of a core whose mass is larger than the neutron star critical mass [367] will generally lead to a black hole characterized by all the three fundamental parameters: the mass-energy $M$, the angular momentum $L$, and the charge $Q$ (see [370]). The creation of critical electric fields and consequent process of pair creation by vacuum polarization are expected to occur in the late phases of gravitational collapse when the gravitational energy of the collapsing core is transformed into an electromagnetic energy and eventually in electron–positron pairs. As of today no process of the gravitational collapse either to a neutron star or to a black hole has reached a satisfactory theoretical understanding. It is a fact that even the theory of a gravitational collapse to a neutron star via a supernova is not able to explain even the ejection of a supernova remnant [371].

In order to estimate the fundamental energetics of these transient phenomena we recall first the metric of a Kerr–Newman black hole, the role of the reversible and irreversible transformations in reaching the mass formula as well as the role of the positive and negative energy states in a quantum analog. We will then estimate the energy emission due to vacuum polarization process. As we will see, such a process occurs on characteristic quantum timescale of $t \sim \frac{\hbar}{(m_e c^2)} \sim 10^{-21}$ sec, which is many orders of magnitude shorter than the characteristic gravitational collapse timescale. Of course the astrophysical progenitor of the black hole will be a neutral one, as all the astrophysical systems. Only during the process of the gravitational collapse and for the above mentioned characteristic gravitational timescale a process of charge separation will occur. The positively charged core would give rise to the electrodynamical process approaching asymptotically in time the horizon of a Kerr–Newman black hole.

A generally charged and rotating, black hole has been considered whose curved space-time is described by the Kerr–Newman geometry [372]. In
Kerr–Newman coordinates \((u, r, \theta, \phi)\) the line element takes the form,

\[
ds^2 = \Sigma d\theta^2 - 2a \sin^2 \theta dr d\phi + 2dr du - 2a\Sigma^{-1}(2Mr - Q^2) \sin^2 \theta d\phi du \\
+ \Sigma^{-1}[(r^2 + a^2)^2 - \Delta a^2 \sin^2 \theta] \sin^2 \theta d\phi^2 - [(1 - \Sigma^{-1}(2Mr - Q^2))] du^2
\]

(8.1.1)

where \(\Delta = r^2 - 2Mr + a^2 + Q^2\) and \(\Sigma = r^2 + a^2 \cos^2 \theta\), \(a = L/M\) being the angular momentum per unit mass of the black hole. The Reissner–Nordström and Kerr geometries are particular cases for a nonrotating, \(a = 0\), and uncharged, \(Q = 0\), black holes respectively. The Kerr–Newman space–time has a horizon at

\[
r = r_+ = M + (M^2 - Q^2 - a^2)^{1/2}
\]

(8.1.2)

where \(\Delta = 0\).

The electromagnetic vector potential around the Kerr–Newman black hole is given by [372]

\[
A = -\Sigma^{-1}Qr(du - a \sin^2 \theta d\phi),
\]

(8.1.3)

the electromagnetic field tensor is then

\[
F = dA = 2Q\Sigma^{-2}[(r^2 - a^2 \cos^2 \theta)dr \wedge du - 2a^2r \cos \theta \sin \theta d\theta \wedge du \\
- a \sin^2 \theta(r^2 - a^2 \cos^2 \theta)dr \wedge d\phi + 2ar(r^2 + a^2) \cos \theta \sin \theta d\theta \wedge d\phi].
\]

(8.1.4)

The equation of motion of a test particle of mass \(m\) and charge \(e\) in the Kerr–Newman geometry reads

\[
u^\mu \nabla_\mu u^\nu = (e/m)u^\mu F_{\mu \nu},
\]

(8.1.5)

where \(u^\mu\) is the 4–velocity of the particle. These equations may be derived from the Lagrangian

\[
L = \frac{1}{2}mg_{\mu\nu}u^\mu u^\nu + eA_\mu u^\mu,
\]

(8.1.6)

or, equivalently, from the Hamiltonian

\[
H = \frac{1}{2}g^{\mu\nu}(p_\mu - eA_\mu)(p_\nu - eA_\nu),
\]

(8.1.7)

where we have introduced the 4–momentum of the particle

\[
p_\mu = mu_\mu + eA_\mu.
\]

(8.1.8)

Note that Hamiltonian (8.1.7) is subject to the constraint

\[
H = -\frac{1}{2}m^2.
\]

(8.1.9)

Carter [67] firstly recognized that the corresponding Hamilton–Jacobi equa-
The extraction of blackholic energy from a black hole by vacuum polarization processes

8.1 Integrable systems

\[ g^{\alpha\beta} \left( \frac{\partial S}{\partial x^\alpha} + eA_\alpha \right) \left( \frac{\partial S}{\partial x^\beta} + eA_\beta \right) + m^2 = 0, \quad (8.1.10) \]

are separable. Correspondingly four integrals of the equation of motion (8.1.5) can be found. Indeed, in addition to the constant of motion (8.1.9) which corresponds to conservation of the rest mass we have the two first integrals

\[ p_u = -\mathcal{E} \quad (8.1.11) \]
\[ p_\phi = \Phi \quad (8.1.12) \]

associated with the stationarity and the axial symmetry of Kerr–Newman space-time respectively. \( \mathcal{E} \) and \( \Phi \) are naturally interpreted as the energy and the angular momentum about the symmetry axis of the test particle. It follows from the separability of Eq. (8.1.10) that the quantities

\[ p_{\theta}^2 + (a\mathcal{E} \sin \theta - \Phi \sin^{-1} \theta)^2 + a^2 m^2 \cos^2 \theta = K \quad (8.1.13) \]
\[ \Delta p_r^2 - 2[(r^2 + a^2)\mathcal{E} + eQr - a\Phi]p_r + m^2 r^2 = -K \quad (8.1.14) \]

are conserved as well. Together with \( \mathcal{E} \) and \( \Phi \) they form a complete set of first integrals of the motion and allow one to integrate Eq. (8.1.5). As an example consider the proper time derivative \( \dot{r} \) of the radial coordinate of the test particle. It follows from Eqs. (8.1.11), (8.1.12), (8.1.13) and (8.1.14), that

\[ \Sigma^2 \dot{r}^2 = (\mathcal{E}(r^2 + a^2) + eQr - \Phi a)^2 - \Delta (m^2 r^2 + K) \quad (8.1.15) \]

which can be numerically integrated using the effective potential technique [373].

8.2 Reversible and irreversible transformations of a black hole: the Christodoulou-Ruffini mass formula

In 1969 Roger Penrose [374] pointed out for the first time the possibility to extract rotational energy from a Kerr black hole. The first example of such an energy extraction was obtained by Ruffini and Wheeler who also introduced the concept of the ergosphere [177, 370, 375], the region between the horizon of the black hole and the surface of infinite redshift. These works has been generalized by Denardo and Ruffini in 1973 [376] and Denardo, Hively and Ruffini in 1974 [377] to the case of a Kerr–Newman black hole. The process described by Denardo, Hively and Ruffini can be described as follows. A neutral particle \( P_0 \) approaches the black hole with positive energy \( \mathcal{E}_0 \) and decays into two oppositely charged particles \( P_1 \) and \( P_2 \) whose energies are
\[ \delta E = E_2 - E_0 \] (8.2.1)

has been extracted from the black hole and deposited on \( P_2 \). The region around the black hole where the energy extraction processes can occur is named effective ergosphere in Refs. [376], [377]. Note that, as the particle \( P_1 \) is swallowed, the black hole undergoes a transformation since its energy, angular momentum and charge change accordingly. When is the extracted energy maximal? In order to answer this question note that the energy \( E \) of a particle of angular momentum \( \Phi \), charge \( e \) and rest mass \( m \) moving around a Kerr–Newman black hole and having a turning point at \( r \) is given by (see Eq. (8.1.15)) the quadratic equation

\[
(r^2 + a^2)E^2 + 2(eQr - \Phi a)(r^2 + a^2)E + (eQr - \Phi a)^2 - \Delta m^2 r^2 + K = 0. \tag{8.2.2}
\]

As recalled in [114, p. 352], in the case of \( Q = 0 \) which corresponds to a pure Kerr solution, the explicit integration of this equation was performed by Ruffini and Wheeler [378]. They introduced the effective potential energy defined by

\[
(r^2 + a^2)E^2 - 2(\Phi a)(r^2 + a^2)E + (\Phi a)^2 - \Delta m^2 r^2 + K = 0. \tag{8.2.3}
\]

The radii of stable orbits are determined by minimum of function \( E(r) \), i.e. by simultaneous solution of equations \( E(r) = E_0, E'(r) = 0 \) for \( E''(r) > 0 \). The orbit closest to the center corresponds to \( E''(r)_{\text{min}} = 0 \); for \( r < r_{\text{min}} \), the function \( E(r) \) has no minima. As a result

- When \( \Phi < 0 \) (motion opposite to the direction of rotation of the collapsing object)

\[
\frac{r_{\text{min}}}{2M} = \frac{9}{2}, \quad \frac{E_0}{m} = \frac{5}{3\sqrt{3}}, \quad \frac{\Phi}{2mM} = \frac{11}{3\sqrt{3}}. \tag{8.2.4}
\]

- For \( \Phi > 0 \) (motion in the direction of rotation of the collapsing object) as \( a \to M \) the radius \( r_{\text{min}} \) tends towards the radius of the horizon. Setting \( a = M(1 + \delta) \), we find in the limit \( \delta \to 0 \):

\[
\frac{r_{\text{hor}}}{2M} = \frac{1}{2}(1 + \sqrt{2\delta}), \quad \frac{r_{\text{min}}}{2m} = \frac{1}{2}[1 + (4\delta)^{1/3}]. \tag{8.2.5}
\]

Then

\[
\frac{E_0}{m} = \frac{\Phi}{2mM} = \frac{1}{\sqrt{3}}[1 + (4\delta)^{1/3}]. \tag{8.2.6}
\]
We call attention to the fact that \( r_{\text{min}} / r_{\text{hor}} \) remains greater than one throughout, i.e. the orbit does not go inside the horizon. This is as it should be: the horizon is a null hypersurface, and no time-like world lines of moving particles can lie on it. Although no general formula exists in the case of the Kerr–Newman geometry the energy and the angular velocity of a test particle in a circular orbit with radius \( R \) in the Reissner–Nordström geometry has been given by Ruffini and Zerilli [373]

\[
\dot{\phi}^2 = \frac{M}{R^3} - \frac{Q^2}{R^4} - \frac{e}{m} \frac{Q}{R^3} \left[ \frac{e}{m} \frac{Q}{2R} + \left( 1 - \frac{3M}{R} + \frac{2Q^2}{R^2} + \frac{e}{m} \frac{Q^2}{4R^2} \right)^{1/2} \right],
\]

(8.2.7)

\[
\mathcal{E} = \frac{1}{m} \left( 1 - \frac{2M}{R} + \frac{Q^2}{R^2} \right) / \left[ \frac{e}{m} \frac{Q}{2R} + \left( 1 - \frac{3M}{R} + \frac{2Q^2}{R^2} + \frac{e}{m} \frac{Q^2}{4R^2} \right)^{1/2} \right]
\]

(8.2.8)

and the limiting cases are there treated.

Eq. (8.2.2) is not only relevant for understanding the fully relativistic stable circular orbit but it also defines the “positive root states” and the “negative root states” for the particle [379]. Such states were first interpreted as limits of states of a quantum field by Deruelle and Ruffini [380]. Such an interpretation will be discussed in the next section. Note that in the case \( eQr - \Phi a < 0 \) there can exist negative energy states of positive root solutions and, as a direct consequence, energy can be extracted from a Kerr–Newman black hole via the Denardo-Ruffini process. Such a process is most efficient when the reduction of mass is greatest for a given reduction in angular momentum. To meet this requirement the energy \( \mathcal{E}_1 \) must be as negative as possible. This happens when \( r = r_+ \), that is the particle has a turning point at the horizon of the black hole. When \( r = r_+ \), \( \Delta = 0 \) and the separation between negative and positive root states vanishes. This implies that capture processes from such an orbit are reversible since they can be inverted bringing the black hole to its original state. Correspondingly the energy of the incoming particle is

\[
\mathcal{E}_1 = \frac{a\Phi + eQr_+}{a^2 + r_+^2}.
\]

(8.2.9)

If we apply the conservation of energy, angular momentum and charge to the capture of the particle \( P_1 \) by the black hole, we find that \( M, L \) and \( Q \) change as for the quantities

\[
dM = \mathcal{E}_1, \quad dL = \Phi, \quad dQ = e.
\]

(8.2.10)
8.2 Reversible and irreversible transformations of a black hole: the Christodoulou-Ruffini mass formula

Thus Eq. (8.2.9) reads

$$dM = \frac{a \, dL + r_+ Q \, dQ}{a^2 + r_+^2}. \quad (8.2.11)$$

Integration of Eq. (8.2.11) gives

$$M^2 c^4 = \left( M_{ir} c^2 + \frac{c^2 Q^2}{4 G M_{ir}} \right)^2 + \frac{L^2 c^8}{4 G^2 M_{ir}^2}, \quad (8.2.12)$$

provided the condition is satisfied

$$\left( \frac{c^2}{16 G^2 M_{ir}^4} \right) \left( Q^4 + 4 L^2 c^4 \right) \leq 1, \quad (8.2.13)$$

where $M_{ir}$ is an integration constant and we restored the physical constants $c$ and $G$. Eq. (8.2.12) is the Christodoulou–Ruffini mass formula [379] and it expresses the contributions to the total energy of the black hole. Extreme black holes satisfy equality (8.2.13). The irreducible mass $M_{ir}$ satisfies the equation [379]

$$S_a = \frac{16 \pi G^2 M_{ir}^2}{c^4} \quad (8.2.14)$$

where $S_a$ is the surface area of the horizon of the black hole, and cannot be decreased by classical processes. Any transformation of the black hole which leaves fixed the irreducible mass (for instance, as we have seen, the capture of a particle having a turning point at the horizon of the black hole) is called reversible [379]. Any transformation of the black hole which increases its irreducible mass, for instance, the capture of a particle with nonzero radial momentum at the horizon, is called irreversible. In irreversible transformations there is always some kinetic energy that is irretrievably lost behind the horizon. Note that energy can be extracted approaching arbitrarily close to reversible transformations which are the most efficient ones. Namely, from Eq. (8.2.12) it follows that up to 29% of the mass-energy of an extreme Kerr black hole ($M^2 = a^2$) can be stored in its rotational energy term $\frac{L c^4}{2 G M_{ir}}$ and can in principle be extracted. Gedankenexperiments have been conceived to extract such energy [374, 381–384]. The first specific example of a process of energy extraction from a black hole can be found in R. Ruffini and J. A. Wheeler, as quoted in [385], see also [386]. Other processes of rotational energy extraction of astrophysical interest based on magnetohydrodynamic mechanism occurring around a rotating Black Hole have also been advanced [381–384] though their reversibility as defined in Ref. [379], and consequently their efficiency of energy extraction, has not been assessed. From the same mass formula (8.2.12) it follows that up to 50% of the mass energy of an extreme black hole with ($Q = M$) can be stored in the electromagnetic term $\frac{c^2 Q^2}{4 G M_{ir}}$ and can be in principle extracted. These extractable energies either rotational or elec-
The extraction of blackholic energy from a black hole by vacuum polarization processes

tromagnetic will be indicated in the following as blackholic energy and they can be the source of some of the most energetic phenomena in the Universe like jets from active galactic nuclei and GRBs.

8.3 Positive and negative root states as limits of quantum field states

In 1974 Deruelle and Ruffini [380] pointed out that negative root solutions of Eq. (8.2.2) can be interpreted in the framework of a fully relativistic quantum field theory as classical limits of antimatter solutions. In this section we briefly review their analysis. The equation of motion of a test particle in a Kerr–Newman geometry can be derived by the Hamilton–Jacobi Eq. (8.1.10). The first quantization of the corresponding theory can be obtained by substituting the Hamilton–Jacobi equation with the generalized Klein–Gordon equation

$$g^{\alpha\beta}(\nabla_\alpha + ieA_\alpha)(\nabla_\beta + ieA_\beta)\Phi + m^2\Phi = 0 \quad (8.3.1)$$

for the wave function $\Phi$. For simplicity we restrict to the Kerr case: $Q = 0$, when Eq. (8.3.1) reduces to

$$g^{\alpha\beta}\nabla_\alpha\nabla_\beta\Phi + m^2\Phi = 0. \quad (8.3.2)$$

In order to solve Eq. (8.3.2) we can separate the variables as follows:

$$\Phi = e^{-imEt}e^{ik\phi}S_{kl}(\theta)R(r) \quad (8.3.3)$$

where $S_{kl}(\theta)$ are spheroidal harmonics. We thus obtain the radial equation

$$\frac{d^2u}{dr^2} = \left\{ -E^2m^2\left(1 + \frac{a^2}{r^2} + \frac{2Ma^2}{r^3}\right) + \frac{4MakEm}{r^3} + m^2\left(1 - \frac{2M}{r} + \frac{a^2}{r^2}\right) - mk\frac{2M}{r^3} - \frac{1}{r^2} - \frac{a^2}{r^4} - \frac{k^2a^2}{r^4} + \frac{2}{r^6}\left[Mr^3 - r^2(a^2 + 2M^2) + 3Ma^2r - a^4\right] \right\} u,\]

where $u = R(r)r$ and $dr/dr^* = \Delta/r^2$. It is natural to look for “resonances” states of the Klein–Gordon equation corresponding to classical bound states (circular or elliptic orbits). Then, impose as boundary conditions (a) an exponential decay of the wave function for $r \to \infty$ and (b) a purely ingoing wave at the horizon $r \to r_+$. The solutions of the corresponding problem can be found numerically [380]. The main conclusions of the integration can be summarized as follows:

1. The continuum spectrum of the classical stable bound states is replaced by a discrete spectrum of resonances with tunneling through the potential barrier giving the finite probability of the particle to be captured by the horizon.
2. In the classical limit \((GM/c^2)/(\hbar/m_ec) \rightarrow \infty\) the separation of the energy levels of the resonances tends to zero. The leakage toward the horizon also decreases and the width of the resonance tends to zero.

3. The negative root solutions of Eq. (8.2.2) correspond to the classical limit \((GM/c^2)/(\hbar/m_ec) \rightarrow \infty\) of the negative energy solutions of the Klein–Gordon Eq. (8.3.2) and consequently they can be thought of as antimatter solutions with an appropriate interchange of the sign of charge, the direction of time and the angular momentum.

4. We can have positive root states of negative energy in the ergosphere, see e.g. [375]. In particular we can have crossing of positive and negative energy root states. This corresponds, at the second quantized theory level to the possibility of particle pair creation à là Klein, Sauter, Heisenberg, Euler and Schwinger [7, 17, 18, 20, 25–27].

Similar considerations can be made in the Kerr–Newman case, \(Q \neq 0\), when the generalized Klein Gordon Eq. (8.3.1) has to be integrated. The resonance states can be obtained imposing the same boundary conditions as above. Once again we can have level crossing inside the effective ergosphere [376, 377] and therefore possible pair creation.

### 8.4 Vacuum polarization in Kerr–Newman geometries

We discussed in the previous Sections the phenomenon of electron–positron pair production in a strong electric field in a flat space-time. Here we study the same phenomenon occurring around a black hole endowed with mass \(M\), charge \(Q\) and the angular momentum \(a\).

The space-time of a Kerr–Newman geometry is described by a metric which in Boyer–Lindquist coordinates \((t, r, \theta, \phi)\) acquires the form

\[
ds^2 = \frac{\Sigma}{\Delta} dr^2 + \Sigma d\theta^2 - \frac{\Delta}{\Sigma} (dt - a \sin^2 \theta d\phi)^2 + \frac{\sin^2 \theta}{\Sigma} \left[ (r^2 + a^2) d\phi - a dt \right]^2,
\]

(8.4.1)

where \(\Delta\) and \(\Sigma\) are defined following (8.1.1). We recall that the Reissner–Nordstrom geometry is the particular case \(a = 0\) of a nonrotating black hole.

The electromagnetic vector potential around the Kerr–Newman black hole is given in Boyer–Lindquist coordinates by

\[
A = -Q\Sigma^{-1} r(dt - a \sin^2 \theta d\phi).
\]

(8.4.2)
The electromagnetic field tensor is then
\[ F = dA = 2Q\Sigma^{-2}[(r^2 - a^2 \cos^2 \theta)dr \wedge dt - 2a^2 r \cos \theta \sin \theta d\theta \wedge dt - a \sin^2 \theta (r^2 - a^2 \cos^2 \theta)dr \wedge d\phi + 2ar(r^2 + a^2) \cos \theta \sin \theta d\theta \wedge d\phi]. \] (8.4.3)

After some preliminary work in Refs. [387–389], the occurrence of pair production in a Kerr–Newman geometry was addressed by Deruelle [390]. In a Reissner–Nordström geometry, QED pair production has been studied by Zaumen [391] and Gibbons [392]. The corresponding problem of QED pair production in the Kerr–Newman geometry was addressed by Damour and Ruffini [68], who obtained the rate of pair production with particular emphasis on:

- the limitations imposed by pair production on the strength of the electromagnetic field of a black hole [373];
- the efficiency of extracting rotational and Coulomb energy (the “black-holic” energy) from a black hole by pair production;
- the possibility of having observational consequences of astrophysical interest.

In the following, we recall the main results of the work by Damour and Ruffini.

In order to study the pair production in the Kerr–Newman geometry, they introduced at each event \((t, r, \theta, \phi)\) a local Lorentz frame associated with a stationary observer \(O\) at the event \((t, r, \theta, \phi)\). A convenient frame is defined by the following orthogonal tetrad [67]

\[
\omega^{(0)} = (\Delta/\Sigma)^{1/2}(dt - a \sin^2 \theta d\phi),
\] (8.4.4)

\[
\omega^{(1)} = (\Sigma/\Delta)^{1/2}dr,
\] (8.4.5)

\[
\omega^{(2)} = \Sigma^{1/2}d\theta,
\] (8.4.6)

\[
\omega^{(3)} = \sin \theta \Sigma^{-1/2}((r^2 + a^2)d\phi - adt).
\] (8.4.7)

In this Lorentz frame, the electric potential \(A_0\), the electric field \(E\) and the magnetic field \(B\) are given by the following formulas (c.e.g. Ref. [393]),

\[
A_0 = \omega^{(0)}_a A_a,
\]

\[
E^a = \omega^{(0)}_\beta F^{a\beta},
\]

\[
B^\beta = \frac{1}{2} \omega^{(0)}_\gamma \epsilon^{\gamma\delta\beta} F_{\delta\beta}.
\]
We then obtain
\[ A_0 = -Q(r\Sigma\Delta)^{-1/2}, \]  
while the electromagnetic fields \( E \) and \( B \) are parallel to the direction of \( \omega^{(1)} \) and have strengths given by
\[
E^{(1)} = Q\Sigma^{-2}(r^2 - a^2\cos^2 \theta), \tag{8.4.9}
\]
\[
B^{(1)} = Q\Sigma^{-2}ar\cos \theta, \tag{8.4.10}
\]
respectively. The maximal strength \( E_{\text{max}} \) of the electric field is obtained in the case \( a = 0 \) at the horizon of the black hole: \( r = r_+ \). We have
\[
E_{\text{max}} = Q/r_+^2. \tag{8.4.11}
\]

In the original paper a limit on the black hole mass \( M_{\text{max}} \simeq 7.2 \cdot 10^6 M_\odot \) was established by requiring that the pair production process would last less than the age of the Universe. For masses much smaller than this absolute maximum mass the pair production process can drastically modify the electromagnetic structure of black hole.

Both the gravitational and the electromagnetic background fields of the Kerr–Newman black hole are stationary when considering the quantum field of the electron. Since \( m_e M \simeq 10^{14} \gg 1 \) the gravitational field of the background black hole is practically constant over the Compton wavelength of the electron characterizing the quantum field. As far as purely QED phenomena such as pair production are concerned, it is possible to consider the electric and magnetic fields defined by Eqs. (8.4.9,8.4.10) as constants in the neighborhood of a few wavelengths around any events \((r, \theta, \phi, t)\). Thus, the analysis and discussion on the Sauter-Euler-Heisenberg-Schwinger process over a flat space-time can be locally applied to the case of the curved Kerr–Newman geometry, based on the equivalence principle.

The rate of pair production around a Kerr–Newman black hole can be obtained from the Schwinger formula (5.7.39) for parallel electromagnetic fields \( \varepsilon = E^{(1)} \) and \( \beta = B^{(1)} \) as:
\[
\tilde{\Gamma} = \frac{aE^{(1)}B^{(1)}}{4\pi^2} \sum_{n=1}^{\infty} \frac{1}{n} \coth \left( \frac{n\pi B^{(1)}}{E^{(1)}} \right) \exp \left( -\frac{n\pi E_e}{E^{(1)}} \right).
\tag{8.4.12}
\]
The total number of pairs produced in a region \( D \) of the space-time is
\[
N = \int_D \, d^4x \sqrt{-g} \frac{\tilde{\Gamma}}{V}, \tag{8.4.13}
\]
where \( \sqrt{-g} = \Sigma \sin \theta \). In Ref. [68], it was assumed that for each created pair the particle (or antiparticle) with the same sign of charge as the black hole was
The extraction of blackholic energy from a black hole by vacuum polarization processes

expelled to infinity with charge $e$, energy $\omega$ and angular momentum $l_\phi$ while the antiparticle was absorbed by the black hole. This implies the decrease of charge, mass and angular momentum of the black hole and a corresponding extraction of all three quantities. These considerations, however, were profoundly modified later by the introduction of the concept of dyadosphere which is presented in the next section. The rates of change of the charge, mass and angular momentum were estimated by

$$\dot{Q} = -Re,$$
$$\dot{M} = -R\langle \omega \rangle,$$
$$\dot{L} = -R\langle l_\phi \rangle,$$  (8.4.14)

where $R = \dot{N}$ is the rate of pair production and $\langle \omega \rangle$ and $\langle l_\phi \rangle$ represent some suitable mean values for the energy and angular momentum carried by the pairs.

Supposing the maximal variation of black hole charge to be $\Delta Q = -Q$, one can estimate the maximal number of pairs created and the maximal mass-energy variation. It was concluded in Ref. [68] that the maximal mass-energy variation in the pair production process is larger than $10^{44}$ erg and up to $10^{58}$ erg, depending on the black hole mass, see Table 1 in [68]. They concluded at the time “this work naturally leads to a most simple model for the explanation of the recently discovered $\gamma$-ray bursts”.

8.5 The “Dyadosphere” in Reissner–Nordström geometry

After the discovery in 1997 of the afterglow of GRBs [394] and the determination of the cosmological distance of their sources, at once of the order of $10^3$ theories explaining them were wiped out on energetic grounds. On the contrary, it was noticed [395,396] the coincidence between their observed energetics and the one theoretically predicted by Damour and Ruffini [68] of $10^{54}$ ergs per burst for $M = M_\odot$. Ruffini and collaborators therefore, indirectly motivated by GRBs, returned to these theoretical results with renewed interest developing some additional basic theoretical concepts [395,397–402] such as the dyadosphere and, more recently, the dyadotorus. In this Section we restore constants $G$, $c$ and $\hbar$ for clarity.

As a first simplifying assumption the case of absence of rotation was considered. The space-time is then described by the Reissner–Nordström geometry, see (8.4.1) whose spherically symmetric metric is given by

$$d^2 s = g_{tt}(r)d^2 t + g_{rr}(r)d^2 r + r^2 d^2 \theta + r^2 \sin^2 \theta d^2 \phi,$$  (8.5.1)
where $g_{tt}(r) = -\left[1 - \frac{2GM}{c^2 r} + \frac{Q^2 c^2}{4 r^2}\right] \equiv -\alpha^2(r)$ and $g_{rr}(r) = \alpha^{-2}(r)$.

The first result obtained is that the pair creation process does not occur at the horizon of the black hole: it extends over the entire region outside the horizon in which the electric field exceeds the value $E^*$ of the order of magnitude of the critical value given by Eq. (2.0.1). We recall the pair creation process is a quantum tunneling between the positive and negative energy states, which needs a level crossing, can occur for $E^* < E_c$ as well, if the field extent to spatial dimension $D^*$ such that $D^* E^* = 2m_e c^2 / e$. The probability of such pair creation process will be exponentially damped by $\exp(-\pi D^* / \lambda_c)$. Clearly, very intense process of pair creation will occur for $E^* > E_c$. In order to give a scale of the phenomenon, and for definiteness, in Ref. [398] it was considered the case of $E^* \equiv E_c$, although later in order to take into due account the tunneling effects we have considered dyadosphere for electric field in the range

$$E^* = \kappa E_c,$$

(8.5.2)

with $\kappa$ in the range 0.1-10. Since the electric field in the Reissner–Nordström geometry has only a radial component given by [403]

$$E(r) = \frac{Q}{r^2},$$

(8.5.3)

this region extends from the horizon radius, for $\kappa = 1$

$$r_+ = 1.47 \cdot 10^5 \mu (1 + \sqrt{1 - \xi^2}) \text{ cm}$$

(8.5.4)

out to an outer radius [395]

$$r^* = \left(\frac{\hbar}{m_e c}\right)^{1/2} \left(\frac{GM}{c^2}\right)^{1/2} \left(\frac{m_p}{m_e}\right)^{1/2} \left(\frac{e}{\eta_p}\right)^{1/2} \left(\frac{Q}{\sqrt{GM}}\right)^{1/2} = 1.12 \cdot 10^8 \sqrt{\mu \xi} \text{ cm},$$

(8.5.5)

where we have introduced the dimensionless mass and charge parameters $\mu = \frac{M}{M_\odot}, \xi = \frac{Q}{(M\sqrt{G})} \leq 1$, see Fig. 8.2.

The second result gave the local number density of electron and positron pairs created in this region as a function of radius

$$n_{e^+e^-}(r) = \frac{Q}{4\pi r^2 \left(\frac{\hbar}{m_e c}\right)^2} e \left[1 - \left(\frac{r}{r^*}\right)^2\right],$$

(8.5.6)

and consequently the total number of electron and positron pairs in this re-
The extraction of blackholic energy from a black hole by vacuum polarization processes

gion is

\[ N_{e^+e^-}^0 \simeq \frac{Q - Q_c}{e} \left[ 1 + \frac{(r^* - r_+)}{\frac{\hbar}{m_e c}} \right], \quad (8.5.7) \]

where \( Q_c = E_c r_+^2 \).

**Figure 8.1:** The energy extracted by the process of vacuum polarization is plotted (solid lines) as a function of the mass \( M \) in solar mass units for selected values of the charge parameter \( \xi = 1, 0.1, 0.01 \) (from top to bottom) for a Reisner-Nordström black hole, the case \( \xi = 1 \) reachable only as a limiting process. For comparison we have also plotted the maximum energy extractable from a black hole (dotted lines) given by Eq. (8.2.12). Details in Ref. [397].

The total number of pairs is larger by an enormous factor \( r^*/(\hbar/m_e c) > 10^{18} \) than the value \( Q/e \) which a naive estimate of the discharge of the black hole would have predicted. Due to this enormous amplification factor in the number of pairs created, the region between the horizon and \( r^* \) is dominated
Figure 8.2: The dyadosphere of a Reissner–Nordström black hole can be represented as equivalent to a concentric set of capacitor shells, each one of thickness $\frac{\hbar}{m_e c}$ and producing a number of $e^+e^-$ pairs of the order of $\sim \frac{Q}{e}$ on a time scale of $10^{-21}$ s, where $Q$ is the black hole charge. The shells extend in a region of thickness $\Delta r$, from the horizon $r_+$ out to the Dyadosphere outer radius $r_{ds}$ (see text). The system evolves to a thermalized plasma configuration.

by an essentially high density neutral plasma of electron–positron pairs. This region was defined [395] as the dyadosphere of the black hole from the Greek duas, duados for pairs. Consequently we have called $r^*$ the dyadosphere radius $r^* \equiv r_{ds}$ [395,397,398]. The vacuum polarization process occurs as if the entire dyadosphere is subdivided into a concentric set of shells of capacitors each of thickness $\hbar/m_e c$ and each producing a number of $e^+e^-$ pairs on the order of $\sim Q/e$ (see Fig. 8.2). The energy density of the electron–positron
The extraction of blackholic energy from a black hole by vacuum polarization processes

Pairs is there given by

\[ \epsilon(r) = \frac{Q^2}{8\pi r^4} \left( 1 - \left( \frac{r}{r_{ds}} \right)^4 \right), \tag{8.5.8} \]

(see Figs. 2–3 of Ref. [397]). The total energy of pairs converted from the static electric energy and deposited within the dyadosphere is then

\[ E_{\text{dya}} = \frac{1}{2} \frac{Q^2}{r_+} \left( 1 - \frac{r_+}{r_{ds}} \right) \left[ 1 - \left( \frac{r_+}{r_{ds}} \right)^4 \right], \tag{8.5.9} \]

In the limit \( \frac{r_+}{r_{ds}} \to 0 \), Eq. (8.5.9) leads to \( E_{\text{dya}} \to \frac{1}{2} \frac{Q^2}{r_+} \), which coincides with the energy extractable from black holes by reversible processes (\( M_{\text{ir}} = \text{const.} \)), namely \( E_{BH} - M_{\text{ir}} = \frac{1}{2} \frac{Q^2}{r_+} \) [379], see Fig. 8.1. Due to the very large pair density given by Eq. (8.5.6) and to the sizes of the cross-sections for the process \( e^+e^- \leftrightarrow \gamma + \gamma \), the system has been assumed to thermalize to a plasma configuration for which

\[ n_{e^+} = n_{e^-} \sim n_\gamma \sim n_{e^+e^-}, \tag{8.5.10} \]

where \( n_{e^+e^-} \) is the total number density of \( e^+e^- \)-pairs created in the dyadosphere [397, 398]. In Fig. 8.3 we show the average energy per pair as a func-

**Figure 8.3:** The average energy per pair is shown here as a function of the black hole mass in solar mass units for \( \xi = 1 \) (solid line), \( \xi = 0.5 \) (dashed line) and \( \xi = 0.1 \) (dashed and dotted line).
8.6 The “dyadotorus”

We turn now to examine how the presence of rotation modifies the geometry of the surface containing the region where electron–positron pairs are created.
The extraction of blackholic energy from a black hole by vacuum polarization processes

as well as the conditions for the existence of such a surface. Due to the axial symmetry of the problem, this region was called the “dyadotorus” [411, 412].

We shall follow the treatment of [411, 412]. As in Damour [383, 413] we introduce at each point of the space-time the orthogonal Carter tetrad (8.4.4-8.4.7).

Figure 8.5: The projections of the dyadotorus on the X-Z plane corresponding to different values of the ratio $E/E_c = \kappa$ are shown (upper panel) for $\mu = 10$ and $\lambda = 1.49 \times 10^4$. The corresponding plot for the dyadosphere with the same mass energy and charge to mass ratio is shown in the lower panel for comparison. Reproduced from [412].

From Eq. (8.4.9) we define the dyadotorus by the condition $|E_{(1)}| = \kappa E_c$, where $10^{-1} \leq \kappa \leq 10$, see Fig. 8.5. Solving for $r$ and introducing the dimensionless quantities $\xi = Q/(\sqrt{GM})$, $\mu = M/M_\odot$, $\bar{\kappa} = ac^2/(GM)$,
The two parts of the surface join at the particular values $\theta^*$ and $\pi - \theta^*$ of the polar angle where

$$\theta^* = \arccos \left( \frac{1}{2\sqrt{2\bar{\alpha}}} \sqrt{\frac{\xi}{\mu \bar{E}}} \right),$$

The requirement that $\cos \theta^* \leq 1$ can be solved for instance for the charge parameter $\xi$, giving a range of values of $\xi$ for which the dyadotorus takes one of the shapes (see fig.8.6)

$$\text{surface} = \begin{cases} 
\text{ellipsoid-like} & \text{if } \xi \geq \xi_* \\
\text{torus-like} & \text{if } \xi < \xi_*
\end{cases} \quad (8.6.2)$$

where $\xi_* = 8\mu \bar{E} \bar{\alpha}^2$.

In Fig. 8.6 we show some examples of the dyadotorus geometry for different sets of parameters for an extreme Kerr–Newman black hole ($a^2 c^5 / G^2 + Q^2 / G = M^2$), we can see the transition from a toroidal geometry to an ellipsoidal one depending on the value of the black hole charge.

Fig. 8.7 shows the projections of the surfaces corresponding to different values of the ratio $|E(1)| / E_c \equiv \kappa$ for the same choice of parameters as in Fig. 8.6 (b), as an example. We see that the region enclosed by such surfaces shrinks for increasing values of $\kappa$.

Equating (8.4.9) and (8.5.2) for $\theta = \pi/2$ and $\bar{\alpha} = 1$ we get

$$\mu = \xi \kappa \times 5 \times 10^5. \quad (8.6.3)$$

8.7 Geometry of gravitationally collapsing cores

In the previous Sections we have focused on the theoretically well defined problem of pair creation in the electric field of an already formed black hole. In this section we shall follow the treatment of Cherubini et al. [405] addressing some specific issues on the dynamical formation of the black hole, recalling first the Oppenheimer-Snyder solution and then considering its generalization to the charged case using the classical work of W. Israel and V. de la Cruz [70,71].
The extraction of blackholic energy from a black hole by vacuum polarization processes

Figure 8.6: The projection of the dyadotorus on the $X - Z$ plane ($X = r \sin \theta$, $Z = r \cos \theta$ are Cartesian-like coordinates built up simply using the Boyer-Lindquist radial and angular coordinates) is shown for an extreme Kerr-Newman black hole with $\mu = 10$ and different values of the charge parameter $\xi = [1, 1.3, 1.49, 1.65] \times 10^{-4}$ (from (a) to (d) respectively). The black circle represents the black hole horizon. Details in [411, 412].

1070
8.7 Geometry of gravitationally collapsing cores

Figure 8.7: The projections of the surfaces corresponding to different values of the ratio $|E_\text{(1)}|/E_c \equiv \kappa$ are shown for the same choice of parameters as in Fig. 8.6 (b), as an example. The gray shaded region is part of the “dyadotorus” corresponding to the case $\kappa = 1$ as plotted in Fig. 8.6 (b). The region delimited by dashed curves corresponds to $\kappa = 0.8$, i.e., to a value of the strength of the electric field smaller than the critical one, and contains the dyadotorus; the latter in turn contains the white region corresponding to $\kappa = 1.4$, i.e., to a value of the strength of the electric field greater than the critical one. Details in [411].

8.7.1 The Tolman-Oppenheimer-Snyder solution

Oppenheimer and Snyder first found a solution of the Einstein equations describing the gravitational collapse of spherically symmetric star of mass greater than $\sim 0.7M_\odot$. In this section we briefly review their pioneering work as presented in Ref. [69].

In a spherically symmetric space-time such coordinates can be found $(t, r, \theta, \phi)$ that the line element takes the form

$$ds^2 = e^\nu dt^2 - e^\lambda dr^2 - r^2 d\Omega^2,$$

$$d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2,$$

where $\nu = \nu(t, r), \lambda = \lambda(t, r)$. However the gravitational collapse problem is better solved in a system of coordinates $(\tau, R, \theta, \phi)$ which are comoving with the matter inside the star. In comoving coordinates the
line element takes the form
\[ ds^2 = d\tau^2 - e^\sigma dR^2 - e^\omega d\Omega^2, \]
\[ \bar{\omega} = \bar{\omega}(\tau, R), \omega = \omega(\tau, R). \]

Einstein equations read
\[ 8\pi T^1_1 = 8\pi T^2_2 = 8\pi T^3_3 = 8\pi T^4_4 = 0, \]
\[ 8\pi e^\sigma T^4_1 = -8\pi T^4_4 = \rho, \]
where \( T_{\mu\nu} \) is the energy–momentum tensor of the stellar matter, a dot denotes a derivative with respect to \( \tau \) and a prime denotes a derivative with respect to \( R \). Oppenheimer and Snyder were only able to integrate Eqs. (8.7.2)–(8.7.5) in the case when the pressure \( p \) of the stellar matter vanishes and no energy is radiated outwards. In the following we thus put \( p = 0 \). Under this hypothesis
\[ T^1_1 = T^2_2 = T^3_3 = T^4_4 = 0, \]
where \( \rho \) is the comoving density of the star. Eq. (8.7.5) was first integrated by Tolman in Ref. [414]. The solution is
\[ e^{\sigma} = e^\omega \omega^2 / 4f^2(R), \]
where \( f = f(R) \) is an arbitrary function. In Ref. [69] the case of \( f(R) = 1 \) was studied. In Section 8.7.2 below the hypothesis \( f(R) = 1 \) is relaxed in the case of a shell of dust. Substitution of Eq. (8.7.6) into Eq. (8.7.2) with \( f(R) = 1 \) gives
\[ \dot{\omega} + \frac{3}{\rho} \omega^2 = 0, \]
which can be integrated to give
\[ e^{\omega} = (F \tau + \tilde{G})^{4/3}, \]
where \( F = F(R) \) and \( \tilde{G} = \tilde{G}(R) \) are arbitrary functions. Substitution of Eq. (8.7.6) into Eq. (8.7.3) gives Eq. (8.7.7) again. From Eqs. (8.7.4), (8.7.6) and (8.7.8) the density \( \rho \) can be found as
\[ 8\pi \rho = \frac{4}{3} \left( \tau + \frac{\tilde{G}}{F} \right)^{-1} \left( \tau + \frac{\tilde{G}}{F} \right)^{-1}. \]
8.7 Geometry of gravitationally collapsing cores

There is still the gauge freedom of choosing $R$ so to have

$$\tilde{G} = R^{3/2}.$$ 

Moreover, arbitrary initial density profile can be chosen, i.e., for the density at the initial time $\tau = 0, \rho_0 = \rho_0(R)$. Eq. (8.7.9) then becomes

$$FF' = 9\pi R^2 \rho_0(R)$$

whose solution contains only one arbitrary integration constant. It is thus seen that the choice of Oppenheimer and Snyder of $f(R) = 1$ allows one to assign only a 1-parameter family of functions for the initial values $\rho_0 = \rho_0(R)$ of $\dot{\rho}$. However in general one should be able to assign the initial values of $\dot{\rho}$ arbitrarily. This will be done in Section (8.7.2) in the case of a shell of dust.

Choosing, for instance,

$$\rho_0 = \begin{cases} \text{const} > 0 & \text{if } R < R_b, \\ 0 & \text{if } R \geq R_b \end{cases},$$

$R_b$ being the comoving radius of the boundary of the star, gives

$$F = \begin{cases} -\frac{3}{2} r_+^{1/2} \left( \frac{R}{R_b} \right)^{3/2} & \text{if } R < R_b, \\ -\frac{3}{2} r_+^{1/2} & \text{if } R \geq R_b \end{cases}$$

where $r_+ = 2M$ is the Schwarzschild radius of the star.

We are finally in the position of performing a coordinate transformation from the comoving coordinates $(\tau, R, \theta, \phi)$ to new coordinates $(t, r, \theta, \phi)$ in which the line elements looks like (8.7.1). The requirement that the line element be the Schwarzschild one outside the star fixes the form of such a coordinate transformation to be

$$r = (F\tau + G)^{2/3},$$

$$t = \begin{cases} \frac{2}{3} r_+^{-1/2} (R_b^{3/2} - r_+^{3/2} y^{3/2}) - 2 r_+ y^{1/2} + r_+ \log \frac{y^{1/2} + 1}{y^{1/2} - 1} & \text{if } R < R_b, \\ \frac{2}{3} r_+^{-1/2} (R^{3/2} - r^{3/2}) - 2 (r r_+)^{1/2} + r_+ \log \frac{r^{1/2} + r_+^{1/2}}{r^{1/2} - r_+^{1/2}} & \text{if } R \geq R_b \end{cases},$$

where

$$y = \frac{1}{2} \left( \left( \frac{R}{R_b} \right)^2 - 1 \right) + \frac{R r}{r_+ R}.$$ 

8.7.2 Gravitational collapse of charged and uncharged shells

It is well known that the role of exact solutions has been fundamental in the development of general relativity. In this section, we present these ex-
act solutions for a charged shell of matter collapsing into a black hole. Such solutions were found in Ref. [405] and are new with respect to the Tolman–Oppenheimer–Snyder class. For simplicity we consider the case of zero angular momentum and spherical symmetry. This problem is relevant on its own account as an addition to the existing family of interesting exact solutions and also represents some progress in understanding the role of the formation of the horizon and of the irreducible mass as will be discussed in Section 8.8.1, see e.g. [406]. It is also essential in improving the treatment of the vacuum polarization processes occurring during the formation of a black hole discussed in [408,415–418] and references therein.

W. Israel and V. de La Cruz [70,71] showed that the problem of a collapsing charged shell can be reduced to a set of ordinary differential equations. We reconsider here the following relativistic system: a spherical shell of electrically charged dust which is moving radially in the Reissner–Nordström background of an already formed nonrotating black hole of mass $M_1$ and charge $Q_1$, with $Q_1 \leq M_1$. The Einstein–Maxwell equations with a charged spherical dust as source are

$$G_{\mu\nu} = 8\pi \left[ T_{\mu\nu}^{(d)} + T_{\mu\nu}^{(em)} \right], \quad \nabla_{\mu} F^{\nu\mu} = 4\pi j^{\nu}, \quad \nabla_{[\mu} F_{\nu\rho]} = 0,$$

(8.7.10)

where

$$T_{\mu\nu}^{(d)} = \epsilon u_{\mu} u_{\nu}, \quad T_{\mu\nu}^{(em)} = \frac{1}{4\pi} \left( F_{\mu}^{\rho} F_{\rho\nu} - \frac{1}{4} g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} \right), \quad j^{\mu} = \sigma u^{\mu}.$$  

(8.7.11)

Here $T_{\mu\nu}^{(d)}$, $T_{\mu\nu}^{(em)}$ and $j^{\mu}$ are respectively the energy-momentum tensor of the dust, the energy-momentum tensor of the electromagnetic field $F_{\mu\nu}$ and the charge 4-current. The mass and charge density in the comoving frame are given by $\epsilon$, $\sigma$ and $u^{a}$ is the 4-velocity of the dust. In spherical–polar coordinates the line element is

$$ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu = -e^{\nu(r,t)} dt^2 + e^{\lambda(r,t)} dr^2 + r^2 d\Omega^2,$$

(8.7.12)

where $d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$.

We describe the shell by using the four-dimensional Dirac distribution $\delta^{(4)}$ normalized as

$$\int \delta^{(4)}(x,x') \sqrt{-g} d^4x = 1$$

(8.7.13)

where $g = \det |g_{\mu\nu}|$. We then have

$$\epsilon(x) = M_0 \int \delta^{(4)}(x,x_0) r^2 d\tau d\Omega,$$

(8.7.14)

$$\sigma(x) = Q_0 \int \delta^{(4)}(x,x_0) r^2 d\tau d\Omega.$$  

(8.7.15)
$M_0$ and $Q_0$ respectively are the rest mass and the charge of the shell and $\tau$ is the proper time along the world surface $S : x_0 = x_0(\tau, \Omega)$ of the shell. $S$ divides the space-time into two regions: an internal one $\mathcal{M}_-$ and an external one $\mathcal{M}_+$. As we will see in the next section for the description of the collapse we can choose either $\mathcal{M}_-$ or $\mathcal{M}_+$. The two descriptions, clearly equivalent, will be relevant for the physical interpretation of the solutions.

Introducing the orthonormal tetrad

$$\omega^{(0)}_\pm = f_\pm^{1/2} dt, \quad \omega^{(1)}_\pm = f_\pm^{-1/2} dr, \quad \omega^{(2)} = r dt, \quad \omega^{(3)} = r \sin \theta d\phi;$$

we obtain the tetrad components of the electric field

$$E = E\omega^{(1)} = \begin{cases} \frac{Q}{f_+} \omega^{(1)}_+ & \text{outside the shell} \\ \frac{Q}{f_-} \omega^{(1)}_- & \text{inside the shell} \end{cases}$$

where $Q = Q_0 + Q_1$ is the total charge of the system. From the $G_{tt}$ Einstein equation we get

$$ds^2 = \begin{cases} -f_+ dt^2 + f_+^{-1} dr^2 + r^2 d\Omega^2 & \text{outside the shell} \\ -f_- dt^2 + f_-^{-1} dr^2 + r^2 d\Omega^2 & \text{inside the shell} \end{cases}$$

where $f_+ = 1 - \frac{2M}{r} + \frac{Q^2}{r^2}, f_- = 1 - \frac{2M_1}{r} + \frac{Q_1^2}{r^2}$ and $t_-$ and $t_+$ are the Schwarzschild-like time coordinates in $\mathcal{M}_-$ and $\mathcal{M}_+$ respectively. Here $M$ is the total mass-energy of the system formed by the shell and the black hole, measured by an observer at rest at infinity.

Indicating by $r_0$ the Schwarzschild-like radial coordinate of the shell and by $t_{0\pm}$ its time coordinate, from the $G_{tr}$ Einstein equation we have

$$\frac{M_0}{r_0} \left[ f_+ (r_0) \frac{dt_{0+}}{dr} + f_- (r_0) \frac{dt_{0-}}{dr} \right] = M - M_1 - \frac{Q_0^2}{2r_0} - \frac{Q_1 Q_0}{r_0}.$$  \hspace{1cm} (8.7.19)

The remaining Einstein equations are identically satisfied. From (8.7.19) and the normalization condition $u_\mu u^\mu = -1$ we find

$$\left( \frac{dr_0}{dr} \right)^2 = \frac{1}{M_0^2} \left( M - M_1 + \frac{M_0^2}{2r_0} - \frac{Q_0^2}{2r_0} - \frac{Q_1 Q_0}{r_0} \right)^2 - f_- (r_0)$$

$$= \frac{1}{M_0^2} \left( M - M_1 + \frac{M_0^2}{2r_0} - \frac{Q_0^2}{2r_0} - \frac{Q_1 Q_0}{r_0} \right)^2 - f_+ (r_0),$$

$$\frac{dt_{0\pm}}{dr} = \frac{1}{M_0 f_{0\pm}(r_0)} \left( M - M_1 + \frac{M_0^2}{2r_0} - \frac{Q_0^2}{2r_0} - \frac{Q_1 Q_0}{r_0} \right).$$  \hspace{1cm} (8.7.20)

We now define, as usual, $r_\pm \equiv M \pm \sqrt{M^2 - Q^2}$: when $Q < M$, $r_\pm$ are real and they correspond to the horizons of the new black hole formed by the gravitational collapse of the shell. We similarly introduce the horizons

1075
8 The extraction of blackholic energy from a black hole by vacuum polarization processes

\[ r_\pm = M_1 \pm \sqrt{M_1^2 - Q_1^2} \] of the already formed black hole. From (8.7.19) we have that the inequality

\[ M - M_1 - \frac{Q_1^2}{r_0} > 0 \] (8.7.22)

holds for \( r_0 > r_+ \) if \( Q < M \) and for \( r_0 > r_1^+ \) if \( Q > M \) since in these cases the left-hand side of (8.7.19) is clearly positive. Eqs. (8.7.20) and (8.7.21) (together with (8.7.18), (8.7.17)) completely describe a 5-parameter \((M, Q, M_1, Q_1, M_0)\) family of solutions of the Einstein-Maxwell equations.

For astrophysical applications [408] the trajectory of the shell \( r_0 = r_0(t_0^+) \) is obtained as a function of the time coordinate \( t_0^+ \) relative to the space-time region \( \mathcal{M}_+ \). In the following we drop the + index from \( t_0^+ \). From (8.7.20) and (8.7.21) we have

\[ \frac{dr_0}{dt_0} = \frac{dx}{d\tau} = \pm \frac{F}{\Omega} \sqrt{\Omega^2 - F}, \] (8.7.23)

where

\[ F = f_+ (r_0) = 1 - \frac{2M}{r_0} + \frac{Q^2}{r_0^3}, \quad \Omega = \Gamma - \frac{M_0^2 + Q^2 - Q_1^2}{2M_0 r_0}, \quad \Gamma = \frac{M - M_1}{M_0}. \] (8.7.24)

Since we are interested in an imploding shell, only the minus sign case in (8.7.23) will be studied. We can give the following physical interpretation of \( \Gamma \). If \( M - M_1 \geq M_0 \), \( \Gamma \) coincides with the Lorentz \( \gamma \) factor of the imploding shell at infinity; from (8.7.23) it satisfies

\[ \Gamma = \frac{1}{\sqrt{1 - \left( \frac{dr_0}{dt_0} \right)^2 \bigg|_{r_0=\infty}}} > 1. \] (8.7.25)

When \( M - M_1 < M_0 \) then there is a turning point \( r_0^* \), defined by \( \left. \frac{dr_0}{dt_0} \right|_{r_0=r_0^*} = 0 \). In this case \( \Gamma \) coincides with the “effective potential” at \( r_0^* \):

\[ \Gamma = \sqrt{f_-(r_0^*)} + M_0^{-1} \left( \frac{M_0^2}{r_0^*} \frac{Q_0^2}{r_0^*} - Q_1 \right) \leq 1. \] (8.7.26)

The solution of the differential equation (8.7.23) is given by:

\[ \int dt_0 = - \int \frac{\Omega}{F \sqrt{\Omega^2 - F}} dr_0. \] (8.7.27)

The functional form of the integral (8.7.27) crucially depends on the degree of the polynomial \( P(r_0) = r_0^2 (\Omega^2 - F) \), which is generically two, but in special cases has lower values. We therefore distinguish the following cases:

1. \( M = M_0 + M_1; Q_1 = M_1; Q = M \): \( P(r_0) \) is equal to 0, we simply have

\[ r_0(t_0) = \text{const}. \] (8.7.28)
2. \( M = M_0 + M_1; M^2 - Q^2 = M_1^2 - Q_1^2; Q \neq M \): \( P(r_0) \) is a constant, we have
\[
t_0 = \text{const} + \frac{1}{2 \sqrt{M^2 - Q^2}} \left[ (r_0 + 2) r_0 + r_0^2 \log \left( \frac{r_0 - r_+}{r_0 - r_-} \right) + r_0^2 \log \left( \frac{r_0 - r_+}{r_0 - r_-} \right) \right].
\] (8.7.29)

3. \( M = M_0 + M_1; M^2 - Q^2 \neq M_1^2 - Q_1^2; P(r_0) \) is a first order polynomial and
\[
t_0 = \text{const} + 2 r_0 \sqrt{\Omega^2 - F} \left[ \frac{M_0 r_0}{3(M^2 - Q^2 - M_1^2 + Q_1^2)} \right.
\+ \frac{(M_1^2 + Q_1^2)^2 - 9MM_0(M_0^2 + Q_0^2) + 12M_0^2 M_0^2 + 2Q_0^2 M_0^2}{3(M^2 - Q^2 - M_1^2 + Q_1^2)} \right.
\- \frac{1}{\sqrt{M^2 - Q^2}} \left[ r_0 \arctanh \left( \frac{r_0 \sqrt{\Omega^2 - F}}{\Omega_+} \right) - r_0 \arctanh \left( \frac{r_0 \sqrt{\Omega^2 - F}}{\Omega_-} \right) \right],
\] (8.7.30)

where \( \Omega_+ \equiv \Omega(r_+) \).

4. \( M \neq M_0 + M_1; P(r_0) \) is a second order polynomial and
\[
t_0 = \text{const} - \frac{1}{2 \sqrt{M^2 - Q^2}} \left\{ \frac{2 \Gamma \sqrt{M^2 - Q^2}}{r_0 \sqrt{\Omega^2 - F}} r_0 \sqrt{\Omega^2 - F} \right.
\+ r_0^2 \log \left[ \frac{r_0 \sqrt{\Omega^2 - F}}{r_0 - r_+} + \frac{r_0^2 (Q^2 - F) + r_0^2 \Omega_+^2 - (\Gamma^2 - 1)(r_0 - r_+)^2}{2(r_0 - r_+)} \right]
\- r_0^2 \log \left[ \frac{r_0 \sqrt{\Omega^2 - F}}{r_0 - r_-} + \frac{r_0^2 (Q^2 - F) + r_0^2 \Omega_-^2 - (\Gamma^2 - 1)(r_0 - r_-)^2}{2(r_0 - r_-)} \right]
\- \frac{2MM_0(2r_0^3 - 3r_0) + M_0^2 + Q_0^2 - Q_1^2)}{M_0(\Gamma^2 - 1)^{3/2}} \sqrt{M^2 - Q^2} \log \left[ \frac{r_0 \sqrt{\Omega^2 - F}}{M_0(\Gamma^2 - 1)} \right]
\+ \frac{2M_0(\Gamma^2 - 1)r_0 - (M_0^2 + Q_0^2 - Q_1^2) \Gamma + 2M_0 M}{2M_0 \sqrt{\Omega^2 - F}} \right\}. \] (8.7.31)

In the case of a shell falling in a flat background \( (M_1 = Q_1 = 0) \) it is of particular interest to study the turning points \( r_0^* \) of the shell trajectory. In this case equation (8.7.20) reduces to
\[
\left( \frac{dr_0}{dt} \right)^2 = \frac{1}{M_0^2} \left( M + \frac{M_0^2}{2r_0^2} - \frac{Q^2}{2r_0} \right)^2 - 1. \] (8.7.32)

Case (2) has no counterpart in this new regime and Eq. (8.7.22) constrains the possible solutions to only the following cases:

1. \( M = M_0; Q = M_0 \). \( r_0 = r_0(0) \) constantly.
2. \( M = M_0; \ Q < M_0 \). There are no turning points, the shell starts at rest at infinity and collapses until a Reissner–Nordström black hole is formed with horizons at \( r_0 = r_\pm = M \pm \sqrt{M^2 - Q^2} \) and the singularity in \( r_0 = 0 \).

3. \( M \neq M_0 \). There is one turning point \( r^*_0 \).

a) \( M < M_0 \), then necessarily is \( Q < M_0 \). Positivity of the right-hand side of (8.7.32) requires \( r_0 \leq r^*_0 \), where \( r^*_0 = \frac{1}{2} \frac{Q^2 - M_0^2}{M - M_0} \) is the unique turning point. Then the shell starts from \( r^*_0 \) and collapses until the singularity at \( r_0 = 0 \) is reached.

b) \( M > M_0 \). The shell has finite radial velocity at infinity.

i. \( Q \leq M_0 \). The dynamics are qualitatively analogous to case (2).

ii. \( Q > M_0 \). Positivity of the right-hand side of (8.7.32) and (8.7.22) requires that \( r_0 \geq r^*_0 \), where \( r^*_0 = \frac{1}{2} \frac{Q^2 - M_0^2}{M - M_0} \). The shell starts from infinity and bounces at \( r_0 = r^*_0 \), reversing its motion.

In this regime the analytic forms of the solutions are given by Eqs. (8.7.30) and (8.7.31), simply setting \( M_1 = Q_1 = 0 \).

Of course, it is of particular interest for the issue of vacuum polarization the time varying electric field \( E_{r_0} = \frac{Q}{r_0^2} \) on the external surface of the shell. In order to study the variability of \( E_{r_0} \) with time it is useful to consider in the tridimensional space of parameters \((r_0, t_0, E_{r_0})\) the parametric curve \( \mathcal{C} : \left( r_0 = \lambda, \ t_0 = t_0(\lambda), \ E_{r_0} = \frac{Q}{\lambda^2} \right) \). In astrophysical applications [408] we are specially interested in the family of solutions such that \( \frac{dr_0}{dt_0} = 0 \) when \( r_0 = \infty \) which implies that \( \Gamma = 1 \). In Fig. 8.8 we plot the collapse curves in the plane \((t_0, r_0)\) for different values of the parameter \( \xi \equiv \frac{Q}{M}, \ 0 < \xi < 1 \). The initial data are chosen so that the integration constant in Eq. (8.7.30) is equal to 0. In all the cases we can follow the details of the approach to the horizon which is reached in an infinite Schwarzschild time coordinate.

In Fig. 8.9 we plot the parametric curves \( \mathcal{C} \) in the space \((r_0, t_0, E_{r_0})\) for different values of \( \tilde{\xi} \). Again we can follow the exact asymptotic behavior of the curves \( \mathcal{C}, E_{r_0} \) reaching the asymptotic value \( \frac{Q}{r_+^2} \). The detailed knowledge of this asymptotic behavior is of relevance for the observational properties of the black hole formation, see e.g. [406], [408].
Figure 8.8: Collapse curves in the plane \((T, R)\) for \(M = 20M_\odot\) and for different values of the parameter \(\xi\). The asymptotic behavior is the clear manifestation of general relativistic effects as the horizon of the black hole is approached. Details in [405].
The extraction of blackholic energy from a black hole by vacuum polarization processes

\[ \xi = 0.1 \]
\[ \xi = 0.8 \]
\[ \xi = 0.9 \]

\begin{align*}
29.97 & \\
29.98 & \\
29.99 & \\
30.00 & \\
30.01 & \\
30.02 & \\
\end{align*}

\begin{align*}
T (s) & \\
5.0 \times 10^7 & \\
1.0 \times 10^8 & \\
1.5 \times 10^8 & \\
2.0 \times 10^8 & \\
2.5 \times 10^8 & \\
3.0 \times 10^8 & \\
\end{align*}

\begin{align*}
R (cm) & \\
1.0 \times 10^{46} & \\
2.0 \times 10^{46} & \\
3.0 \times 10^{46} & \\
4.0 \times 10^{46} & \\
5.0 \times 10^{46} & \\
6.0 \times 10^{46} & \\
7.0 \times 10^{46} & \\
8.0 \times 10^{46} & \\
\end{align*}

\[ \varepsilon_R \text{ (esu/cm}^2) \]

\textbf{Figure 8.9:} Electric field behaviour at the surface of the shell for } M = 20M_\odot \text{ and for different values of the parameter } \xi. \text{ The asymptotic behavior is the clear manifestation of general relativistic effects as the horizon of the black hole is approached. Details in [405].}
8.8 The maximum energy extractable from a black hole

The theoretical analysis of the collapsing shell considered in the previous section allows to reach a deeper understanding of the mass formula of black holes at least in the case of a Reissner–Nordström black hole. This allows as well to give an expression of the irreducible mass of the black hole only in terms of its kinetic energy of the initial rest mass undergoing gravitational collapse and its gravitational energy and kinetic energy at the crossing of the black hole horizon. It also allows to create a scenario for acceleration of the ultrahigh energy cosmic rays with energy typically $10^{21}$ eV from black holes, as opposed to the process of vacuum polarization producing pairs with energies in the MeV region. We shall follow in this Section the treatment by Ruffini and Vitagliano [407].

8.8.1 The formula of the irreducible mass of a black hole

The main objective of this section is to clarify the interpretation of the mass-energy formula [379] for a black hole. For simplicity we study the case of a nonrotating black hole using the results presented in the previous section. As we saw there, the collapse of a nonrotating charged shell can be described by exact analytic solutions of the Einstein–Maxwell equations. Consider to two complementary regions in which the world surface of the shell divides the space-time: $M_-$ and $M_+$. They are static space-times; we denote their time-like Killing vectors by $\xi^-_\mu$ and $\xi^+_\mu$ respectively. $M_+$ is foliated by the family $\{\Sigma^+_t: t_+ = t\}$ of space-like hypersurfaces of constant $t_+$.

The splitting of the space-time into the regions $M_-$ and $M_+$ allows two physically equivalent descriptions of the collapse and the use of one or the other depends on the question one is studying. The use of $M_-$ proves helpful for the identification of the physical constituents of the irreducible mass while $M_+$ is needed to describe the energy extraction process from black hole. The equation of motion for the shell, Eq. 8.7.20, reduces in this case to

$$\left( M_0 \frac{d\tau_0}{\pi r} \right)^2 = \left( M + \frac{M^2_0}{2r_0} - \frac{Q^2}{2r_0} \right)^2 - M^2_0 \quad (8.8.1)$$

in $M_-$ and

$$\left( M_0 \frac{d\tau_0}{\pi r} \right)^2 = \left( M - \frac{M^2_0}{2r_0} - \frac{Q^2}{2r_0} \right)^2 - M^2_0 f_+ \quad (8.8.2)$$

in $M_+$. The constraint (8.7.22) becomes

$$M - \frac{Q^2}{2r_0} > 0. \quad (8.8.3)$$
The extraction of black holic energy from a black hole by vacuum polarization processes

Since $M_-$ is a flat space-time we can interpret $-\frac{M_0^2}{2r_0}$ in (8.8.1) as the gravitational binding energy of the system. $\frac{Q^2}{2r_0}$ is its electromagnetic energy. Then Eqs. (8.8.1), (8.8.2) differ by the gravitational and electromagnetic self-energy terms from the corresponding equations of motion of a test particle.

Introducing the total radial momentum $P^r \equiv M_0 u^r = M_0 \frac{dr_0}{d\tau}$ of the shell, we can express the kinetic energy of the shell as measured by static observers in $M_-$ as $T \equiv -M_0 u^\mu \zeta^\mu_0 - M_0 = \sqrt{(P^r)^2 + M_0^2} - M_0$. Then from Eq. (8.8.1) we have

$$M = -\frac{M_0^2}{2r_0} + \frac{Q^2}{2r_0} + \sqrt{(P^r)^2 + M_0^2} = M_0 + T - \frac{M_0^2}{2r_0} + \frac{Q^2}{2r_0}. \quad (8.8.4)$$

where we choose the positive root solution due to the constraint (8.8.3). Eq. (8.8.4) is the mass formula of the shell, which depends on the time-dependent radial coordinate $r_0$ and kinetic energy $T$. If $M \geq Q$, a black hole is formed and we have

$$M = M_0 + T_+ - \frac{M_0^2}{2r_+} + \frac{Q^2}{2r_+}, \quad (8.8.5)$$

where $T_+ \equiv T(r_+)$ and $r_+ = M + \sqrt{M^2 - Q^2}$ is the radius of the external horizon of that

$$M = M_{ir} + \frac{Q^2}{2r_+}, \quad (8.8.6)$$

so it follows that

$$M_{ir} = M_0 - \frac{M_0^2}{2r_+} + T_+, \quad (8.8.7)$$

namely that $M_{ir}$ is the sum of only three contributions: the rest mass $M_0$, the gravitational potential energy and the kinetic energy of the rest mass evaluated at the horizon. $M_{ir}$ is independent of the electromagnetic energy, a fact noticed by Bekenstein [419]. We have taken one further step here by identifying the independent physical contributions to $M_{ir}$. This has important consequences for the energetics of black hole formation (see [406]).

Next we consider the physical interpretation of the electromagnetic term $\frac{Q^2}{2r_0}$, which can be obtained by evaluating the Killing integral

$$\int_{\Sigma^+} \zeta^\mu_+ T^\text{(em)}_{\mu\nu} d\Sigma^\nu = \int_{r_0}^{\infty} r^2 dr \int_0^1 d\cos \theta \int_0^{2\pi} d\phi \, T^\text{(em)}_{\mu\nu} \, d\Sigma^\nu_0 \equiv \frac{Q^2}{2r_0}, \quad (8.8.8)$$

where $\Sigma^+_+ \equiv t = \text{const}$ is the time-like hypersurface in $M_+$ described by the equation $t_+ = t = \text{const}$, with $d\Sigma^\nu$ as its surface element vector. The quantity in Eq. (8.8.8) differs from the purely electromagnetic energy

$$\int_{\Sigma^+} n^\mu_+ T^\text{(em)}_{\mu\nu} d\Sigma^\nu = \frac{1}{2} \int_{r_0}^{\infty} dr \sqrt{g_{\tau\tau}} Q^2, \quad (8.8.9)$$
where \( n_+^{\mu} = f_+^{-1/2} \xi_+^{\mu} \) is the unit normal to the integration hypersurface and \( g_{rr} = f_+ \). This is similar to the analogous situation for the total energy of a static spherical star of energy density \( \epsilon \) within a radius \( r_0 \),

\[
m_0 (r_0) = 4\pi \int_0^{r_0} dr \, r^2 \epsilon. 
\]

by the gravitational energy (see [393]). Therefore the term \( \frac{Q^2}{2r_0} \) in the mass formula (8.8.4) is the total energy of the electromagnetic field and includes its own gravitational binding energy. This energy is stored throughout the region \( M_+ \), extending from \( r_0 \) to infinity.

### 8.8.2 Extracting electromagnetic energy from a subcritical and overcritical black hole

We now turn to the problem of extracting the electromagnetic energy from a black hole (see [379]). We can distinguish between two conceptually physically different processes, depending on whether the electric field strength \( E = \frac{Q}{r^2} \) is smaller or greater than the critical value \( E_c \). The maximum value \( E_+ = \frac{Q}{r_+^2} \) of the electric field around a black hole is reached at the horizon. In what follows we restore \( G, \hbar \) and \( c \).

For \( E_+ < E_c \) the leading energy extraction mechanism consists of a sequence of discrete elementary decay processes of a particle into two oppositely charged particles. The condition \( E_+ < E_c \) implies

\[
\xi \equiv \frac{Q}{\sqrt{GM}} \lesssim \left\{ \begin{array}{ll}
\frac{GM/c^2}{\lambda_C} \left( \frac{\sqrt{GM}}{\lambda_C} \right)^{-1} \sim 10^{-6} \frac{M}{M_\odot} & \text{if } \frac{M}{M_\odot} \leq 10^6 \\
1 & \text{if } \frac{M}{M_\odot} > 10^6
\end{array} \right. ,
\]

(8.8.10)

where \( \lambda_C \) is the Compton wavelength of the electron. Denardo and Ruffini [376] and Denardo, Hively and Ruffini [377] have defined as the effective ergosphere the region around a black hole where the energy extraction processes occur. This region extends from the horizon \( r_+ \) up to a radius

\[
r_{\text{Erg}} = \frac{GM}{c^2} \left[ 1 + \sqrt{1 - \frac{\xi^2}{1 - \frac{\xi^2}{c^2 Gm_e}}} \right] \sim \frac{\xi}{m_e c}. 
\]

(8.8.11)

The energy extraction occurs in a finite number \( N_{\text{PD}} \) of such discrete elementary processes, each one corresponding to a decrease of the black hole charge. We have

\[
N_{\text{PD}} \simeq \frac{Q}{\epsilon}. 
\]

(8.8.12)
8 The extraction of blackholic energy from a black hole by vacuum polarization processes

Since the total extracted energy is (see Eq. (8.8.6)) $E_{\text{tot}} = \frac{Q^2}{2r_+}$, we obtain for the mean energy per accelerated particle $\langle E \rangle_{\text{PD}} = \frac{E_{\text{tot}}}{N_{\text{PD}}}$

$$\langle E \rangle_{\text{PD}} = \frac{Qe}{2r_+} = \frac{1}{2} \left( \frac{\xi}{\sqrt{1-\xi^2}} \right) \frac{e}{\sqrt{Gm_e}} m_e c^2 \approx \frac{1}{2} \left( \frac{\xi}{\sqrt{Gm_e}} \right) m_e c^2,$$

(8.8.13)

which gives

$$\langle E \rangle_{\text{PD}} \lesssim \begin{cases} \frac{M}{M_\odot} \times 10^{21} \text{eV} & \text{if } \frac{M}{M_\odot} \leq 10^6 \\ 10^{27} \text{eV} & \text{if } \frac{M}{M_\odot} > 10^6 \end{cases}. \quad (8.8.14)$$

One of the crucial aspects of the energy extraction process from a black hole is its back reaction on the irreducible mass expressed in [379]. Although the energy extraction processes can occur in the entire effective ergosphere defined by Eq. (8.8.11), only the limiting processes occurring on the horizon with zero kinetic energy can reach the maximum efficiency while approaching the condition of total reversibility (see Fig. 2 in [379] for details). The farther from the horizon that a decay occurs, the more it increases the irreducible mass and loses efficiency. Only in the complete reversibility limit [379] can the energy extraction process from an extreme black hole reach the upper value of 50% of the total black hole energy.

For $E_+ \geq E_c$ the leading extraction process is the collective process based on the generation of the optically thick electron–positron plasma by the vacuum polarization. The condition $E_+ \geq E_c$ implies

$$\frac{GM/c^2}{\lambda_C} \left( \frac{e}{\sqrt{Gm_e}} \right)^{-1} \approx 2 \cdot 10^{-6} \frac{M}{M_\odot} \leq \xi \leq 1. \quad (8.8.15)$$

This vacuum polarization process can occur only for a black hole with mass smaller than $5 \cdot 10^5 M_\odot$. The electron–positron pairs are now produced in the dyadosphere of the black hole. We have

$$r_{\text{dya}} \ll r_{\text{Erg}}. \quad (8.8.16)$$

The number of particles created [398] is then

$$N_{\text{dya}} = \frac{1}{3} \left( \frac{r_{\text{dya}}}{\lambda_C} \right) \left( 1 - \frac{r_+}{r_{\text{dya}}} \right) \left[ 4 + \frac{r_+}{r_{\text{dya}}} + \left( \frac{r_+}{r_{\text{dya}}} \right)^2 \right] \frac{Q}{c} \approx \frac{4}{3} \left( \frac{r_{\text{dya}}}{\lambda_C} \right) \frac{Q}{c}. \quad (8.8.17)$$

The total energy stored in the dyadosphere is [398]

$$E_{\text{dya}} = \left( 1 - \frac{r_+}{r_{\text{dya}}} \right) \left[ 1 - \left( \frac{r_+}{r_{\text{dya}}} \right)^4 \right] \frac{Q^2}{2r_+} \approx \frac{Q^2}{2r_+}. \quad (8.8.18)$$

The mean energy per particle produced in the dyadosphere $\langle E \rangle_{\text{dya}} = \frac{E_{\text{dya}}}{N_{\text{dya}}}$ is
8.9 A theorem on a possible disagreement between black holes and thermodynamics

Then

\[ \langle E \rangle_{\text{dya}} = \frac{3}{2} \left( \frac{r_+}{r_{\text{dya}}} \right)^4 \left( \frac{\lambda C}{r_{\text{dya}}} \right) \frac{Q e}{r_+} \approx \frac{3}{8} \left( \frac{\lambda C}{r_{\text{dya}}} \right) \frac{Q e}{r_+}, \]  

(8.8.19)

which can be also rewritten as

\[ \langle E \rangle_{\text{dya}} \approx \frac{3}{8} \left( \frac{r_{\text{dya}}}{r_+} \right) m_e c^2 \sim \sqrt{\frac{x}{M/M_\odot}} 10^5 \text{keV}. \]  

(8.8.20)

We stress again that the vacuum polarization around a black hole has been observed to reach theoretically the maximum efficiency limit of 50% of the total mass-energy of an extreme black hole (see e.g. [398]).

Let us now compare and contrast these two processes. We have

\[ r_{\text{Eerg}} \sim \left( \frac{r_{\text{dya}}}{\lambda C} \right) r_{\text{dya}}, \quad N_{\text{dya}} \sim \left( \frac{r_{\text{dya}}}{\lambda C} \right) N_{\text{PD}}, \quad \langle E \rangle_{\text{dya}} \sim \left( \frac{\lambda C}{r_{\text{dya}}} \right) \langle E \rangle_{\text{PD}}. \]  

(8.8.21)

Moreover we see (Eqs. (8.8.14), (8.8.20)) that \( \langle E \rangle_{\text{PD}} \) is in the range of energies of UHECR (see [420] and references therein), while for \( \xi \sim 0.1 \) and \( M \sim 10 M_\odot, \langle E \rangle_{\text{dya}} \) is in the \( \gamma \)-ray range. In other words, the discrete particle decay process involves a small number of particles with ultrahigh energies \( (\sim 10^{21} \text{eV}) \), while vacuum polarization involves a much larger number of particles with lower mean energies \( (\sim 10 \text{MeV}) \).

8.9 A theorem on a possible disagreement between black holes and thermodynamics

This analysis of vacuum polarization process around black holes is so general that it allows as well to look back to traditional results on black hole physics with an alternative point of view. We quote in particular a result which allows to overcome a claimed inconsistency between general relativity and thermodynamics in the field of black holes.

It is well known that if a spherically symmetric mass distribution without any electromagnetic structure undergoes free gravitational collapse, its total mass-energy \( M \) is conserved according to the Birkhoff theorem: the increase in the kinetic energy of implosion is balanced by the increase in the gravitational energy of the system. If one considers the possibility that part of the kinetic energy of implosion is extracted then the situation is very different: configurations of smaller mass-energy and greater density can be attained without violating Birkhoff theorem in view of the radiation process.

From a theoretical physics point of view it is still an open question how far such a sequence can go: using causality nonviolating interactions, can one find a sequence of braking and energy extraction processes by which the
density and the gravitational binding energy can increase indefinitely and the mass-energy of the collapsed object be reduced at will? This question can also be formulated in the mass formula language [379] (see also Ref. [406]): given a collapsing core of nucleons with a given rest mass-energy \( M_0 \), what is the minimum irreducible mass of the black hole which is formed?

Following the previous two sections, consider a spherical shell of rest mass \( M_0 \) collapsing in a flat space-time. In the neutral case the irreducible mass of the final black hole satisfies Eq. 8.8.7. The minimum irreducible mass \( M_{\text{irr}}^{(\min)} \) is obtained when the kinetic energy at the horizon \( T_+ \) is 0, that is when the entire kinetic energy \( T_+ \) has been extracted. We then obtain, form Eq. 8.8.7, the simple result

\[
M_{\text{irr}}^{(\min)} = \frac{M_0}{2}. \tag{8.9.1}
\]

We conclude that in the gravitational collapse of a spherical shell of rest mass \( M_0 \) at rest at infinity (initial energy \( M_i = M_0 \)), an energy up to 50\% of \( M_0c^2 \) can in principle be extracted, by braking processes of the kinetic energy. In this limiting case the shell crosses the horizon with \( T_+ = 0 \). The limit \( \frac{M_0}{2} \) in the extractable kinetic energy can further increase if the collapsing shell is endowed with kinetic energy at infinity, since all that kinetic energy is in principle extractable.

We have represented in Fig. 8.10 the world lines of spherical shells of the same rest mass \( M_0 \), starting their gravitational collapse at rest at selected radii \( r^*_0 \). These initial conditions can be implemented by performing suitable braking of the collapsing shell and concurrent kinetic energy extraction processes at progressively smaller radii (see also Fig. 8.11). The reason for the existence of the minimum (8.9.1) in the black hole mass is the “self-closure” occurring by the formation of a horizon in the initial configuration (thick line in Fig. 8.10).

Is the limit \( M_{\text{irr}} \to \frac{M_0}{2} \) actually attainable without violating causality? Let us consider a collapsing shell with charge \( Q \). If \( M \geq Q \) a black hole is formed. As pointed out in the previous section the irreducible mass of the final black hole does not depend on the charge \( Q \). Therefore Eqs. (8.8.7) and (8.9.1) still hold in the charged case. In Fig. 8.11 we consider the special case in which the shell is initially at rest at infinity, i.e. has initial energy \( M_i = M_0 \), for three different values of the charge \( Q \). We plot the initial energy \( M_i \), the energy of the system when all the kinetic energy of implosion has been extracted as well as the sum of the rest mass energy and the gravitational binding energy \(-\frac{M_0^2}{2r_0}\) of the system (here \( r_0 \) is the radius of the shell). In the extreme case \( Q = M_0 \), the shell is in equilibrium at all radii (see Ref. [405]) and the kinetic energy is identically zero. In all three cases, the sum of the extractable kinetic energy \( T \) and the electromagnetic energy \( \frac{Q^2}{2r_0} \) reaches 50\% of the rest mass energy at the horizon, according to Eq. (8.9.1).

What is the role of the electromagnetic field here? If we consider the case
8.9 A theorem on a possible disagreement between black holes and thermodynamics

![Figure 8.10: Collapse curves for neutral shells with rest mass $M_0$ starting at rest at selected radii $R^*$ computed by using the exact solutions given in Ref. [405]. A different value of $M_{\text{irr}}$ (and therefore of $r_+$) corresponds to each curve. The time parameter is the Schwarzschild time coordinate $t$ and the asymptotic behaviour at the respective horizons is evident. The limiting configuration $M_{\text{irr}} = \frac{M_0}{2}$ (solid line) corresponds to the case in which the shell is trapped, at the very beginning of its motion, by the formation of the horizon.](image-url)
Figure 8.11: Energetics of a shell such that $M_i = M_0$, for selected values of the charge. In the first diagram $Q = 0$; the dashed line represents the total energy for a gravitational collapse without any braking process as a function of the radius $R$ of the shell; the solid, stepwise line represents a collapse with suitable braking of the kinetic energy of implosion at selected radii; the dotted line represents the rest mass energy plus the gravitational binding energy. In the second and third diagram $Q/M_0 = 0.7$, $Q/M_0 = 1$ respectively; the dashed and the dotted lines have the same meaning as above; the solid lines represent the total energy minus the kinetic energy. The region between the solid line and the dotted line corresponds to the stored electromagnetic energy. The region between the dashed line and the solid line corresponds to the kinetic energy of collapse. In all the cases the sum of the kinetic energy and the electromagnetic energy at the horizon is 50% of $M_0$. Both the electromagnetic and the kinetic energy are extractable. It is most remarkable that the same underlying process occurs in the three cases: the role of the electromagnetic interaction is twofold: a) to reduce the kinetic energy of implosion by the Coulomb repulsion of the shell; b) to store such an energy in the region around the black hole. The stored electromagnetic energy is extractable as shown in Ref. [406]
of a charged shell with \( Q \approx M_0 \), the electromagnetic repulsion implements the braking process and the extractable energy is entirely stored in the electromagnetic field surrounding the black hole (see Ref. [406]). We emphasize here that the extraction of 50% of the mass-energy of a black hole is not specifically linked to the electromagnetic field but depends on three factors: a) the increase of the gravitational energy during the collapse, b) the formation of a horizon, c) the reduction of the kinetic energy of implosion. Such conditions are naturally met during the formation of an extreme black hole with \( Q = M \), but as we have seen, they are more general and can indeed occur in a variety of different situations, e.g. during the formation of a Schwarzschild black hole by a suitable extraction of the kinetic energy of implosion (see Fig. 8.10 and Fig. 8.11).

Before closing let us consider a test particle of mass \( m \) in the gravitational field of an already formed Schwarzschild black hole of mass \( M \) and go through such a sequence of braking and energy extraction processes. Kaplan [421] found for the energy \( E \) of the particle as a function of the radius \( r \)

\[
E = m \sqrt{1 - \frac{2M}{r}}. \tag{8.9.2}
\]

It would appear from this formula that the entire energy of a particle could be extracted in the limit \( r \rightarrow 2M \). Such 100% efficiency of energy extraction has often been quoted as evidence for incompatibility between General Relativity and the second principle of Thermodynamics (see Ref. [422] and references therein). J. Bekenstein and S. Hawking have gone as far as to consider General Relativity not to be a complete theory and to conclude that in order to avoid inconsistencies with thermodynamics, the theory should be implemented through a quantum description [422–425]. Einstein himself often expressed the opposite point of view (see, e.g., Ref. [426] and references therein).

The analytic treatment presented in Section 8.7.2 can clarify this fundamental issue. It allows to express the energy increase \( \mathcal{E} \) of a black hole of mass \( M_1 \) through the accretion of a shell of mass \( M_0 \) starting its motion at rest at a radius \( r_0 \) in the following formula which generalizes Eq. (8.9.2):

\[
\mathcal{E} \equiv M - M_1 = -\frac{M_0^2}{2r_0} + M_0 \sqrt{1 - \frac{2M_1}{r_0}}, \tag{8.9.3}
\]

where \( M = M_1 + \mathcal{E} \) is clearly the mass-energy of the final black hole. This formula differs from the Kaplan formula (8.9.2) in three respects: (a) it takes into account the increase of the horizon area due to the accretion of the shell; (b) it shows the role of the gravitational self-energy of the imploping shell; (c) it expresses the combined effects of (a) and (b) in an exact closed formula.

The minimum value \( \mathcal{E}_{\text{min}} \) of \( \mathcal{E} \) is attained for the minimum value of the radius \( r_0 = 2M \): the horizon of the final black hole. This corresponds to the
maximum efficiency of the energy extraction. We have
\[ \varepsilon_{\text{min}} = -\frac{M_0^2}{4M} + M_0 \sqrt{1 - \frac{M_1}{M}} = -\frac{M_0^2}{4(M_1 + \varepsilon_{\text{min}})} + M_0 \sqrt{1 - \frac{M_1}{M_1 + \varepsilon_{\text{min}}}}, \]  
(8.9.4)
or solving the quadratic equation and choosing the positive solution for physical reasons
\[ \varepsilon_{\text{min}} = \frac{1}{2} \left( \sqrt{M_1^2 + M_0^2 - M_1} \right). \]
(8.9.5)
The corresponding efficiency of energy extraction is
\[ \eta_{\text{max}} = \frac{M_0 - \varepsilon_{\text{min}}}{M_0} = 1 - \frac{1}{2} \frac{M_1}{M_0} \left( \sqrt{1 + \frac{M_0^2}{M_1^2}} - 1 \right), \]
(8.9.6)
which is strictly smaller than 100% for any given \( M_0 \neq 0 \). It is interesting that this analytic formula, in the limit \( M_1 \ll M_0 \), properly reproduces the result of equation (8.9.1), corresponding to an efficiency of 50%. In the opposite limit \( M_1 \gg M_0 \) we have
\[ \eta_{\text{max}} \simeq 1 - \frac{1}{4} \frac{M_0}{M_1}. \]
(8.9.7)
Only for \( M_0 \rightarrow 0 \), Eq. (8.9.6) corresponds to an efficiency of 100% and correctly represents the limiting reversible transformations. It seems that the difficulties of reconciling General Relativity and Thermodynamics are ascribable not to an incompleteness of General Relativity but to the use of the Kaplan formula in a regime in which it is not valid.

### 8.10 Astrophysical gravitational collapse and black holes

The time evolution of the gravitational collapse (occurring on characteristic gravitational timescale \( \tau = GM/c^3 \approx 5 \times 10^{-5} M/M_\odot \) s) and the associated electrodynamical process are too complex for direct description. We addressed here a more confined problem: the vacuum polarization process around an already formed Kerr–Newman black hole. This is a well defined problem which deserves attention. It is theoretically expected to represent a physical state asymptotically reached in the process of gravitational collapse. Such an asymptotic configuration will be reached when all multipoles departing from the Kerr–Newman geometry have been radiated away either by process of vacuum polarization or electromagnetic and gravitational waves. What is most important is that by performing this theoretical analysis we can have a direct evaluation of the energetics and of the spectra and dynamics of the \( e^+e^- \) plasma created on the extremely short timescale due to the quantum phenomena of \( \Delta t = \hbar/(m_e c^2) \approx 10^{-21} \) s. This entire transient phenomena, starting from an initial neutral condition of the core in the progenitor star,
undergoes the formation of the Kerr–Newman black hole by the collective effects of gravitation, strong, weak, electromagnetic interactions during a fraction of the above mentioned gravitational characteristic timescale of collapse.

After the process of vacuum polarization all the electromagnetic energy of incipient Kerr–Newman black hole will be radiated away and almost neutral Kerr solution will be left and reached asymptotically in time. In a realistic gravitational collapse the theoretical picture described above will be further amplified by the presence of high-energy processes including neutrino emission and gravitational waves emission with their electromagnetic coupling: the gravitationally induced electromagnetic radiation and electromagnetically induced gravitational radiation [427, 428].

Similarly, in the next Section we proceed to a deeper understanding of other collective plasma phenomena also studied in idealized theoretically well defined cases. They will play an essential role in the astrophysical description of the dynamical phase of gravitational collapse.
9 Plasma oscillations in electric fields

We have seen in the previous Sections the application of the Sauter-Heisenberg-Euler-Schwinger process for electron–positron pair production in the heavy nuclei, in the laser and in the last Section in the field of black holes. The case of black holes is drastically different from all the previous ones. The number of electron–positron pairs created is of the order of $10^{60}$, the plasma expected is optically thick and is very different from the nuclear collisions and laser case where pairs are very few and therefore optically thin. The following dynamical aspects need to be addressed.

1. the back reaction of pair production on the external electric field;
2. the screening effect of pairs on the external electric field strengths;
3. the motion of pairs and their interactions.

When these dynamical effects are considered, the pair production in an external electric field is no longer only a process of quantum tunneling in a constant static electric field. In fact, it turns out to be a much more complex process during which all the three above mentioned effects play an important role. More precisely, a phenomenon of electron–positron oscillation, plasma oscillation, takes place. We are going to discuss such plasma oscillation phenomenon in this Section. As we will see in this Section these phenomena can become also relevant for heavy-ion collisions. After giving the basic equations for description of plasma oscillations we give first some applications in the field of heavy ions. In this Section in all formulas we use $c = \hbar = 1$.

9.1 Semiclassical theory of plasma oscillations in electric fields

In the semi-classical QED [429, 430], one quantizes only the Dirac field $\psi(x)$, while an external electromagnetic field $A^\mu(x)$ is treated classically as a mean field. This is the self-consistent mean field or Gaussian approximation that can be formally derived as the leading term in the large-$N$ limit of QED, where $N$ is the number of charged matter fields [430–434]. The motion of
these electrons can be described by a Dirac equation in an external classical electromagnetic potential \( A^\mu(x) \)

\[
[\gamma_\mu (i\partial^\mu - eA^\mu) - m]\psi(x) = 0 \tag{9.1.1}
\]

and the semi-classical Maxwell equation

\[
\partial_\mu F^{\mu\nu} = \langle j^\nu(x) \rangle, \quad j^\nu(x) = i\frac{e}{2}[\bar{\psi}(x), \gamma^\nu \psi(x)], \tag{9.1.2}
\]

where \( j^\nu(x) \) is the electron and positron current and the expectation value is with respect to the quantum states of the electron field. The dynamics that these equations describe is not only the motion of electron and positron pairs, but also their back reaction on the external electromagnetic field. The resultant phenomenon is the so-called plasma oscillation that we will discuss based on both a simplified model of semi-classical scalar QED and kinetic Boltzmann-Vlasov equation as presented in Refs. [429, 430, 435–437].

A scheme for solving the back reaction problem in scalar QED was offered in Refs. [429, 430]. Based on this scheme, a numerical analysis was made in (1+1)-dimensional case [435]. Eqs. (9.1.1), (9.1.2) are replaced by the scalar QED coupled equations for a charged scalar field \( \Phi(x) \)

\[
[(i\partial^\mu - eA^\mu)^2 - m^2]\Phi(x) = 0. \tag{9.1.3}
\]

The current \( j^\nu(x) \) of the charged scalar field in the semi-classical Maxwell equations (9.1.2) is

\[
 j^\nu(x) = i\frac{e}{2} [\Phi^*(x)\partial^\nu\Phi(x) - \Phi(x)\partial^\nu\Phi^*(x)]. \tag{9.1.4}
\]

Now, consider a spatially homogeneous electric field \( E = E_z(t)\hat{z} \) in the \( \hat{z} \)-direction. A corresponding gauge potential is \( A = A_z(t)\hat{z}, A_0 = 0 \). Defining \( E \equiv E_z, A \equiv A_z \) and \( j \equiv j_z \), the Maxwell equations (9.1.2) reduce to the single equation

\[
\frac{d^2 A}{dt^2} = \langle j(x) \rangle, \tag{9.1.5}
\]

for the potential and \( E = -dA/dt \).

The quantized scalar field \( \Phi(x) \) in Eq. (9.1.3) can be expanded in terms of plane waves with operator-valued amplitudes \( f_k(t)a_k \) and \( f_{-k}^*(t)b_k^\dagger \)

\[
\Phi(x) = \frac{1}{V^{1/2}} \sum_k [f_k(t)a_k + f_{-k}^*(t)b_k^\dagger]e^{-ikx}, \tag{9.1.6}
\]

where \( V \) is the volume of the system and the time-independent creation and
annihilation operators obey the commutation relations

\[ [a_k, a^{\dagger}_{k'}] = [b_k, b^{\dagger}_{k'}] = \delta_{k,k'}, \quad (9.1.7) \]

and each \( k \)-mode function \( f_k \) obeys the Wronskian condition,

\[ f_k \dot{f}^*_k - \dot{f}_k f^*_k = i. \quad (9.1.8) \]

The time dependency in this basis \((a_k, b^{\dagger}_k)\) \((9.1.6,9.1.7)\) is carried by the complex mode functions \( f_k(t) \) that satisfy the following equation of motion, as demanded from the QED coupled Eq. \( (9.1.3) \) of Klein–Gordon type

\[ \left( \frac{d^2}{dt^2} + \omega^2_k(t) \right) f_k(t) = 0, \quad (9.1.9) \]

where the time-dependent frequency \( \omega^2_k(t) \) is given by

\[ \omega^2_k(t) \equiv \left[ k - eA(t) \right]^2 + m_e^2 = \left[ k - eA(t) \right]^2 + k_\perp^2 + m_e^2. \quad (9.1.10) \]

Here \( k \) is the constant canonical momentum in the \( \hat{z} \)-direction which should be distinguished from the gauge invariant, but time-dependent kinetic momentum

\[ p(t) = k - eA(t), \quad \frac{dp}{dt} = eE, \quad (9.1.11) \]

which reflects the acceleration of the charged particles due to the electric field, while in the directions transverse to the electric field the kinetic and canonical momenta are the same \( k_\perp = p_\perp \).

The mean value of electromagnetic current \((9.1.4)\) in the \( \hat{z} \)-direction is then

\[ \langle j(t) \rangle = 2e \int \frac{d^3k}{(2\pi)^3} |k - eA(t)||f_k(t)|^2 \left[ 1 + N_+(k) + N_-(\mathbf{-k}) \right], \quad (9.1.12) \]

where \( N_+(k) = \langle a^\dagger_k a_k \rangle \) and \( N_-(\mathbf{k}) = \langle b^\dagger_k b_k \rangle \) are the mean numbers of particles and antiparticles in the time-independent basis \((9.1.6,9.1.7)\). The mean charge density must vanish

\[ \langle j^0(t) \rangle = e \int \frac{d^3k}{(2\pi)^3} \left[ N_+(k) - N_-(\mathbf{-k}) \right] = 0, \quad \int \frac{d^3k}{(2\pi)^3} \equiv \frac{1}{V} \sum_k \]

by the Gauss law for a spatially homogeneous electric field \( (i.e., \nabla \cdot E = 0) \). As a result, \( N_+(k) = N_-(-\mathbf{k}) \equiv N_k \). For the vacuum state, \( N_k = 0 \). The Maxwell equation \((9.1.5)\) for the evolution of electric field becomes

\[ \frac{d^2A}{dt^2} = 2e \int \frac{d^3k}{(2\pi)^3} |k - eA(t)||f_k(t)|^2 \sigma_k, \quad \sigma_k = (1 + 2N_k). \quad (9.1.13) \]
These two scalar QED coupled Eqs. (9.1.9) and (9.1.13) in (1+1)-dimensional case were numerically integrated in Ref. [435]. The results are shown in Fig. 9.1, where the time evolutions of the scaled electric field $\tilde{E} \equiv E/E_c$ and current $\tilde{j} \equiv j\hbar/(E_cm_e^2)$ are shown as functions of time $\tau \equiv (m_e^2/c^2)t$ in unit of Compton time $(\hbar/m_e^2c^2)$. Starting with a strong electric field, one clearly finds the phenomenon of oscillating electric field $E(t)$ and current $j(t)$, i.e., plasma oscillation.

![Figure 9.1](image_url)

**Figure 9.1:** Time evolution of scaled electric field $\tilde{E}$ and current $\tilde{j}$, with initial value $\tilde{E} = 1.0$ and coupling $e^2/m_e^2 = 0.1$. The solid line is semi classical scalar QED, and the dashed line is the Boltzmann-Vlasov model. This figure is reproduced from Fig. 1 (a) in Ref. [435].

This phenomenon of plasma oscillation is shown in Fig. 9.1 and is easy to understand as follows. In a classical kinetic picture, we have the electric current $j = 2en\langle v \rangle$ where $n$ is the density of electrons (or positrons) and $\langle v \rangle$ is their mean velocity. Driven by the external electric field, the velocity $\langle v \rangle$ of electrons (or positrons) continuously increases, until the electric field of electron and positron pairs screens the external electric field down to zero, and the kinetic energy of electrons (or positrons) reaches its maximum. The electric current $j$ saturates as the velocity $\langle v \rangle$ is close to the speed of light.
Afterward, these electrons and positrons continuously move apart from each other further, their electric field, whose direction is opposite to the direction of the external electric field, increases and decelerates electrons and positrons themselves. Thus the velocity $\langle v \rangle$ of electrons and positrons decreases, until the electric field reaches negative maximum and the velocity vanishes. Then the velocity $\langle v \rangle$ of electrons and positrons starts to increase in backward direction and the electric field starts to decrease for another oscillation cycle.

### 9.2 Kinetic theory of plasma oscillations in electric fields

In describing the same system as in the previous section it can be used, alternatively to the semi-classical theory, a phenomenological model based on the following relativistic Boltzmann-Vlasov equation [435]

$$\frac{d\mathcal{F}}{dt} = \partial_t \mathcal{F} + eE \frac{\partial \mathcal{F}}{\partial p} = \frac{dN}{dtdVd^3p}, \quad (9.2.1)$$

where the $x$-independent function $\mathcal{F}(p, t)$ is a classical distribution function of particle and antiparticle pairs in phase space. The source term in the right-hand side of Eq. (9.2.1) is related to the Schwinger rate $\mathcal{P}_{\text{boson}}$ (5.7.26) for the pair production of spin-0 particle and antiparticle

$$\frac{dN}{dtdVd^3p} = \left[1 + 2\mathcal{F}(p, t)\right]\mathcal{P}_{\text{boson}}\delta^3(p), \quad (9.2.2)$$

where the $\delta$-function $\delta^3(p)$ expresses the fact that particles are produced at rest and the factor $\left[1 + 2\mathcal{F}(p, t)\right]$ accounts for stimulated pair production (Bose enhancement). The Boltzmann-Vlasov equation (9.2.1) for the distribution function $\mathcal{F}(p, t)$ is in fact attributed to the conservation of particle number in phase space.

In the field equation (9.1.5) for the classical gauge potential $A$, the electric current $\langle j(x) \rangle$ is contributed from the conduction current

$$j_{\text{cond}} = 2e \int \frac{d^3p}{(2\pi)^3} \frac{[p - eA(t)]}{\omega_p} \mathcal{F}(p, t), \quad (9.2.3)$$

and the polarization current [438]

$$j_{\text{pol}} = \frac{2}{E} \int \frac{d^3p}{(2\pi)^3} \omega_p \frac{dN}{dtdVd^3p}. \quad (9.2.4)$$

The relativistic Boltzmann-Vlasov equation (9.2.1) and field equation (9.1.5)
with the conduction current (9.2.3) and the polarization current (9.2.4) were
numerically integrated [435] in (1+1)-dimensional case. The numerical in-
tegration shows that the system undergoes plasma oscillations. In Fig. 9.1
the results of the semi-classical analysis and the numerical integration of the
Boltzmann equation are compared. We see that they are in good quantita-
tive agreement. The discrepancies are because in addition to spontaneous
pair production, the quantum theory takes into account pair production via
bremsstrahlung ("induced" pair production), which are neglected in Eq. (9.2.1)).

In Refs. [439, 440], the study of plasma oscillation was extended to the
fermionic case. On the basis of semi-classical theory of spinor QED, express-
ing the solution of the Dirac equation (9.1.1) as
\[
\psi(x) = \gamma_\mu (i\partial^\mu - eA^\mu) + m_e \phi(x),
\]
where \( \phi(x) \) is a four-component spinor, one finds that \( \phi(x) \) satisfies the quadratic
Dirac equation,
\[
\left[(i\partial^\mu - eA^\mu)^2 - \frac{e^2}{2}\sigma^{\mu\nu}F_{\mu\nu} - m_e^2\right] \phi(x) = 0. \tag{9.2.5}
\]
The electric current of spinor field \( \phi(x) \) couples to the external electric field
that obeys the field equation (9.1.5).

The source term in the right-hand side of the kinetic Boltzmann-Vlasov
equation (9.2.1) is changed to the Schwinger rate \( P_{\text{fermion}} \) (5.7.25) for the pair
production of electrons and positrons,
\[
d\!N/\!dtdVd^3p = [1 - 2F(p,t)]P_{\text{fermion}}\delta^3(p), \tag{9.2.6}
\]
where the Pauli blocking is taken into account by the factor \( [1 - 2F(p,t)] \).
Analogously to the scalar QED case, both semi-classical theory of spinor QED
and kinetic Boltzmann-Vlasov equation have been analyzed and numerical
integration was made in the (1+1)-dimensional case [439,440]. The numerical
results show that plasma oscillations of electric field, electron and positron
currents are similar to that plotted in Fig. 9.1.

### 9.3 Plasma oscillations in the color electric field of
heavy ions

The Relativistic Heavy-Ion Collider (RIHC) at Brookhaven National Labora-
tory and Large Hadron Collider (LHC) at CERN are designed with the goal of
producing a phase of deconfined hadronic matter: the quark-gluon plasma.
A popular theoretical model for studying high-energy heavy-ion collisions
begins with the creation of a flux tube containing a strong color electric field
Plasma oscillations in the color electric field of heavy ions

The field energy is converted into particles such as quark and antiquark pairs and gluons that are created by the Sauter–Euler–Heisenberg–Schwinger quantum tunneling mechanism. A relativistic Boltzmann–Vlasov equation coupling to such particle creation source is phenomenologically adopted in a kinetic theory model for the hydrodynamics of quark and gluon plasma.

In the collision of heavy-ion beams, one is clearly dealing with a situation that is not spatially homogeneous. However, particle production in the central rapidity region can be modeled as hydrodynamical system with longitudinal boost invariance. To express the longitudinal boost invariance of the hydrodynamical system, one introduces the comoving coordinates: fluid proper time $\tau$ and rapidity $\eta$ by the relationships:

$$z = \tau \sinh(\eta), \quad t = \tau \cosh(\eta),$$

in terms of the Minkowski time $t$ and the coordinate $z$ along the beam direction $\hat{z}$ in the ordinary laboratory frame. The line element and metric tensor in these coordinates are given by:

$$ds^2 = d\tau^2 - dx^2 - dy^2 - \tau^2 d\eta^2, \quad g_{\mu\nu} = \text{diag}(1, -1, -1, -\tau^2),$$

and

$$g_{\mu\nu} = V^a_{\mu} V^b_{\nu} \eta_{ab}, \quad V^a_{\mu} = \text{diag}(1, 1, 1, \tau).$$

where the vierbein $V^a_{\mu}$ transforms the curvilinear coordinates to Minkowski coordinates and $\det(V) = \sqrt{-g} = \tau$. The covariant derivative on fermion field $\psi(x)$ is given by:

$$\nabla_\mu \psi(x) \equiv [(i\partial_\mu - eA_\mu) + \Gamma_\mu] \psi(x)$$

and the spin connection $\Gamma_\mu$ is

$$\Gamma_\mu = \frac{1}{2} \Sigma^{ab} V_{\alpha\beta}(\partial_\mu V^\alpha_{\beta} + \Gamma^\alpha_{\mu\lambda} V^\lambda_{\beta}), \quad \Sigma^{ab} = \frac{1}{4}[\gamma^a, \gamma^b],$$

with $\Gamma^\nu_{\mu\lambda}$ the usual Christoffel symbols and $\gamma^a$ the usual coordinate-independent Dirac gamma matrices. The coordinate-dependent gamma matrices $\tilde{\gamma}^a_{\mu}$ are obtained via

$$\tilde{\gamma}^a_{\mu} = \gamma^a V^a_{\mu}.$$

In the curved space time (9.3.2), the Dirac equation (9.1.1) and semi-classical Maxwell equation (9.1.2) are modified as

$$[i\tilde{\gamma}_\mu \nabla^\mu - m_e] \psi(x) = 0$$
and
\[
\frac{1}{\sqrt{-g}} \partial_\mu \sqrt{-g} F^{\mu \nu} = \langle j^\nu (x) \rangle, \quad j^\nu (x) = \frac{e}{2} [\bar{\psi} (x), \gamma^\nu \psi (x)]. \tag{9.3.5}
\]

The phenomenological Boltzmann–Vlasov equation in (3+1)-dimensions can be also written covariantly as
\[
\frac{DF}{D\tau} \equiv p^\mu \frac{\partial F}{\partial q^\mu} - e p^\mu F_{\mu \nu} \frac{\partial F}{\partial p^\nu} = \frac{dN}{\sqrt{-g} d^3 q d^3 p}, \tag{9.3.6}
\]
where \( D / D\tau \) is the total proper time derivative. This kinetic transport equation is written in the comoving coordinates and their conjugate momenta:
\[
q^\mu = (\tau, x, y, \eta), \quad p^\mu = (p_\tau, p_x, p_y, p_\eta).
\]

Due to the longitudinal boost invariance, energy density and color electric field are spatially homogeneous, i.e., they are functions of the proper time \( \tau \) only \[453\]. Consequently, the approach for spatially homogeneous electric field presented in Refs. \[430, 435, 439, 440\] and discussed in the previous Section 9.2 is applicable to the phenomenon of plasma oscillations in ultrarelativistic heavy-ion collisions \[453\] using Eq. 9.3.4 (resp. Eq. 9.3.6) in the place of Eq. 9.1.1 (resp. Eq. 9.2.1).

### 9.4 Quantum Vlasov equation

To understand the connection between the two frameworks of semi-classical field theory and classical kinetic theory, both of which describe the plasma oscillations, one can try to study a quantum transport equation in the semi-classical theory \[437, 455–457\]. For this purpose, a Bogoliubov transformation from the time-independent number basis \((a_k, b_k^\dagger)\) (9.1.6, 9.1.7) to a time-dependent number basis \([\tilde{a}_k(t), \tilde{b}_k^\dagger(t)]\) is introduced \[437, 455, 456\],
\[
f_k (t) = a_k (t) \tilde{f}_k (t) + \beta_k (t) \tilde{f}_k^*(t), \quad \tilde{f}_k^* (t) = -i \omega_k a_k (t) \tilde{f}_k (t) + i \omega_k \beta_k (t) \tilde{f}_k^* (t), \tag{9.4.1}
\]
and
\[
a_k (t) = a_k^+ (t) \tilde{a}_k (t) - \beta_k^* (t) \tilde{b}_{-k}^+ (t), \quad \tilde{b}_{-k}^+ (t) = a_k (t) \tilde{b}_{-k}^+ (t) - \beta_k (t) \tilde{a}_k (t), \tag{9.4.3}
\]
where \( a_k (t) \) and \( \beta_k (t) \) are the Bogoliubov coefficients. They obey
\[
|a_k (t)|^2 - |\beta_k (t)|^2 = 1,
\]
for each mode $k$. In the limit of very slowly varying $\omega_k(t)$ as a function of time $t$ (9.1.10), i.e., $\dot{\omega}_k \ll \omega_k^2$ and $\ddot{\omega}_k \ll \omega_k^3$, the adiabatic number basis $[\tilde{a}_k(t), \tilde{b}_k^\dagger(t)]$ is defined by first constructing the adiabatic mode functions,

$$\tilde{f}_k(t) = \left(\frac{\hbar}{2\omega_k}\right)^{1/2} \exp[-i\Theta_k(t)], \quad \Theta_k(t) = \int^t \omega_k(t') dt'. \quad (9.4.5)$$

The particle number $N_k(t)$ in the time-dependent adiabatic number basis is given by

$$N_k(t) = \langle \tilde{a}_k^\dagger(t) \tilde{a}_k(t) \rangle = \langle \tilde{b}_k^\dagger(t) \tilde{b}_k(t) \rangle = |\alpha_k(t)|^2 N_k + |\beta_k(t)|^2 [1 + N_k], \quad (9.4.6)$$

which though time-dependent, is an adiabatic invariant of the motion. Consequently, it is a natural candidate for a particle number density distribution function $\mathcal{F}(\mathbf{p}, t)$ in the phase space, that is needed in a kinetic description.

By differentiating $N_k(t)$ (9.4.6) and using the basic relationships (9.1.10,9.4.2,9.4.5), one obtains,

$$\dot{N}_k(t) = \frac{\dot{\omega}_k}{\omega_k} \text{Re} \{ \tilde{c}_k(t) \exp[-2i\Theta_k(t)] \}, \quad \tilde{c}_k(t) = (1 + 2N_k)\alpha_k(t)\beta_k^*(t), \quad (9.4.7)$$

and

$$\dot{\tilde{c}}_k(t) = \frac{\dot{\omega}_k}{2\omega_k} [1 + 2N_k(t)] \exp[2i\Theta_k(t)]. \quad (9.4.8)$$

These two equations give rise to the quantum Vlasov equations [437,455,456],

$$\dot{N}_k(t) = S_k(t), \quad (9.4.9)$$

$$S_k(t) = \frac{\dot{\omega}_k}{2\omega_k} \int_{-\infty}^{t} dt' \frac{\dot{\omega}_k(t')}{\omega_k(t')} [1 \pm 2N_k(t')] \cos[2\Theta_k(t) - 2\Theta_k(t')] \quad (9.4.10)$$

describing the time evolution of the adiabatic particle number $N_k(t)$ of the mean field theory. $S_k(t)$ describes the quantum creation rate of particle number in an arbitrary slowly varying mean field. The Bose enhancement and Pauli blocking factors $[1 \pm 2N_k(t')]$ appear in Eq. (9.4.10) so that both spontaneous and induced particle creation are included automatically in the quantum treatment. The most important feature of Eq. (9.4.10) is that the source term $S_k(t)$ is nonlocal in time, indicating the particle creation rate depending on the entire history of the system. This means that the time evolution of the particle number $N_k(t)$ governed by the quantum Vlasov equation is a non-Markovian process.

The mean electric current $\langle j(t) \rangle$ (9.1.12) in the basis of adiabatic number
Plasma oscillations in electric fields

\[ \langle j(t) \rangle = j_{\text{cond}} + j_{\text{pol}}, \quad (9.4.11) \]

\[ j_{\text{cond}} = 2e \int \frac{d^3k}{(2\pi)^3} \frac{[k - eA(t)]}{\omega_k} N_k(t), \quad (9.4.12) \]

\[ j_{\text{pol}} = \frac{2}{E} \int \frac{d^3k}{(2\pi)^3} \omega_k N_k(t), \quad (9.4.13) \]

by using Eqs. (9.4.6), (9.4.7). This means electric current \( \langle j(t) \rangle \) enters into the right-hand side of the field equation (9.1.5).

For the comparison between the quantum Vlasov equation (9.4.10) and the classical Boltzmann–Vlasov equation (9.2.1), the adiabatic particle number \( N_k(t) \) has to be understood as the counterpart of the classical distribution function \( f(p,t) \) of particle number in the phase space. The source term, that is composed by the Schwinger rate of pair production and the factor \([1 \pm 2N_k(t')]\) for either the Bose enhancement or Pauli blocking, is phenomenologically added into the Boltzmann–Vlasov equation (9.2.1). Such source term is local in time, indicating that the time evolution of the classical distribution function \( f(p,t) \) is a Markovian process. In the limit of a very slowly varying uniform electric field \( E \) and at very large time \( t \) the source term \( S_k(t) \) (9.4.10) integrated over momenta \( k \) reduces to the source term in the Boltzmann–Vlasov equation (9.2.1) [437,455,456]. As a result, the conduction current \( j_{\text{cond}} \) and the polarization current \( j_{\text{pol}} \) in Eq. (9.4.11) are reduced to their counterparts Eqs. (9.2.3), (9.2.4) in the phenomenological model of kinetic theory. In Ref. [267,437,458–460], the quantum Vlasov equation has been numerically studied to show the plasma oscillations and the non-Markovian effects. They are also compared with the Boltzmann–Vlasov equation (9.2.1) that corresponds to the Markovian limit.

9.5 Quantum decoherence in plasma oscillations

As showed in Fig. 9.1, the collective oscillations of electric field \( E(t) \) and associate electric current \( \langle j(t) \rangle \) are damped in their amplitude. Moreover, as time increases, a decoherence in their oscillating frequency occurs [435,461]. This indicates that plasma oscillations decay in time. This effective energy dissipation or time irreversibility is the phenomenon of quantum decoherence [462] in the process of creation and oscillation of particles, in the sense that energy flows from collective motion of the classical electromagnetic field to the quantum fluctuations of charged matter fields without returning back over times of physical interest [461,463]. This means that the characteristic frequency \( \omega_k \) of the quantum fluctuation mode “\( k \)” is much larger than the frequency \( \omega_{pl} \) of the classical electric field: \( \omega_k \gg \omega_{pl} \) and \( \omega_{pl}^2 \sim 2e^2\hbar n_p/(m_e c^2) \) [437,461].
9.5 Quantum decoherence in plasma oscillations

where \( n_p \) is the number density of particles and antiparticles. The study of quantum decoherence and energy dissipation associated with particle production to understand the plasma oscillation frequency and damping can be found in Refs. [437, 461, 463].

To understand the energy dissipation from the collective oscillation of classical mean fields to rapid fluctuations of quantum fields, it is necessary to use the Hamiltonian formalism of semi-classical theory. One defines the quantum fluctuation \( \xi_k(t) \) upon classical mean field \( \langle \Phi_k(t) \rangle \) in the semi-classical scalar theory [437, 461, 463],

\[
\xi_k^2(t) = \langle [\Phi_k(t) - \langle \Phi_k(t) \rangle]^2 \rangle = \langle \Phi_k(t) \Phi_k^*(t) \rangle - [\langle \Phi_k(t) \rangle]^2, \tag{9.5.1}
\]

where \( \Phi_k(t) \) is the Fourier \( k \)-component of the quantized scalar field \( \Phi(x) \) (9.1.6). In the time-independent basis (9.1.7,9.1.8), one has (see Section 9),

\[
\xi_k^2(t) = \sigma_k |f_k(t)|^2, \tag{9.5.2}
\]

and the mode equation (9.1.9) for \( f_k(t) \) can be rewritten as

\[
\dot{\eta}_k(t) = \ddot{\xi}_k(t) = -\omega_k^2(t)\xi_k(t) + \frac{\hbar^2 \sigma_k^2}{4\xi_k^2(t)}, \tag{9.5.3}
\]

where \( \eta_k(t) = \xi_k(t) \) is the momentum canonically conjugate to \( \xi_k(t) \). Moreover, the semi-classical Maxwell equation (9.1.13) is rewritten as

\[
\frac{d^2A}{dt^2} = 2e \int \frac{d^3k}{(2\pi)^3} [k - eA(t)]\xi_k^2(t). \tag{9.5.4}
\]

Eqs. (9.5.3) and (9.5.4) actually are Hamilton equations of motion,

\[
\dot{\eta}_k(t) = -\frac{\delta\mathcal{H}_{\text{eff}}}{\delta\xi_k(t)}, \quad \dot{P}_A(t) = -\frac{\delta\mathcal{H}_{\text{eff}}}{\delta A(t)}, \tag{9.5.5}
\]

for a closed system with Hamiltonian,

\[
\mathcal{H}_{\text{eff}}(A, P_A, \xi, \eta, \sigma) = \frac{E^2}{2} + V \int \frac{d^3k}{(2\pi)^3} \left( \eta_k^2 + \omega_k^2(A)\xi_k^2 + \frac{\hbar^2 \sigma_k^2}{4\xi_k^2} \right), \tag{9.5.6}
\]

where \( P_A = \dot{A} = -E \) is the momentum canonically conjugate to \( A \), \( \omega_k(A) \) is the field-dependent frequency of quantum fluctuations given by Eq. (9.1.10) and the value of mean field (9.5.1) vanishes, \( \langle \Phi_k(t) \rangle = 0 \), for each \( k \)-mode. In Eq. (9.5.6), the first term is the electric energy and the second term is the energy of quantum fluctuations of charged matter field, interacting with electric field.

Quantum decoherence can be studied within this Hamiltonian framework.
If one considers only the time evolution of classical electric field \( A(t) \), that is influenced by the quantum fluctuating modes \( f_k(t) \), the latter can be treated as a heat bath “environment”. Quantitative information about the quantum decoherence is contained in the so-called influence functional, which is a functional of two time evolution trajectories \( A_1(t) \) and \( A_2(t) \) [461]

\[
F_{12}(t) = \exp[i\Gamma_{12}(t)] = \text{Tr}(|A_1(t)\rangle\langle A_2(t)|), \tag{9.5.7}
\]

where \(|A_{1,2}(t)\rangle\) are different time evolution states determined by Eq. (9.5.5), starting with the same initial state \(|A(0)\rangle\) and initial vacuum condition \( N_k = 0, \sigma_k = 1 \) (9.1.13). One finds [461]

\[
\Gamma_{12}(t) = -\frac{i}{2} \ln \left[ \frac{\hbar}{|f_1 f_2|} \left( \frac{f_1 f_2^*}{f_1^* f_2 - f_1^* f_2^*} \right) \right], \tag{9.5.8}
\]

in terms of the two sets of mode functions \( \{f_1(t)\} \) and \( \{f_2(t)\} \) (the subscript \( k \) is omitted). This \( \Gamma_{12} \) (9.5.8) is precisely the closed time path (CTP) effective action functional which generates the connected real \( n \)-points vertices in the quantum theory [464–471]. The absolute value of \( F_{12} \) (9.5.7) measures the influence of quantum fluctuations \( f_k(t) \) on the time evolution of the classical electric field \( A(t) \), i.e., the effect of quantum decoherence. If there is no influence of quantum fluctuations on \( A(t) \), then \(|A_1(t)\rangle = |A_2(t)\rangle\) and \(|F_{12}| = 1\), otherwise \( A_2(t) \) deviates from \( A_1(t) \), \(|A_1(t)\rangle \neq |A_2(t)\rangle\) and \(|F_{12}| < 1\). Numerical results about the damping and the decoherence of the electric field are presented in Refs. [435, 461] (see Figs. 9.1 and 9.2).

The effective damping discussed above is certainly collisionless, since the charged particle modes \( f_k(t) \) interact only with the electric field but not directly with each other. The damping of plasma oscillation attributed to the collisions between charged particle modes \( f_k(t) \) will be discussed in the next section.

## 9.6 Collision decoherence in plasma oscillations

If there are interactions between different modes \( k \) and species of particles, the time evolutions of electric field and the distribution function \( F(p, t) \) of particle number in the phase space are certainly changed. This can be phenomenologically studied in the relativistic Boltzmann-Vlasov equation (9.2.1) by adding collision terms \( \mathcal{C}(p, t) \),

\[
\frac{dF}{dt} = \frac{\partial F}{\partial t} + eE \frac{\partial F}{\partial p} = S_p(t) + \mathcal{C}(p, t). \tag{9.6.1}
\]
9.6 Collision decoherence in plasma oscillations

Figure 9.2: Absolute values of the decoherence functional $|F_{12}|$ as a function of time. The two field values are $E$ and $E - \Delta$. The top figure shows (for fixed $\Delta$) the sharp dependence of decoherence on particle production when $|E| \geq 0.2E_c$. The second illustrates the relatively milder dependence on $\Delta$. These figures are reproduced from Fig. 2 in Ref. [461].

These collision terms $C(p,t)$ describe not only the interactions of different modes $k$ of particles, but also interactions of different species of particles, for example, electron and positron annihilation to two photons and vice versa. In Refs. [458, 460], the following equilibrating collision terms were considered,

$$ C(p,t) = \frac{1}{\tau_r} [F_{\text{eq}}(p,t,T) - F(p,t)], \quad (9.6.2) $$

where $F_{\text{eq}}(p,t,T)$ is the thermal (Fermi or Bose) distribution function of particle number (fermions or bosons) at temperature $T$, the relaxation time $\tau_r$ is determined by the mean free path $\lambda(t)$ and mean velocity $\bar{v}(t)$ of particles, through

$$ \tau_r(t) = \tau_c \frac{\lambda(t)}{\bar{v}(t)}, \quad (9.6.3) $$
and \( \tau_c \) is a dimensionless parameter. \( \lambda(t) \) is computed from the number density \( n(t) \) of particles,

\[
\lambda(t) = \frac{1}{[n(t)]^{1/3}}, \quad n(t) = \int \frac{d^3k}{(2\pi)^3} F(p, t). \tag{9.6.4}
\]

Whereas, the mean velocity of particles is given by \( \bar{v}(t) = \bar{p}(t)/\bar{\epsilon}(t) \), expressed in terms of mean kinetic momentum \( \bar{p}(t) \) and energy \( \bar{\epsilon}(t) \) of particles. The mean values of momentum \( \bar{p}(t) \) and energy \( \bar{\epsilon}(t) \) are computed \cite{460} by using distribution function \( F(p, t) \) regularized via the procedure described in Ref. \cite{458} that yields the renormalized electric current (9.4.11). The temperature \( T(t) \) in Eq. (9.6.2) is the “instantaneous temperature”, which is determined by requiring that at each time \( t \) the mean particle energy \( \bar{\epsilon}(t) \) is identical to that in an equilibrium distribution \( F_{\text{eq}}(k, t, T) \) at the temperature \( T(t) \),

\[
\bar{\epsilon}(t) = \int \frac{d^3k}{(2\pi)^3} \epsilon(k) F_{\text{eq}}[k, t, T(t)]. \tag{9.6.5}
\]

This system of two coupled equations: (i) the field equation (9.1.5) with renormalized electric currents (9.4.11); (ii) the relativistic Boltzmann-Vlasov equation (9.6.1) with the source term \( S_p(t) \) (9.4.10) and equilibrating collision term \( C(p, t) \) (9.6.2), are numerically integrated in Refs. \cite{458,460}. One of these numerical results is presented in Fig. 9.3. It shows \cite{458} that when the collision timescale \( \tau_c \) is much larger than the plasma oscillation timescale \( \tau_{pl} \), \( \tau_c \gg \tau_{pl} \), the plasma oscillations are unaffected. On the other hand, when \( \tau_c \sim \tau_{pl} \) the collision term has a significant impact on both the amplitude and the frequency of the oscillations that result damped. There is a value of \( \tau_c \) below which no oscillations arise and the system evolves quickly and directly to thermal equilibrium. It is worthwhile to contrast this collision damping of plasma oscillation with the collisionless damping effect due to rapid quantum fluctuations described in Section 9.5.

### 9.7 \( e^+e^-\gamma \) interactions in plasma oscillations in electric fields

In this section, a detailed report of the studies \cite{72} of the relativistic Boltzmann-Vlasov equations for electrons, positrons and photons with collision terms originated from annihilation of electrons and positrons pair into two photons and vice versa is presented. These collision terms lead to the damping of plasma oscillation and possibly to energy equipartition between different types of particles. We focus on the evolution of a system of \( e^+e^- \) pairs created in a strongly overcritical electric field \( (E \sim 10E_c) \), explicitly taking into account the process \( e^+e^- \rightarrow \gamma\gamma \). Since it is far from equilibrium the collisions
cannot be modeled by an effective relaxation time term in the transport equations, as discussed in the previous section. Rather the actual, time varying, collision integrals have to be used.

Furthermore we are mainly interested in a system in which the electric field varies on macroscopic length scale and therefore one can approximate an electric field as a homogeneous one. Also, transport equations are used for electrons, positrons and photons, with collision terms, coupled to Maxwell equations, as introduced in Section 9.1 and 9.2. There is no free parameter here: the collision terms can be exactly computed, since the QED cross-sections are known. Starting from a regime which is far from thermal equilibrium, one finds that collisions do not prevent plasma oscillations in the initial phase of the evolution and analyze the issue of the timescale of the approach to an $e^+e^-\gamma$ plasma equilibrium configuration.

As discussed in Section 9.1 and 9.2 one can describe positrons (electrons) created by vacuum polarization in a strong homogeneous electric field $E$ through the distribution function $f_{e^+}$ ($f_{e^-}$) in the phase space of positrons

---

**Figure 9.3:** Time evolution for electric field obtained using different relaxation times $\tau_r$ in the collision term of Eq. (9.6.2) and with the impulse external field $E_{ex}(t) = -A_0[b \cosh^2(t/b)]^{-1}$, where $A_0 = 0.7$ and $b = 0.5$. All dimensioned quantities are given in units of the mass-scale $m$. This Figure is reproduced from Fig. 6 in Ref. [458].
Because of homogeneity $f_{e^+} (f_{e^-})$ only depend on the time $t$ and the positron (electron) 3–momentum $p$:

$$f_{e^+,-} = f_{e^+,-} (t, p).$$ (9.7.1)

Moreover, because of particle–antiparticle symmetry, one also has

$$f_{e^+} (t, p) = f_{e^-} (t, -p) \equiv f_e (t, p).$$ (9.7.2)

Analogously photons created by pair annihilation are described through the distribution function $f_\gamma$ in the phase space of photons. $k$ is the photon 3–momentum , then

$$f_\gamma = f_\gamma (t, k).$$ (9.7.3)

$f_e$ and $f_\gamma$ are normalized so that

$$\int \frac{d^3 p}{(2\pi)^3} f_e (t, p) = n_e (t),$$ (9.7.4)

$$\int \frac{d^3 k}{(2\pi)^3} f_\gamma (t, k) = n_\gamma (t),$$ (9.7.5)

where $n_e$ and $n_\gamma$ are number densities of positrons (electrons) and photons, respectively. For any function of the momenta, one can denote by

$$\langle F (p_1, ..., p_n) \rangle_e \equiv n_e^{-n} \int \frac{d^3 p_1}{(2\pi)^3} ... \frac{d^3 p_n}{(2\pi)^3} F (p_1, ..., p_n) \cdot f_e (p_1) \cdot ... \cdot f_e (p_n),$$ (9.7.6)

or

$$\langle G (k_1, ..., k_l) \rangle_\gamma \equiv n_\gamma^{-l} \int \frac{d^3 k_1}{(2\pi)^3} ... \frac{d^3 k_l}{(2\pi)^3} G (k_1, ..., k_l) \cdot f_\gamma (k_1) \cdot ... \cdot f_\gamma (k_l),$$ (9.7.7)

its mean value in the phase space of positrons (electrons) or photons, respectively.

The motion of positrons (electrons) is the result of three contributions: the pair creation, the electric acceleration and the annihilation damping. The probability density rate $S (E, p)$ for the creation of a pair with 3–momentum $p$ in the electric field $E$ is given by the Schwinger formula (see also Refs. [435, 439]):

$$S (E, p) = (2\pi)^3 \frac{dN}{dtd^3x d^4p} = - |eE| \log \left[ 1 - \exp \left( - \frac{\pi (m_e^2 + p_\perp^2)}{|eE|} \right) \right] \delta (p_\parallel),$$ (9.7.8)

where $p_\parallel$ and $p_\perp$ are the components of the 3-momentum $p$ parallel and or-
9.7 $e^+e^-\gamma$ interactions in plasma oscillations in electric fields

thogonal to $E$, respectively. Also the energy is introduced

$$\epsilon_p = \left( p \cdot p + m_e^2 \right)^{1/2} \quad (9.7.9)$$

of an electron of 3-momentum $p$ and the energy

$$\epsilon_k = (k \cdot k)^{1/2} \quad (9.7.10)$$

of a photon of 3-momentum $k$. Then, the probability density rate $C_e(t, p)$ for the creation (destruction) of a fermion with 3-momentum $p$ is given by

$$C_e(t, p) \simeq \frac{1}{\epsilon_p} \int \frac{d^3 p_1}{(2\pi)^3 \epsilon_{p_1}} \frac{d^3 p_2}{(2\pi)^3 \epsilon_{p_2}} \frac{d^3 k_1}{(2\pi)^3 \epsilon_{k_1}} \frac{d^3 k_2}{(2\pi)^3 \epsilon_{k_2}} (2\pi)^4 \delta^{(4)}(p + p_1 - k_1 - k_2)$$

$$\times \left| M_{e^+ e^- (p_1) \rightarrow \gamma (k_1) \gamma (k_2)} \right|^2 \left[ f_e(p_1) f_e(p_2) - f_\gamma(k_1) f_\gamma(k_2) \right], (9.7.11)$$

where

$$M_{e^+ e^- (p_1) \rightarrow \gamma (k_1) \gamma (k_2)} = M_{e^+ e^- (p_2) \rightarrow \gamma (k_1) \gamma (k_2)} \equiv M \quad (9.7.12)$$

is the matrix element for the process

$$e^+ (p_1) e^- (p_2) \rightarrow \gamma (k_1) \gamma (k_2) \quad (9.7.13)$$

and as a first approximation, Pauli blocking and Bose enhancement (see, for instance, Ref. [439]) are neglected. Analogously the probability density rate $C_\gamma(t, k)$ for the creation (annihilation) of a photon with 3-momentum $k$ is given by

$$C_\gamma(t, k) \simeq \frac{1}{\epsilon_k} \int \frac{d^3 p_1}{(2\pi)^3 \epsilon_{p_1}} \frac{d^3 p_2}{(2\pi)^3 \epsilon_{p_2}} \frac{d^3 k_1}{(2\pi)^3 \epsilon_{k_1}} \frac{d^3 k_2}{(2\pi)^3 \epsilon_{k_2}} (2\pi)^4 \delta^{(4)}(p_1 + p_2 - k - k_1)$$

$$\times \left| M_{e^+ e^- (p_1) \rightarrow \gamma (k) \gamma (k_1)} \right|^2 \left[ f_e(p_1) f_e(p_2) - f_\gamma(k) f_\gamma(k_1) \right], (9.7.14)$$

Finally the evolution of the pairs is governed by the transport Boltzmann–Vlasov equations

$$\partial_t f_e + eE \cdot \nabla_p f_e = S_E(E, p) - C_e(t, p), \quad (9.7.15)$$

$$\partial_t f_\gamma = 2C_\gamma(t, k), \quad (9.7.16)$$

Note that the collisional terms (9.7.11) and (9.7.14) are negligible, when created pairs do not produce a dense plasma.

Because pair creation back reacts on the electric field, as seen in Section 9.1 and 9.2, Vlasov equations are coupled with the homogeneous Maxwell equa-
Plasma oscillations in electric fields

\[ \partial_t E = -j_p(E) - j_c(t), \tag{9.7.17} \]

where

\[ j_p(E) = 2 \frac{E}{E^2} \int \frac{d^3p}{(2\pi)^3} \epsilon_p S(E, p) \tag{9.7.18} \]

is the polarization current and

\[ j_c(t) = 2en_e \int \frac{d^3p}{(2\pi)^3} \epsilon_p f_e(p) \tag{9.7.19} \]

is the conduction current (see Ref. [438]).

Eqs. (9.7.15), (9.7.16) and (9.7.17) describe the dynamical evolution of the electron–positron pairs, the photons and the strong homogeneous electric field due to the Schwinger process of pair creation, the pair annihilation into photons and the two photons annihilation into pairs. It is hard to (even numerically) solve this system of integral and partial differential equations. It is therefore useful to introduce a simplification procedure of such a system through an approximation scheme. First of all note that Eqs. (9.7.15) and (9.7.16) can be suitably integrated over the phase spaces of positrons (electrons) and photons to get differential equations for mean values. The following exact equations for mean values are obtained:

\[ \frac{d}{dt} n_e = S(E) - n_e^2 \langle \sigma_1 v' \rangle_e + n_e^2 \langle \sigma_2 v'' \rangle_\gamma, \]
\[ \frac{d}{dt} n_\gamma = 2n_\gamma^2 \langle \sigma_1 v' \rangle_e - 2n_\gamma^2 \langle \sigma_2 v'' \rangle_\gamma, \]
\[ \frac{d}{dt} \langle \epsilon_p \rangle_e = en_e E \cdot \langle v \rangle_e + \frac{1}{2} E \cdot j_p - n_e^2 \langle \epsilon_p \sigma_1 v'' \rangle_e + n_\gamma^2 \langle \epsilon_\gamma \sigma_2 v'' \rangle_\gamma, \]
\[ \frac{d}{dt} \langle \epsilon_k \rangle_\gamma = 2n_\gamma^2 \langle \epsilon_k \sigma_1 v' \rangle_e - 2n_\gamma^2 \langle \epsilon_k \sigma_2 v'' \rangle_\gamma, \]
\[ \frac{d}{dt} \langle p \rangle_e = en_e E - n_e^2 \langle p \sigma_1 v' \rangle_e, \]
\[ \frac{d}{dt} E = -2en_e \langle v \rangle_e - j_p(E), \tag{9.7.20} \]

where

\[ S(E) = \int \frac{d^3p}{(2\pi)^3} S(E, p) \tag{9.7.21} \]
\[ \equiv \frac{dN}{dt d^3x}, \tag{9.7.22} \]

is the total probability rate for Schwinger pair production. In Eqs. (9.7.20), \( v'' = c \) the velocity of light and \( v' = 2|p/\epsilon_p^{\text{CoM}}| \) is the relative velocity between electrons and positrons in the reference frame of the center of mass, where \( p = |p_{e^+}|, \) \( p_{e^-} = -p_{e^+} \) are 3–momenta of electron and positron and \( \epsilon_p^{\text{CoM}} \) are their energies. \( \sigma_1 = \sigma_1(\epsilon_p^{\text{CoM}}) \) is the total cross-section for the process \( e^+e^- \rightarrow \gamma \gamma \), and \( \sigma_2 = \sigma_2(\epsilon_\gamma^{\text{CoM}}) \) is the total cross-section for the process \( \gamma \gamma \rightarrow e^+e^- \), here \( \epsilon_\gamma^{\text{CoM}} \) is the energy of a particle in the reference
9.7 $e^+e^-\gamma$ interactions in plasma oscillations in electric fields

frame of the center of mass.

In order to evaluate the mean values in system (9.7.20) some further hypotheses on the distribution functions are needed. One defines $\bar{p}_\|$, $\bar{\epsilon}_p$ and $\bar{p}_\perp^2$ such that

$$\langle p_\| \rangle_e \equiv \bar{p}_\|,$$  \hfill (9.7.23)

$$\langle \epsilon_p \rangle_e \equiv \bar{\epsilon}_p \equiv (\bar{p}_\|^2 + \bar{p}_\perp^2 + m_e^2)^{1/2}. \hfill (9.7.24)$$

It is assumed

$$f_e (t, p) \propto n_e (t) \delta (p_\| - \bar{p}_\|) \delta (p_\perp^2 - \bar{p}_\perp^2). \hfill (9.7.25)$$

Since in the scattering $e^+e^- \rightarrow \gamma\gamma$ the coincidence of the scattering direction with the incidence direction is statistically favored, it is also assumed

$$f_\gamma (t, k) \propto n_\gamma (t) \delta (k_\|^2 - \bar{k}_\|^2) \left[ \delta (k_\| - \bar{k}_\|) + \delta (k_\| + \bar{k}_\|) \right], \hfill (9.7.26)$$

where $k_\|$ and $k_\perp$ have analogous meaning as $p_\|$ and $p_\perp$ and the terms $\delta (k_\| - \bar{k}_\|)$ and $\delta (k_\| + \bar{k}_\|)$ account for the probability of producing, respectively, forwardly scattered and backwardly scattered photons. Since the Schwinger source term (9.7.8) implies that the positrons (electrons) have initially fixed $p_\|$, namely $p_\| = 0$, assumption (9.7.25) ((9.7.26)) means that the distribution of $p_\|$ (or $k_\|$) does not spread too much with time and, analogously, that the distribution of energies is sufficiently peaked to be describable by a $\delta-$function. As long as this condition is fulfilled, approximations (9.7.25) and (9.7.26) are applicable. The actual dependence on the momentum of the distribution functions has been discussed in Ref. [437, 439]. If Eqs. (9.7.25) and (9.7.26) are substituted into the system (9.7.20) one gets a new system of ordinary differential equations. One can introduce the inertial reference frame which on average coincides with the center of mass frame for the processes $e^+e^- \rightarrow \gamma\gamma$, and has $e^{\text{CoM}} \simeq \bar{\epsilon}$ for each species, and therefore substituting
Plasma oscillations in electric fields

Eqs. (9.7.25) and (9.7.26) into Eqs. (9.7.20) one finds

\[ \frac{d}{dt} n_e = S(E) - 2n_e^2 \sigma_1 \rho_e^{-1} |\pi_e| + 2n_e^2 \sigma_2, \]

\[ \frac{d}{dt} n_\gamma = 4n_e^2 \sigma_1 \rho_e^{-1} |\pi_e| - 2n_\gamma \rho_\gamma \sigma_2, \]

\[ \frac{d}{dt} \rho_e = e n_e E \rho_e^{-1} |\pi_e| + \frac{1}{2} E j_p - 2n_e \rho_e \sigma_1 \rho_e^{-1} |\pi_e| + 2n_\gamma \rho_\gamma \sigma_2, \]

\[ \frac{d}{dt} \rho_\gamma = 4n_e \rho_e \sigma_1 \rho_e^{-1} |\pi_e| - 4n_\gamma \rho_\gamma \sigma_2, \]

\[ \frac{d}{dt} \pi_e = e n_e E - 2n_e \pi_e \sigma_1 \rho_e^{-1} |\pi_e|, \]

\[ \frac{d}{dt} E = -2en_e \rho_e^{-1} |\pi_e| - j_p(E), \] (9.7.27)

where

\[ \rho_e = n_e \bar{\epsilon}_p, \]

\[ \rho_\gamma = n_\gamma \bar{\epsilon}_k, \]

\[ \pi_e = n_e \bar{\epsilon}_p \] (9.7.28)

\[ \pi_e = n_e \bar{\epsilon}_p \] (9.7.29)

\[ \pi_e = n_e \bar{\epsilon}_p \] (9.7.30)

are the energy density of positrons (electrons), the energy density of photons and the density of “parallel momentum” of positrons (electrons), E is the electric field strength and \( j_p \) the unique component of \( j_p \) parallel to \( E \). \( \sigma_1 \) and \( \sigma_2 \) are evaluated at \( \epsilon_{\text{CoM}} = \bar{\epsilon} \) for each species. Note that Eqs. (9.7.27) are “classical” in the sense that the only quantum information is encoded in the terms describing pair creation and scattering probabilities. Finally Eqs. (9.7.27) are duly consistent with energy density conservation:

\[ \frac{d}{dt} \left( \rho_e + \rho_\gamma + \frac{1}{2} E^2 \right) = 0. \] (9.7.31)

Eqs. (9.7.27) have to be integrated with the following initial conditions

\[ n_e = 0, \]

\[ n_\gamma = 0, \]

\[ \rho_e = 0, \]

\[ \rho_\gamma = 0, \]

\[ \pi_e = 0, \]

\[ E = E_0. \]

In Fig. 9.4 the results of the numerical integration for \( E_0 = 9E_c \) is showed. The integration stops at \( t = 150 \tau_C \) (where, as usual, \( \tau_C = \hbar/m_e c^2 \) is the Compton time of the electron). Each quantity is represented in units of \( m_e \) and \( \lambda_C = \hbar/m_e c, \) the Compton length of the electron.
Figure 9.4: Plasma oscillations in a strong homogeneous electric field: early times behavior. Setting $E_0 = 9E_c$, $t < 150\tau_C$ and it is plotted, from the top to the bottom panel: a) electromagnetic field strength; b) electrons energy density; c) electrons number density; d) photons energy density; e) photons number density as functions of time.
The numerical integration confirms [435, 439] that the system undergoes plasma oscillations:

1. the electric field does not abruptly reach the equilibrium value but rather oscillates with decreasing amplitude;
2. electrons and positrons oscillates in the electric field direction, reaching ultrarelativistic velocities;
3. the role of the $e^+e^- \rightarrow \gamma\gamma$ scatterings is marginal in the early time of the evolution.

This last point can be easily explained as follows: since the electrons are too extremely relativistic, the annihilation probability is very low and consequently the density of photons builds up very slowly (see details in Fig. 9.4).

At late times the system is expected to relax to an equilibrium configuration and assumptions (9.7.25) and (9.7.26) have to be generalized to take into account quantum spreading of the distribution functions. It is nevertheless instructive to look at the solutions of Eqs. (9.7.27) in this regime. Moreover, such a solution should give information at least at the order of magnitude level. In Fig. 9.5 the numerical solution of Eqs. (9.7.27) is plotted, but the integration extends here all the way up to $t = 7000 \tau_C$ (the timescale of oscillations is not resolved in these plots).

It is interesting that the leading term recovers the expected asymptotic behavior:

1. the electric field is screened to about the critical value: $E \simeq E_c$ for $t \sim 10^3 - 10^4 \tau_C \gg \tau_C$;
2. the initial electromagnetic energy density is distributed over electron–positron pairs and photons, indicating energy equipartition;
3. photons and electron–positron pairs number densities are asymptotically comparable, indicating number equipartition.

At such late times a regime of thermalized electrons–positrons–photons plasma is expected to begin (as qualitatively indicated by points 2 and 3 above) during which the system is describable by hydrodynamic equations [399,408].

Let us summarize the results in this section. A very simple formalism is provided to describe simultaneously the creation of electron–positron pairs by a strong electric field $E \gtrsim E_c$ and the pairs annihilation into photons. As discussed in literature, one finds plasma oscillations. In particular the collisions do not prevent such a feature. This is because the momentum of electrons (positrons) is very high, therefore the cross-section for the process $e^+e^- \rightarrow \gamma\gamma$ is small and the annihilation into photons is negligible in the very first phase of the evolution. As a result, the system takes some time ($t \sim 10^3 - 10^4 \tau_C$) to reach an equilibrium $e^+e^-\gamma$ plasma configuration.
9.7 $e^+e^-\gamma$ interactions in plasma oscillations in electric fields

Figure 9.5: Plasma oscillations in a strong homogeneous electric field: late time expected behaviour. Setting $E_0 = 9E_C$, $t < 7000\tau_C$ and it is plotted, from the top to the bottom panel: a) electromagnetic field strength; b) electrons energy density; c) electrons number density; d) photons energy density; e) photons number density as functions of time – the oscillation period is not resolved in these plots. The model should have a breakdown at a time much earlier than $7000\tau_C$ and therefore this plot contains no more than qualitative informations.
9.8 Electro-fluidodynamics of the pair plasma

In the previous section, collisional terms in the Vlasov-Boltzmann equation are introduced, describing interaction of pairs and photons via the reaction $e^+e^- \leftrightarrow \gamma \gamma$. These results have been considered of interest in the studies of pair production in free electron lasers [60, 238, 239, 261, 264], in optical lasers [472], of millicharged fermions in extensions of the standard model of particle physics [473], electromagnetic wave propagation in a plasma [269], as well in astrophysics [408].

In this section, following [73, 474], the case of undercritical electric field is explored. It is usually expected that for $E < E_c$ back reaction of the created electrons and positrons on the external electric field can be neglected and electrons and positrons would move as test particles along lines of force of the electric field. Here it is shown that this is not the case in a uniform unbounded field. This work is important since the first observation of oscillations effects should be first detectable in experiments for the regime $E < E_c$, in view of the rapid developments in experimental techniques, see e.g. [270, 273, 274].

An approach is introduced based on continuity, energy-momentum conservation and Maxwell equations in order to account for the back reaction of the created pairs. This approach is more simple than the one, presented in the previous section. However, the final equations coincide with (9.7.27) when the interaction with photons can be neglected. By this treatment one can analyze the new case of undercritical field, $E < E_c$, and recover the old results for overcritical field, $E > E_c$. In particular, the range $0.15E_c < E < 10E_c$ is focused.

It is generally assumed that electrons and positrons are created at rest in pairs, due to vacuum polarization in uniform electric field with strength $E$ [7, 20, 25–27, 33, 203, 204], with the average rate per unit volume and per unit time (5.7.29)

$$S(E) \equiv \frac{dN}{dVdt} = \frac{m_e^4}{4\pi^3} \left( \frac{E}{E_c} \right)^2 \exp \left( -\pi \frac{E}{E_c} \right). \quad (9.8.1)$$

This formula is derived for uniform constant in time electric field. However, it still can be used for slowly time varying electric field provided the inverse adiabaticity parameter [30–32, 36, 203, 204, 244], see Eq. (7.2.8), is much larger than one,

$$\eta = \frac{m_e E_{\text{peak}}}{\omega E_c} = \tilde{T} E_{\text{peak}} \gg 1, \quad (9.8.2)$$

where $\omega$ is the frequency of oscillations, $\tilde{T} = m_e/\omega$ is dimensionless period of oscillations. Equation (9.8.2) implies that time variation of the electric field is much slower than the rate of pair production. In two specific cases considered in this section, $E = 10E_c$ and $E = 0.15E_c$ one finds for the first oscillation $\eta = 334$ and $\eta = 3.1 \times 10^6$ respectively. This demonstrates applicability of
9.8 Electro-fluidodynamics of the pair plasma

the formula (9.8.1) in this case.

From the continuity, energy-momentum conservation and Maxwell equations written for electrons, positrons and electromagnetic field one can have

\[
\frac{\partial (\bar{n} U^\mu)}{\partial x^\mu} = S, \quad (9.8.3)
\]

\[
\frac{\partial T^{\mu\nu}}{\partial x^\nu} = -F^{\mu\nu} J_\nu, \quad (9.8.4)
\]

\[
\frac{\partial F^{\mu\nu}}{\partial x^\nu} = -4\pi J^\mu, \quad (9.8.5)
\]

where \(\bar{n}\) is the comoving number density of electrons, \(T^{\mu\nu}\) is energy-momentum tensor of electrons and positrons

\[
T^{\mu\nu} = m_e \bar{n} \left( U^\mu_{(+)} U^\nu_{(+)} + U^\mu_{(-)} U^\nu_{(-)} \right), \quad (9.8.6)
\]

\(F^{\mu\nu}\) is electromagnetic field tensor, \(J^\mu\) is the total 4-current density, \(U^\mu\) is four velocity respectively of positrons and electrons

\[
U^\mu_{(+)} = U^\mu = \gamma (1, v, 0, 0), \quad U^\mu_{(-)} = \gamma (1, -v, 0, 0), \quad (9.8.7)
\]

\(v\) is the average velocity of electrons, \(\gamma = (1 - v^2)^{-1/2}\) is relativistic Lorentz factor. Electrons and positrons move along the electric field lines in opposite directions.

One can choose a coordinate frame where pairs are created at rest. Electric field in this frame is directed along \(x\)-axis. In spatially homogeneous case from (9.8.3) one has for coordinate number density \(n = \bar{n} \gamma\)

\[
\dot{n} = S. \quad (9.8.8)
\]

With definitions (9.8.6) from (9.8.4) and equation of motion for positrons and electrons

\[
m_e \frac{\partial U^\mu_{(\pm)}}{\partial x^\nu} = \mp e F^\mu_{\nu}, \quad (9.8.9)
\]

one finds

\[
\frac{\partial T^{\mu\nu}}{\partial x^\nu} = -e\bar{n} \left( U^\nu_{(+)} - U^\nu_{(-)} \right) F^\mu_{\nu} + m_e S \left( U^\mu_{(+)} + U^\mu_{(-)} \right) = -F^{\mu\nu} J_\nu, \quad (9.8.10)
\]

where the total 4-current density is the sum of conducting \(J^\mu_{\text{cond}}\) and polariza-
tion $J_{\mu}^{pol}$ currents \cite{438} densities

$$J_{\mu} = J_{\mu}^{cond} + J_{\mu}^{pol}, \quad (9.8.11)$$

$$J_{\mu}^{cond} = e\bar{n} \left( U_{(+)}^{\mu} - U_{(-)}^{\mu} \right), \quad (9.8.12)$$

$$J_{\mu}^{pol} = \frac{2m_eS}{E} \gamma (0, 1, 0, 0). \quad (9.8.13)$$

Energy-momentum tensor in \cite{9.8.4} and electromagnetic field tensor in \cite{9.8.5} change for two reasons: 1) electrons and positrons acceleration in the electric field, given by the term $J_{\mu}^{\mu, \text{cond}}$, 2) particle creation, described by the term $J_{\mu}^{\mu, \text{pol}}$. Equation \cite{9.8.3} is satisfied separately for electrons and positrons.

Defining energy density of positrons

$$\rho = \frac{1}{2} T_{00}^{00} = m_e \bar{n} \gamma, \quad (9.8.14)$$

one can find from \cite{9.8.4}

$$\dot{\rho} = envE + m_e \gamma S. \quad (9.8.15)$$

Due to homogeneity of the electric field and plasma, electrons and positrons have the same energy and absolute value of the momentum density $p$, but their momenta have opposite directions. The definitions also imply for velocity and momentum densities of electrons and positrons

$$v = \frac{p}{\rho}, \quad (9.8.16)$$

and

$$\rho^2 = p^2 + m_e^2 n^2, \quad (9.8.17)$$

which is just relativistic relation between the energy, momentum and mass densities of particles.

Gathering together the above equations one then has the following equations

$$\dot{n} = S, \quad (9.8.18)$$

$$\dot{\rho} = E \left( env + \frac{m_e \gamma S}{E} \right), \quad (9.8.19)$$

$$\dot{p} = enE + m_e v \gamma S, \quad (9.8.20)$$

$$\dot{E} = -8\pi \left( env + \frac{m_e \gamma S}{E} \right). \quad (9.8.21)$$
From (9.8.19) and (9.8.21) one obtains the energy conservation equation

\[
\frac{E_0^2 - E^2}{8\pi} + 2\rho = 0, \quad (9.8.22)
\]

where \(E_0\) is the constant of integration, so the particle energy density vanishes for initial value of the electric field, \(E_0\).

These equations give also the maximum number of the pair density asymptotically attainable consistently with the above rate equation and energy conservation

\[
n_0 = \frac{E_0^2}{8\pi m_e}. \quad (9.8.23)
\]

For simplicity dimensionless variables \(n = m_3^2 \tilde{n}, \rho = m_4^2 \tilde{\rho}, \ p = m_4^2 \tilde{p}, \ E = E_c \tilde{E}, \) and \(t = m_e^{-1} \tilde{t}\) are introduced. With these variables the system of equations (9.8.18)-(9.8.21) takes the form

\[
\frac{d \tilde{n}}{d \tilde{t}} = \tilde{S}, \\
\frac{d \tilde{\rho}}{d \tilde{t}} = \tilde{n} \tilde{E} \tilde{\rho} + \tilde{\gamma} \tilde{S}, \\
\frac{d \tilde{p}}{d \tilde{t}} = \tilde{n} \tilde{E} + \tilde{\gamma} \tilde{\rho} \tilde{S}, \\
\frac{d \tilde{E}}{d \tilde{t}} = -8\pi \alpha \left( \tilde{n} \tilde{\rho} + \tilde{\gamma} \frac{\tilde{S}}{\tilde{E}} \right),
\]

where \(\tilde{S} = \frac{1}{4\pi^2} \tilde{E}^2 \exp \left( -\frac{\tilde{n}}{\tilde{E}} \right), \tilde{\sigma} = \tilde{\rho}/\tilde{n}\) and \(\tilde{\gamma} = (1 - \tilde{\sigma}^2)^{-1/2}, \alpha = e^2/(\hbar c)\) as before.

The system of equations (9.8.24) is solved numerically with the initial conditions \(n(0) = \rho(0) = v(0) = 0\), and the electric field \(E(0) = E_0\).

In fig. 9.6, electric field strength, number density, velocity and Lorentz gamma factor of electrons as functions of time, are presented for initial values of the electric field \(E_0 = 10E_c\) (left column) and \(E_0 = 0.15E_c\) (right column). Slowly decaying plasma oscillations develop in both cases. The half-life of oscillations to be \(10^3t_c\) for \(E_0 = 10E_c\) and \(10^5t_c\) for \(E_0 = 0.8E_c\) are estimated respectively. The period of the fist oscillation is 50\(t_c\) and \(3 \times 10^7t_c\), the Lorentz factor of electrons and positrons in the first oscillation equals 75 and \(3 \times 10^5\) respectively for \(E_0 = 10E_c\) and \(E_0 = 0.15E_c\). Therefore, in contrast to the case \(E > E_c\), for \(E < E_c\) plasma oscillations develop on a much longer timescale, electrons and positrons reach extremely relativistic velocities.

In fig. 9.7 the characteristic length of oscillations is shown together with the distance between the pairs at the moment of their creation. For constant electric field the formation length for the electron–positron pairs, or the quantum tunneling length, is not simply \(m_e/(eE)\), as expected from a semi-classical
9 Plasma oscillations in electric fields

Figure 9.6: Electric field strength, number density of electrons, their velocity and Lorentz gamma factor depending on time with $E_0 = 10E_c$ (left column) and $E_0 = 0.15E_c$ (right column). Electric field, number density and velocity of positron are measured respectively in terms of the critical field $E_c$, Compton volume $\lambda^3_C = \left(\frac{\hbar}{mc}\right)^3$, and the speed of light $c$. The length of oscillation is defined as $D = c\tau$, where $\tau$ is the time needed for the first half-oscillation, shown above.
approximation, but [28,46]

\[ D^* = \frac{m_e}{eE} \left( \frac{E_c}{E} \right)^{1/2}. \]  

(9.8.25)

Thus, given initial electric field strength one can define two characteristic distances: \( D^* \), the distance between created pairs, above which pair creation is possible, and the length of oscillations, \( D = c\tau \), above which plasma oscillations occur in a uniform electric field. The length of oscillations is the maximal distance between two turning points in the motion of electrons and positrons (see fig. 9.7). From fig. 9.7 it is clear that \( D \gg D^* \). In the oscillation phenomena the larger electric field is, the larger becomes the density of pairs and therefore the back reaction, or the screening effect, is stronger. Thus the period of oscillations becomes shorter. Note that the frequency of oscillation is not equal to the plasma frequency, so it cannot be used as the measure of the latter. Notice that for \( E \ll E_c \) the length of oscillations becomes macroscopically large.

**Figure 9.7**: Maximum length of oscillations (black curves) together with the distance between electron and positron in a pair (red curve) computed from (9.8.25), depending on initial value of electric field strength. The solid black curve is obtained from solutions of exact equations (9.8.24), while the dotted black curve corresponds to solutions of approximate equation (9.8.27).

At fig. 9.8 maximum Lorentz gamma factor in the first oscillation is presented depending on initial value of the electric field. Since in the successive oscillations the maximal value of the Lorentz factor is monotonically decreasing (see fig. 9.6) It is concluded that for every initial value of the elec-
tric field there exists a maximum Lorentz factor attainable by the electrons and positrons in the plasma. It is interesting to stress the dependence of the Lorentz factor on initial electric field strength. The kinetic energy contribution becomes overwhelming in the $E < E_c$ case. On the contrary, in the case $E > E_c$ the electromagnetic energy of the field goes mainly into the rest mass energy of the pairs.

This diagram clearly shows that never in this process the test particle approximation for the electrons and positrons motion in uniform electric field can be applied. Without considering back reaction on the initial field, electrons and positrons moving in a uniform electric field would experience constant acceleration reaching $v \sim c$ for $E = E_c$ on the timescale $t_c$ and keep that speed thereafter. Therefore, the back reaction effects in a uniform field are essential both in the case of $E > E_c$ and $E < E_c$.

The average rate of pair creation is compared for two cases: when the electric field value is constant in time (an external energy source keeps the field unchanged) and when it is self-regulated by equations (9.8.24). The result is represented in fig. 9.9. It is clear from fig. 9.9 that when the back reaction effects are taken into account, the effective rate of the pair production is smaller than the corresponding rate (9.8.1) in a uniform field $E_0$. At the same time, discharge of the field takes much longer time. In order to quantify this effect we need to compute the efficiency of the pair production defined as $\epsilon = n(t_S)/n_0$ where $t_S$ is the time when pair creation with the constant rate

![Figure 9.8: Maximum Lorentz gamma factor $\gamma$ reached at the first oscillation depending on initial value of the electric field strength.](image-url)
\( S(E_0) \) would stop, and \( n_0 \) is defined above, see (9.8.23). For \( E_0 = E_c \) one finds \( \epsilon = 14\% \), while for \( E_0 = 0.3E_c \) one has \( \epsilon = 1\% \).

![Figure 9.9: The average rate of pair production \( n/t \) is shown as function of time (thick curve), comparing to its initial value \( S(E_0) \) (thin line) for \( E_0 = E_c \). The dashed line marks the time when the energy of electric field would have exhausted if the rate kept constant.](image)

It is clear from the structure of the above equations that for \( E < E_c \) the number of pairs is small, electrons and positrons are accelerated in electric field and the conducting current dominates. Assuming electric field to be weak, the polarization current is neglected in energy conservation (9.8.19) and in Maxwell equation (9.8.21). This means energy density change due to acceleration is much larger than the one due to pair creation,

\[
E_{\text{env}} \gg m_e \gamma S. \quad (9.8.26)
\]

In this case oscillations equations (9.8.18)-(9.8.21) simplify. From (9.8.19) and (9.8.20) one has \( \dot{\rho} = v \dot{\varphi} \), and using (9.8.16) obtains \( v = \pm 1 \). This is the limit when rest mass energy is much smaller than the kinetic energy, \( \gamma \gg 1 \).

One may therefore use only the first and the last equations from the above set. Taking time derivative of the Maxwell equation one arrives to a single second order differential equation

\[
\ddot{E} + \frac{2em_e^2}{\pi^2} \left( \frac{E}{E_c} \right) \left| \frac{E}{E_c} \right| \exp \left( -\pi \left| \frac{E_c}{E} \right| \right) = 0. \quad (9.8.27)
\]

Equation (9.8.27) is integrated numerically to find the length of oscillations shown in fig. 9.7 for \( E < E_c \). Notice that condition (9.8.26) means ultrarelativistic approximation for electrons and positrons, so that although according
to (9.8.18) there is creation of pairs with rest mass $2m$ for each pair, the corresponding increase of plasma energy is neglected, as can be seen from (9.8.26).

Now one can turn to qualitative properties of the system (9.8.18)–(9.8.21). These nonlinear ordinary differential equations describe certain dynamical system which can be studied by using methods of qualitative analysis of dynamical systems. The presence of the two integrals (9.8.17) and (9.8.22) allows reduction of the system to two dimensions. It is useful to work with the variables $\tilde{v}$ and $\tilde{E}$. In these variables one has

$$\frac{d\tilde{\varphi}}{dt} = \left(1 - \tilde{\varphi}^2\right)^{3/2} \tilde{E}, \quad (9.8.28)$$

$$\frac{d\tilde{E}}{dt} = -\frac{1}{2} \tilde{\varphi} \left(1 - \tilde{\varphi}^2\right)^{1/2} \left(\tilde{E}_0^2 - \tilde{E}^2\right) - 8\pi\alpha \frac{\tilde{S}}{\tilde{E} \left(1 - \tilde{\varphi}^2\right)^{1/2}}. \quad (9.8.29)$$

Introducing the new time variable $\tau$

$$\frac{d\tau}{dt} = \left(1 - \tilde{\varphi}^2\right)^{-1/2} \quad (9.8.30)$$

one arrives at

$$\frac{d\tilde{\varphi}}{d\tau} = \left(1 - \tilde{\varphi}^2\right)^2 \tilde{E}, \quad (9.8.31)$$

$$\frac{d\tilde{E}}{d\tau} = -\frac{1}{2} \tilde{\varphi} \left(1 - \tilde{\varphi}^2\right) \left(\tilde{E}_0^2 - \tilde{E}^2\right) - 8\pi\alpha \frac{\tilde{S}}{\tilde{E}}. \quad (9.8.32)$$

Clearly the phase space is bounded by the two curves $\tilde{\varphi} = \pm 1$. Moreover, physical requirement $\rho \geq 0$ leads to existence of two other bounds $\tilde{E} = \pm \tilde{E}_0$. This system has only one singular point in the physical region, of the type focus at $\tilde{E} = 0$ and $\tilde{\varphi} = 0$.

The phase portrait of the dynamical system (9.8.31),(9.8.32) is represented at fig. 9.10. Thus, every phase trajectory tends asymptotically to the only singular point at $\tilde{E} = 0$ and $\tilde{\varphi} = 0$. This means oscillations stop only when electric field vanishes. At that point clearly

$$\rho = m_e n. \quad (9.8.33)$$

is valid. i.e. all the energy in the system transforms just to the rest mass of the pairs.

In order to illustrate details of the phase trajectories shown at fig. 9.10 only 1.5 cycles are plotted at fig. 9.11. One can see that the deviation from closed curves, representing undamped oscillations and shown by dashed curves, is maximal when the field peaks, namely when the pair production rate is maximal.

The above treatment has been done by considering uniquely back reaction
Figure 9.10: Phase portrait of the two-dimensional dynamical system (9.8.31), (9.8.32). Tildes are omitted. Notice that phase trajectories are not closed curves and with each cycle they approach the point with $\tilde{E} = 0$ and $\tilde{\varphi} = 0$. 
Figure 9.11: Phase trajectory for 1.5 cycles (thick curve) compared with solutions where the Schwinger pair production is switched off (dashed curves).
of the electron–positron pairs on the external uniform electric field. The only source of damping of the oscillations is pair production, i.e. creation of mass. As analysis shows the damping in this case is exponentially weak. However, since electrons and positrons are strongly accelerated in electric field the bremsstrahlung radiation may give significant contribution to the damping of oscillations and further reduce the pair creation rate. Therefore, the effective rate shown in fig. 9.9 will represent an upper limit. In order to estimate the effect of bremsstrahlung, the classical formula for the radiation loss in electric field is recalled

\[
I = \frac{2}{3} \frac{e^4}{m_e^2} E^2 = \frac{2}{3} \alpha m_e^2 \left( \frac{E}{E_c} \right)^2. \tag{9.8.34}
\]

Thus the equations (9.8.19) and (9.8.20), generalized for bremsstrahlung, are

\[
\dot{\rho} = E \left( e n v + \frac{m_e \gamma S}{E} \right) - \frac{2}{3} e^4 m_e E^2, \tag{9.8.35}
\]

\[
\dot{\rho} = e n E + m_e v \gamma S - \frac{2}{3} e^4 m_e E^2 v. \tag{9.8.36}
\]

while equations (9.8.18) and (9.8.21) remain unchanged. Assuming that new terms are small, relations (9.8.17) and (9.8.22) are still approximately satisfied.

Now damping of the oscillations is caused by two terms:

\[
\frac{\bar{\gamma}}{4\pi^2} \bar{E}^2 \exp \left( -\frac{\pi}{E} \right) \quad \text{and} \quad \frac{2}{3} \alpha \bar{E}^2. \tag{9.8.37}
\]

The modified system of equations is integrated, taking into account radiation loss, starting with \(E_0 = 10E_c\). The results are presented in Fig. 9.12 where the sum of the energy of electric field and electron–positron pairs normalized to the initial energy is shown as a function of time. The energy loss reaches 20 percent for 400 Compton times. Thus the effect of bremsstrahlung is as important as the effect of collisions considered in [72] for \(E > E_c\), leading to comparable energy loss for pairs on the same timescale. For \(E < E_c\) one expects that the damping due to bremsstrahlung dominates, but the correct description in this case requires Vlasov–Boltzmann treatment [74].

The damping of the plasma oscillations due to electron–positron annihilation into photons has been addressed in [72]. There it was found that the system evolves towards an electron–positron-photon plasma reaching energy equipartition. Such a system undergoes self-acceleration process following the work of [399].

Therefore the following conclusions are reached:

- It is usually assumed that for \(E < E_c\) electron–positron pairs, created by the vacuum polarization process, move as charged particles in exter-
9 Plasma oscillations in electric fields

Figure 9.12: Losses of the energy due to classical bremsstrahlung radiation. The energy density of the system of electrons, positrons and the electric field normalized to the initial energy density is shown without (solid line) and with (dashed line) the effect of bremsstrahlung.

The existence of plasma oscillations of the electron–positron pairs also for $E \lesssim E_c$ is demonstrated. The corresponding results for $E > E_c$ are well known in the literature. For both cases the maximum Lorentz factors $\gamma_{\text{max}}$ reached by electrons and positrons are determined. The length of oscillations is $10 \ h/(m_e c)$ for $E_0 = 10E_c$, and $10^7 \ h/(m_e c)$ for $E_0 = 0.15E_c$. The asymptotic behavior in time, $t \to \infty$, of the plasma oscillations by the phase portrait technique is also studied.

- For $E > E_c$ the vacuum polarization process transforms the electromagnetic energy of the field mainly in the rest mass of pairs, with moderate contribution to their kinetic energy: for $E_0 = 10E_c$ one finds $\gamma_{\text{max}} = 76$. For $E < E_c$ the kinetic energy contribution is maximized with respect to the rest mass of pairs: $\gamma_{\text{max}} = 8 \times 10^5$ for $E_0 = 0.15E_c$.

- In the case of oscillations the effective rate of pair production is smaller than the rate in uniform electric field constant in time, and consequently, the discharge process lasts longer. The half-life of oscillations is $10^3 t_c$ for $E_0 = 10E_c$ and $10^5 t_c$ for $E_0 = 0.8E_c$. The efficiency of pair production is computed with respect to the one in a uniform constant field. For $E = 0.3E_c$ the efficiency is reduced to one percent, decreasing further for smaller initial electric field.

All these considerations apply to a uniform electric field unbounded in space. The presence of a boundary or a gradient in electric field would re-
quire the use of partial differential equations, in contrast to the ordinary dif-
ferential equations used here. This topic needs further study. The effect of
bremsstrahlung for $E > E_c$ is also estimated, and it is found that it represents
comparable contribution to the damping of the plasma oscillations caused by
collisions [72]. It is therefore clear, that the effect of oscillations introduces a
new and firm upper limit to the rate of pair production which would be fur-
ther reduced if one takes into account bremsstrahlung, collisions and bound-
ary effects.
10 Thermalization of the mildly relativistic pair plasma

An electron–positron plasma is of interest in many fields of physics and astrophysics. In the early Universe [454,475] during the lepton era, ultrarelativistic electron–positron pairs contributed to the matter contents of the Universe. In GRBs electron–positron pairs play an essential role in the dynamics of expansion [476,477], [399]. Indications exist on the presence of the pair plasma also in active galactic nuclei [478], in the center of our Galaxy [479], around hypothetical quark stars [480]. In the laboratory pair plasma is expected to appear in the fields of ultraintense lasers [472].

In many stationary astrophysical sources the pair plasma is thought to be in thermodynamic equilibrium. A detailed study of the relevant processes [481–486], radiation mechanisms [487], possible equilibrium configurations [483, 488] and spectra [489] in an optically thin pair plasma has been carried out. Particular attention has been given to collisional relaxation process [490,491], pair production and annihilation [492], relativistic bremsstrahlung [493,494], double Compton scattering [495,496].

An equilibrium occurs if the sum of all reaction rates vanishes. For instance, electron–positron pairs are in equilibrium when the net pair production (annihilation) rate is zero. This can be achieved by variety of ways and the corresponding condition can be represented as a system of algebraic equations [497]. However, the main assumption made in all the above mentioned works is that the plasma is in thermodynamic equilibrium.

At the same time, in some cases considered above the pair plasma can be optically thick. Although moderately thick plasmas have been considered in the literature [498], only qualitative description is available for large optical depths. Assumption of thermal equilibrium is often adopted for rapidly evolving systems without explicit proof [476,477], [399,499]. Then hydrodynamic approximation is usually applied both for leptons and photons. However, particles may not be in equilibrium initially. Moreover, it is very likely situation, especially in the early Universe or in transient events when the energy is released on a very short timescale and there is not enough time for the system to relax to thermal equilibrium configuration.

Ultrarelativistic expansion of GRB sources is unprecedented in astrophysics. There are indications that relativistic jets in X-ray binaries have Lorentz factors $\gamma \sim 2 - 10$ while in active galactic nuclei $\gamma \sim 10 - 20$ [500], but some bursts sources have $\gamma \sim 400$ and possibly larger [501]. There is a consensus
in the literature that the acceleration required to reach ultrarelativistic velocity in astrophysical flows comes from the radiation pressure, namely from photons and electron–positron pairs. Therefore, the source does not move as a whole, but expands from a compact region, almost reaching the speed of light. The bulk of radiation is emitted far from the region of formation of the plasma, when it becomes transparent for photons, trapped initially inside by the huge optical depth. Thus the plasma is optically thick at the moment of its formation and intense interactions between electrons, positrons and photons take place in it. Even if initially the energy is released in the form of only photons, or only pairs, the process of creation and annihilation of pairs soon redistribute the energy between particles in such a way that the final state will be a mixture of pairs and photons. The main question arises: what is the initial state, prior to expansion, of the pair plasma? Is it in a kind of equilibrium and, if so, is it thermal equilibrium, as expected from the optically thick plasma? Stationary sources in astrophysics have enough time for such an equilibrium to be achieved. On the contrary, for transient sources with the timescale of expansion of the order of milliseconds it is not at all clear that equilibrium can be reached.

In the literature there is no consensus on this point. Some authors considered thermal equilibrium as the initial state prior to expansion [399, 476], while others did not [77]. In fact, the study of the pair plasma equilibrium configurations in detail, performed in [488], cannot answer this question, because essentially nonequilibrium processes have to be considered.

Thus, observations provide motivation for theoretical analysis of physical conditions taking place in the sources of GRBs, and more generally, in nonequilibrium optically thick pair plasma. Notice that there is substantial difference between the ion-electron plasma on the one hand and electron–positron plasma on the other hand. Firstly, the former is collisionless in the wide range of parameters, while collisions are always essential in the latter. Secondly, when collisions are important relevant interactions in the former case are Coulomb scattering of particles which are usually described by the classical Rutherford cross-section. In contrast, interactions in the pair plasma are described by quantum cross-sections even if the plasma itself can be still considered as classical one.

The study reported in [74, 75] in the case of pure pair plasma clarified the issue of initial state of the pair plasma in GRB sources. The numerical calculations show that the pair plasma quickly reach thermal equilibrium prior to expansion, due to intense binary and triple collisions. In this Section details about the computational scheme adopted in [74] are given. Generalization to the presence of proton loading is given in [502].
10.1 Qualitative description of the pair plasma

First of all the domain of parameters characterizing the pair plasma considered in this Section is specified. It is convenient to use dimensionless parameters usually adopted for this purpose.

Mildly relativistic pair plasma is considered, thus the average energy per particle $\bar{\epsilon}$ brackets the electron rest mass energy

$$0.1 \text{ MeV} \lesssim \bar{\epsilon} \lesssim 10 \text{ MeV}. \quad (10.1.1)$$

The lower boundary is required for significant concentrations of pairs, while the upper boundary is set to avoid substantial production of other particles such as muons.

The plasma parameter is $g = \left( n_- d^3 \right)^{-1}$, where $d = \sqrt{\frac{k_B T_-}{4 \pi n_-}} = \frac{\varepsilon}{\omega} \sqrt{\theta_-}$ is the Debye length, $k_B$ is Boltzmann’s constant, $n_-$ and $T_-$ are electron number density and temperature respectively, $\theta_- = k_B T_- / (m_e c^2)$ is dimensionless temperature, $\omega = \sqrt{4 \pi e^2 n_- / m_e}$ is the plasma frequency. To ensure applicability of kinetic approach it is necessary that the plasma parameter is small, $g \ll 1$. This condition means that kinetic energy of particles dominates their potential energy due to mutual interaction. For the pair plasma considered in this Section this condition is satisfied.

Further, the classicality parameter, defined as $\kappa = \frac{e^2}{(\hbar v_r)} = \frac{\alpha}{\beta_r}$, where $v_r = \beta_r c$ is mean relative velocity of particles, see (10.7.12). The condition $\kappa \gg 1$ means that particles collisions can be considered classically, while for $\kappa \ll 1$ quantum description is required. Both for pairs and protons quantum cross-sections are used since $\kappa < 1$.

The strength of screening of the Coulomb interactions is characterized by the Coulomb logarithm $\Lambda = m_e d v_r / \hbar$. Coulomb logarithm varies with mean particle velocity, and it cannot be set a constant as in most of studies of the pair plasma.

Finally, pair plasma is considered with linear dimensions $R$ exceeding the mean free path of photons $l = (n_- \sigma)^{-1}$, where $n_-$ is concentration of electrons and $\sigma$ is the corresponding total cross-section. Thus the optical depth $\tau = n \sigma R \gg 1$ is large, and interactions between photons and other particles have to be taken in due account. These interaction are reviewed in the next section.

Note that natural parameter for perturbative expansion in the problem under consideration is the fine structure constant $\alpha$.

Pure pair plasma composed of electrons $e^-$, positrons $e^+$, and photons $\gamma$ is considered. It is assumed that pairs or photons appear by some physical process in the region with a size $R$ and on a timescale $t < R / c$. We assume that distribution functions of particles depend on neither spatial coordinates nor direction of momentum $f_i = f_i(\bar{\epsilon}, t)$, i.e. isotropic in momentum space.
and uniform plasma is considered.

To make sure that classical kinetic description is adequate the dimensionless degeneracy temperature is estimated

\[ \theta_F = \left[ \left( \frac{\hbar}{m_e c} \right)^2 \left( 3\pi^2 n_e \right)^{\frac{2}{3}} + 1 \right]^{1/2} - 1, \] (10.1.2)

and compared with the estimated temperature in thermal equilibrium. With initial conditions (10.1.1) the degeneracy temperature is always smaller than the temperature in thermal equilibrium and therefore the classical kinetic approach is applied. Besides, since ideal plasma is considered with the plasma parameter \( g \sim 10^{-3} \) it is possible to use one-particle distribution functions. These considerations justify the computational approach based on classical relativistic Boltzmann equation. At the same time the right-hand side of Boltzmann equations contains collisional integrals with quantum and not classical matrix elements, as discussed above.

Relativistic Boltzmann equations [503], [504] in spherically symmetric case are

\[ \frac{1}{c} \frac{\partial f_i}{\partial t} + \beta_i \left( \frac{\mu}{r} \frac{\partial f_i}{\partial \mu} + \frac{1 - \mu^2}{r} \frac{\partial f_i}{\partial \mu} \right) - \nabla U \frac{\partial f_i}{\partial p} = \sum_q \left( \eta_i^q - \chi_i^q f_i \right), \] (10.1.3)

where \( \mu = \cos \theta, \theta \) is the angle between the radius vector \( r \) from the origin and the particle momentum \( p, U \) is a potential due to some external force, \( \beta_i = v_i/c \) are particles velocities, \( f_i(\epsilon, t) \) are their distribution functions, the index \( i \) denotes the type of particle, \( \epsilon \) is their energy, and \( \eta_i^q \) and \( \chi_i^q \) are the emission and the absorption coefficients for the production of a particle of type “i” via the physical process labeled by \( q \). This is a coupled system of partial–integro-differential equations. For homogeneous and isotropic distribution functions of electrons, positrons and photons Eqs. (10.1.3) reduce to

\[ \frac{1}{c} \frac{\partial f_i}{\partial t} = \sum_q \left( \eta_i^q - \chi_i^q f_i \right), \] (10.1.4)

which is a coupled system of integro-differential equations. In (10.1.4) the Vlasov term \( \nabla U \frac{\partial f_i}{\partial p} \) is explicitly neglected.

Therefore, the left-hand side of the Boltzmann equation is reduced to partial derivative of the distribution function with respect to time. The right-hand side contains collisional integrals, representing interactions between electrons, positrons and photons.

Differential probability for all processes per unit time and unit volume [90]
is defined as
\[ dw = c (2\pi \hbar)^4 \delta^{(4)} (p_f - p_i) \left| M_{fi} \right|^2 V \times \left[ \prod_b \frac{\hbar c}{2\epsilon_b V} \right] \left[ \prod_a \frac{d\epsilon_a'}{(2\pi \hbar)^3 2\epsilon_a'} \right], \]

where \( p_a' \) and \( \epsilon_a' \) are respectively momenta and energies of outgoing particles, \( \epsilon_b \) are energies of particles before interaction, \( M_{fi} \) are the corresponding matrix elements, \( \delta^{(4)} \) stands for energy-momentum conservation, \( V \) is the normalization volume. The matrix elements are related to the scattering amplitudes by
\[ M_{fi} = \left[ \prod_b \frac{\hbar c}{2\epsilon_b V} \right] \left[ \prod_a \frac{\hbar c}{2\epsilon_a' V} \right] T_{fi}. \]

As example consider absorption coefficient for Compton scattering which is given by
\[ \chi_{\text{fe}}^{\pm \rightarrow e^\pm} f_\gamma = \int d\epsilon' \frac{d\epsilon'}{d\epsilon'} w_{k',p',k,p} f_\gamma (k,t) f_\pm (p,t), \]

where \( p \) and \( k \) are momenta of electron (positron) and photon respectively, \( d\epsilon' = d\epsilon_\gamma^2 / \beta c^3 \), \( d\epsilon' = d\epsilon_\gamma^2 / \beta c^3 \) and the differential probability \( w_{k',p',k,p} \) is given by (10.5.3).

In (10.1.7) one can perform one integration over \( d\epsilon' \)
\[ \int d\epsilon' \delta (k + p - k' - p') \to 1, \]

but it is necessary to take into account the momentum conservation in the next integration over \( d\epsilon' \), so
\[ \int d\epsilon_\gamma' \delta (\epsilon_\gamma + \epsilon_\pm - \epsilon_\gamma' - \epsilon_\pm') = \int d(\epsilon_\gamma' + \epsilon_\pm') \frac{1}{|\partial(\epsilon_\gamma' + \epsilon_\pm') / \partial \epsilon_\gamma'|} \delta (\epsilon_\gamma + \epsilon_\pm - \epsilon_\gamma' - \epsilon_\pm') \to \frac{1}{|\partial(\epsilon_\gamma' + \epsilon_\pm') / \partial \epsilon_\gamma'|} \equiv J_{cs}, \]

where the Jacobian of the transformation is
\[ J_{cs} = \frac{1}{1 - \beta_+ b'_+ \cdot b'_-}, \]

where \( b_i = p_i / p, b'_i = p'_i / p', b'_\pm = (\beta_\pm \epsilon_\pm b_\pm + \epsilon_\gamma b_\gamma - \epsilon_\gamma' b'_\gamma) / (\beta_\pm \epsilon_\pm'). \]
Finally, for the absorption coefficient

$$\chi_{cs} f_\gamma = - \int d\epsilon'_1 d\epsilon'_2 |M_{fi}|^2 \frac{\hbar^2 c^2}{16\epsilon_+ \epsilon'_+ \epsilon'_-} J_{cs, f_\gamma}(k, t) f_{\pm}(p, t),$$

(10.1.11)

where the matrix element here is dimensionless. This integral is evaluated numerically.

For all binary interactions exact QED matrix elements are used which can be found in the standard textbooks, e.g. in [90], [322, 505], and are given below.

In order to account for charge screening in Coulomb scattering the minimal scattering angles are introduced following [506]. This allows to apply the same scheme for the computation of emission and absorption coefficients even for Coulomb scattering, while many treatments in the literature use Fokker-Planck approximation [507].

For such a dense plasma collisional integrals in (10.1.4) should include not only binary interactions, having order $\alpha^2$ in Feynman diagrams, but also triple ones, having order $\alpha^3$ [90]. Consider relativistic bremsstrahlung

$$e_1 + e_2 \leftrightarrow e'_1 + e'_2 + \gamma'.$$

(10.1.12)

For the time derivative, for instance, of the distribution function $f_2$ in the direct and in the inverse reactions (10.1.12) one has

$$f_2 = \int dp'_1 dp'_2 dp''_2 dk' \left[W_{P'_1, P'_2, k'; P'_2, P'_2, f'_1 f'_2 f''_k} - W_{P_2, P'_1, P'_2, f_1 f_2} \right],$$

(10.1.13)

$$= \frac{\epsilon^3 \hbar^3}{(2\pi)^2} \frac{\delta^{(4)}(P_f - P_i) |M_{fi}|^2}{2^4 \epsilon_1 \epsilon_2 \epsilon'_1 \epsilon'_2 \epsilon'_\gamma} \left[f'_1 f'_2 f''_k - \frac{1}{(2\pi \hbar)^3} f_1 f_2 \right],$$

where

$$dp'_1 dp'_2 W_{P'_1, P'_2, k'; P_2, P'_2} \equiv V^2 dw_1,$$

$$dp'_1 dp'_2 dk' W_{P_2, P'_1, P'_2, k'} \equiv V dw_2,$$

and $dw_1$ and $dw_2$ are differential probabilities given by (10.1.5). The matrix element here has dimensions of the length squared.

In the case of the distribution functions (10.1.18), see below, there are multipliers proportional to $\exp \frac{\varphi}{k_B T}$ in front of the integrals, where $\varphi$ are chemical potentials. The calculation of emission and absorption coefficients is then reduced to the well known thermal equilibrium case [497]. In fact, since reaction rates of triple interactions are $\alpha$ times smaller than binary reaction rates, it is expected that binary reactions come to detailed balance first. Only when binary reactions are all balanced, triple interactions become important. In addition, when binary reactions come into balance, distribution functions al-
10.1 Qualitative description of the pair plasma

<table>
<thead>
<tr>
<th>Binary interactions</th>
<th>Radiative and pair producing variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reactions with pairs</td>
<td></td>
</tr>
<tr>
<td><strong>Møller and Bhabha scattering</strong></td>
<td><strong>Bremsstrahlung</strong></td>
</tr>
<tr>
<td>(e^+_1 e^-_2 \rightarrow e^+_1' e^-_2')</td>
<td>(e^+_1 e^-_2 \rightarrow e^+_1' e^-_2' \gamma)</td>
</tr>
<tr>
<td>(e^+ e^- \rightarrow e^+ e^-)</td>
<td>(e^+ e^- \rightarrow e^+ e^- \gamma)</td>
</tr>
<tr>
<td><strong>Single Compton scattering</strong></td>
<td><strong>Double Compton scattering</strong></td>
</tr>
<tr>
<td>(e^+ \gamma \rightarrow e^+ \gamma')</td>
<td>(e^+ \gamma \leftrightarrow e^+ \gamma' \gamma'')</td>
</tr>
<tr>
<td><strong>Pair production and annihilation</strong></td>
<td><strong>Radiative pair production and three photon annihilation</strong></td>
</tr>
<tr>
<td>(\gamma \gamma' \leftrightarrow e^+ e^-)</td>
<td>(\gamma \gamma' \leftrightarrow e^+ e^- \gamma'')</td>
</tr>
<tr>
<td>(e^+ e^- \leftrightarrow \gamma \gamma)</td>
<td>(e^+ e^- \leftrightarrow \gamma \gamma' \gamma'')</td>
</tr>
<tr>
<td>(e^+ e^- \leftrightarrow e^+ e^- \gamma)</td>
<td>(e^+ e^- \leftrightarrow e^+ e^- \gamma \gamma)</td>
</tr>
</tbody>
</table>

Table 10.1: Microphysical processes in the pair plasma.

ready acquire the form (10.1.18). Although there is no principle difficulty in computations using exact matrix elements for triple reactions as well, the simplified scheme allows for much faster numerical computation.

All possible binary and triple interactions between electrons, positrons and photons are considered as summarized in Tab. 10.1.

Each of the above mentioned reactions is characterized by the corresponding timescale and optical depth. For Compton scattering of a photon, for instance

\[
  t_{cs} = \frac{1}{\sigma_T n \pm c}, \quad \tau_{cs} = \sigma_T n \pm R, \tag{10.1.14}
\]

where \(\sigma_T = \frac{8\pi}{3} \alpha^2 \left( \frac{\hbar}{m_e c} \right)^2\) is the Thomson cross-section. There are two timescales in the problem that characterize the condition of detailed balance between direct and inverse reactions, \(t_{cs}\) for binary and \(\alpha^{-1} t_{cs}\) for triple interactions respectively.

Notice, that electron–positron pair can annihilate into neutrino channel with the main contribution from the reaction \(e^\pm e^\mp \rightarrow \nu \bar{\nu}\). By this process the energy could leak out from the plasma if it is transparent for neutrinos. The optical depth and energy loss for this process can be estimated following [508] by using Fermi theory, see also [509, 510] for calculations within electroweak theory.

The optical depth is given by (10.1.14) with the cross-section

\[
  \sigma_{\nu \bar{\nu}} \sim \frac{g^2}{\pi} \left( \frac{\hbar}{m_e c} \right)^2, \tag{10.1.15}
\]

where \(g \approx 10^{-12}\) is the weak interaction coupling constant and it is assumed that typical energies of electron and positron are \(\sim m_e c^2\) and their relative
velocities $v \sim c$. Numerically $\sigma_{v^2} / \sigma_T = \frac{3}{8\pi^2} (g/\alpha)^2 \simeq 7 \times 10^{-22}$. For astrophysical sources the plasma may be both transparent and opaque to neutrino production. The energy loss when pairs are relativistic and nondegenerate is

$$\frac{d\rho}{dt} = \frac{128g^2}{\pi^3} \zeta(5)\zeta(4)\theta^9 m_e c^2 \left( \frac{m_e c}{\hbar} \right)^3 \left( \frac{m_e c^2}{\hbar} \right).$$  

(10.1.16)

The ratio between the energy lost due to neutrinos and the energy of photons in thermal equilibrium is then

$$\frac{1}{\rho} \frac{d\rho}{dt} \Delta t = \frac{128g^2}{\pi^3} \zeta(5)\zeta(4)\theta^5 \left( \frac{m_e c^2}{\hbar} \right) \Delta t \simeq 3.6 \times 10^{-3} \theta^5 \Delta t \text{ sec.}$$  

(10.1.17)

For astrophysical sources with the dynamical time $\Delta t \sim 10^{-3}$ sec, the energy loss due to neutrinos becomes relevant [511] for high temperatures $\theta > 10$. However, on the timescale of relaxation to thermal equilibrium $\Delta t \sim 10^{-12}$ sec the energy loss is negligible.

Starting from arbitrary distribution functions a common development is found: at the time $t_{cs}$ the distribution functions always have evolved in a functional form on the entire energy range, and depend only on two parameters. In fact it is found for the distribution functions

$$f_i(\epsilon) = \frac{2}{(2\pi\hbar)^3} \exp \left( -\frac{\epsilon - v_i}{\theta_i} \right),$$  

(10.1.18)

with chemical potential $v_i \equiv \frac{\varphi_i}{m_e c^2}$ and temperature $\theta_i \equiv \frac{k_B T_i}{m_e c^2}$, where $\epsilon \equiv \frac{\epsilon}{m_e c^2}$ is the energy of the particle. Such a configuration corresponds to a kinetic equilibrium [76, 475, 507] in which all particles acquire a common temperature and nonzero chemical potentials. Triple interactions become essential for $t > t_{cs}$, after the establishment of kinetic equilibrium. In strict mathematical sense the sufficient condition for reaching thermal equilibrium is when all direct reactions are exactly balanced with their inverse. Therefore, in principle, not only triple, but also four-particle, five-particle and so on reaction have to be accounted for in equation (10.1.4). The timescale for reaching thermal equilibrium will be then determined by the slowest reaction which is not balanced with its inverse. The necessary condition here is the detailed balance at least in triple interactions, since binary reactions do not change chemical potentials at all.

Notice that similar method to ours was applied in [507] in order to compute spectra of particles in kinetic equilibrium. However, it was never shown how particles evolve down to thermal equilibrium.

In the case of pure pair plasma chemical potentials in (10.1.18) represent
deviations from the thermal equilibrium through the relation

$$\nu = \theta \ln (n / n_{th}),$$  \hspace{2cm} (10.1.19)

where $n_{th}$ are concentrations of particles in thermal equilibrium.

### 10.2 The discretization procedure and the computational scheme

In order to solve equations (10.1.4) a finite difference method is used by introducing a computational grid in the phase space to represent the distribution functions and to compute collisional integrals following [512]. The goal is to construct the scheme implementing energy, baryon number and electric charge conservation laws. For this reason instead of distribution functions $f_i$, spectral energy densities are used

$$E_i(\epsilon_i) = \frac{4\pi\epsilon_i^3\beta_i f_i}{c^3}, \hspace{2cm} (10.2.1)$$

where $\beta_i = \sqrt{1 - (m_i c^2 / \epsilon_i)^2}$, in the phase space $\epsilon_i$. Then

$$\epsilon_i f_i(p, t) d\tau dp = \frac{4\pi\epsilon_i^3\beta_i f_i}{c^3} r d\epsilon_i = E_i d\tau d\epsilon_i \hspace{2cm} (10.2.2)$$

is the energy in the volume of the phase space $d\tau dp$. The particle density is

$$n_i = \int f_i d\mathbf{p} = \int \frac{E_i}{\epsilon_i} d\epsilon_i, \quad dn_i = f_i d\mathbf{p}, \hspace{2cm} (10.2.3)$$

while the corresponding energy density is

$$\rho_i = \int \epsilon_i f_i d\mathbf{p} = \int E_i d\epsilon_i.$$  

Boltzmann equations (10.1.4) can be rewritten in the form

$$\frac{1}{c} \frac{\partial E_i}{\partial t} = \sum_q (\eta_i^q - \chi_i^q E_i), \hspace{2cm} (10.2.4)$$

where $\eta_i^q = (4\pi\epsilon_i^3\beta_i / c^3) \eta_i^q$.

The computational grid for phase space is \{\(\epsilon_i, \mu, \phi\)}, where $\mu = \cos \theta$, $\theta$ and $\phi$ are angles between radius vector $r$ and the particle momentum $p$. The zone boundaries are $\epsilon_i, \mu_{\pm1/2}, \mu_{k\pm1/2}, \phi_{l\pm1/2}$ for $1 \leq \omega \leq \omega_{\max}$, $1 \leq k \leq k_{\max}$, $1 \leq l \leq l_{\max}$. The length of the $i$th interval is $\Delta \epsilon_i, \omega \equiv \epsilon_{i, \omega+1/2} - \epsilon_{i, \omega-1/2}$. On
the finite grid the functions (10.2.1) become

\[ E_{i,\omega} \equiv \frac{1}{\Delta \epsilon_{i,\omega}} \int_{\Delta \epsilon_{i,\omega}} d\epsilon E_i(\epsilon). \]  

(10.2.5)

Now the collisional integrals in (10.2.4) are replaced by the corresponding sums.

After this procedure the set of ordinary differential equations (ODE’s) is obtained, instead of the system of partial differential equations for the quantities \( E_{i,\omega} \) to be solved. There are several characteristic times for different processes in the problem, and therefore the system of differential equations is stiff. (Eigenvalues of Jacobi matrix differs significantly, and the real parts of eigenvalues are negative.) Gear’s method [513] is used to integrate ODE’s numerically. This high order implicit method was developed for the solution of stiff ODE’s.

In this method exact energy conservation law is satisfied. For binary interactions the particles number conservation law is satisfied as interpolation of grid functions \( E_{i,\omega} \) inside the energy intervals is adopted.

### 10.3 Conservation laws

Conservation laws consist of charge and energy conservations. In addition, in binary reactions particle number is conserved.

Energy conservation law can be rewritten for the spectral density

\[ \frac{d}{dt} \sum_i \rho_i = 0, \quad \text{or} \quad \frac{d}{dt} \sum_{i,\omega} Y_{i,\omega} = 0, \]  

(10.3.1)

where

\[ Y_{i,\omega} = \int_{\epsilon_{i,\omega}-\Delta \epsilon_{i,\omega}/2}^{\epsilon_{i,\omega}+\Delta \epsilon_{i,\omega}/2} E_i d\epsilon. \]  

(10.3.2)

Particle’s conservation law in binary reactions reduces to

\[ \frac{d}{dt} \sum_i n_i = 0, \quad \text{or} \quad \frac{d}{dt} \sum_{i,\omega} \frac{Y_{i,\omega}}{\epsilon_{i,\omega}} = 0. \]  

(10.3.3)

For electrically neutral plasma considered in this Section charge conservation implies

\[ n_- = n_+. \]  

(10.3.4)
10.4 Determination of temperature and chemical potentials in kinetic equilibrium

Consider distribution functions for photons and pairs in the most general form (10.1.18). If one supposes that reaction rate for the Bhabha scattering vanishes, i.e. there is equilibrium with respect to reaction

\[ e^+ + e^- \leftrightarrow +e^{+'} + e^{-'}, \] (10.4.1)

and the corresponding condition can be written in the following way

\[ f_+(1 - f_+')(1 - f_-') = f_+(1 - f_+)f_-'(1 + f_-), \] (10.4.2)

where Bose-Einstein enhancement along with Pauli blocking factors are taken into account, it can be shown that electrons and positrons have the same temperature

\[ \theta_+ = \theta_\equiv \theta_\pm, \] (10.4.3)

and they have arbitrary chemical potentials.

With (10.4.3) analogous consideration for the Compton scattering

\[ e^\pm + \gamma \leftrightarrow +e^{\pm'} + \gamma', \] (10.4.4)

gives

\[ f_\pm(1 - f_\pm')(1 + f_\gamma') = f_\pm'(1 - f_\pm)f_\gamma'(1 + f_\gamma), \] (10.4.5)

and leads to the same temperature of pairs and photons

\[ \theta_\pm = \theta_\gamma \equiv \theta_k, \] (10.4.6)

with arbitrary chemical potentials. If, in addition, reaction rate in the pair creation and annihilation process

\[ e^\pm + e^\mp \leftrightarrow \gamma + \gamma' \] (10.4.7)

vanishes too, i.e. there is equilibrium with respect to pair production and annihilation, with the corresponding condition,

\[ f_+f_- (1 + f_\gamma)(1 + f_\gamma') = f_\gamma f_\gamma'(1 - f_+)(1 - f_-), \] (10.4.8)

it turns out that also chemical potentials for pairs and photons satisfy the following condition for the chemical potentials

\[ \nu_+ + \nu_- = 2\nu_\gamma. \] (10.4.9)

However, since in general \( \nu_\gamma \neq 0 \) the condition (10.4.9) does not imply \( \nu_+ = \nu_- \).
Interaction & Parameters of DFs
---
I $e^+e^-$ scattering & $\theta_+ = \theta_-, \forall \nu_+\nu_-$
II $e^\pm\gamma$ scattering & $\theta_\gamma = \theta_\pm, \forall \nu_\gamma\nu_\pm$
III pair production & $\nu_+ + \nu_- = 2\nu_\gamma$, if $\theta_\gamma = \theta_\pm$
IV Tripe interactions & $\nu_\gamma, \nu_\pm = 0$, if $\theta_\gamma = \theta_\pm$

**Table 10.2:** Relations between parameters of equilibrium DFs fulfilling detailed balance conditions for the reactions shown in Tab. 10.1.

In general, the detailed balance conditions for different reactions lead to relations between temperatures and chemical potentials summarized in table 10.2.

Kinetic equilibrium is first established simultaneously for electrons, positrons and photons. Thus they reach the same temperature, but with chemical potentials different from zero. Later on, protons reach the same temperature.

In order to find temperatures and chemical potentials the following constraints are implemented: energy conservation (10.3.1), particle number conservation (10.3.3), charge conservation (10.3.4), condition for the chemical potentials (10.4.9).

Given (10.1.18) it is found for photons

$$\frac{\rho_\gamma}{n_\gamma m_\gamma c^2} = 3\theta_\gamma, \quad n_\gamma = \frac{1}{V_0} \exp\left(\frac{\nu_\gamma}{\theta_\gamma}\right) 2\theta_\gamma^3,$$

and for pairs

$$\frac{\rho_\pm}{n_\pm m_\gamma c^2} = j_2(\theta_\pm), \quad n_\pm = \frac{1}{V_0} \exp\left(\frac{\nu_\pm}{\theta_\pm}\right) j_1(\theta_\pm),$$

where the Compton volume is

$$V_0 = \frac{1}{8\pi} \left(\frac{2\pi\hbar}{m_\gamma c}\right)^3$$

and functions $j_1$ and $j_2$ are defined as

$$j_1(\theta) = \theta K_2(\theta^{-1}) \rightarrow \left\{ \begin{array}{ll} \sqrt{\frac{\pi}{2}} e^{\frac{1}{2}\theta^{3/2}}, & \theta \rightarrow 0 \\ 2\theta^3, & \theta \rightarrow \infty \end{array} \right.,$$

$$j_2(\theta) = \frac{3K_3(\theta^{-1}) + K_1(\theta^{-1})}{4K_2(\theta^{-1})} \rightarrow \left\{ \begin{array}{ll} 1 + \frac{8\theta^2}{27}, & \theta \rightarrow 0 \\ 3\theta, & \theta \rightarrow \infty \end{array} \right..$$

For pure electron–positron-photon plasma in kinetic equilibrium, summing up energy densities in (10.4.10),(10.4.11) and using (10.4.3),(10.4.6) and (10.4.9)
it is found
\[ \sum_{e^+, e^-, \gamma} \rho_i = \frac{2m_e c^2}{V_0} \exp \left( \frac{\nu_k}{\theta_k} \right) \left[ 3\theta^4 + j_1(\theta_k)j_2(\theta_k) \right], \quad (10.4.15) \]
and analogously for number densities
\[ \sum_{e^+, e^-, \gamma} n_i = \frac{2}{V_0} \exp \left( \frac{\nu_k}{\theta_k} \right) \left[ \theta^3 + j_1(\theta_k) \right]. \quad (10.4.16) \]
Therefore, two unknowns, \( \nu_k \) and \( \theta_k \) can be found.
In thermal equilibrium \( \nu_\gamma \) vanishes and one has
\[ \nu_+ = \nu_- = 0. \quad (10.4.17) \]

**10.5 Binary interactions**

In this section the expressions for emission and absorption coefficients in Compton scattering, pair creation and annihilation with two photons, Møller and Bhabha scattering are obtained.

**10.5.1 Compton scattering**

The time evolution of the distribution functions of photons and pair particles due to Compton scattering may be described by [514], [515]
\[ \left( \frac{\partial f_\gamma(k, t)}{\partial t} \right)_{\gamma e^{\pm} \rightarrow \gamma' e^{\pm}} = \int dk'dp dp' V w_{k', p', k, p} \times \left[ f_\gamma(k', t)f_{\pm}(p', t) - f_\gamma(k, t)f_{\pm}(p, t) \right], \quad (10.5.1) \]
\[ \left( \frac{\partial f_{\pm}(p, t)}{\partial t} \right)_{\gamma e^{\pm} \rightarrow \gamma' e^{\pm}} = \int dk dk' dp' V w_{k', p', k, p} \times \left[ f_\gamma(k', t)f_{\pm}(p', t) - f_\gamma(k, t)f_{\pm}(p, t) \right], \quad (10.5.2) \]
where
\[ w_{k', p', k, p} = \frac{\hbar^2 c^6}{(2\pi)^2 V} \delta(e_\gamma - e_\pm - e'_\gamma - e'_\pm) \delta(k + p - k' - p') \frac{|M_{fi}|^2}{16e_\gamma e_\pm e'_\gamma e'_\pm}, \quad (10.5.3) \]
is the probability of the process,
\[ |M_{fi}|^2 = 2^6 \pi^2 \alpha^2 \left[ \frac{m_e^2 c^2}{s - m_e^2 c^2} + \frac{m_e^2 c^2}{u - m_e^2 c^2} + \left( \frac{m_e^2 c^2}{s - m_e^2 c^2} + \frac{m_e^2 c^2}{u - m_e^2 c^2} \right)^2 \right] - \frac{1}{4} \left( \frac{s - m_e^2 c^2}{u - m_e^2 c^2} + \frac{u - m_e^2 c^2}{s - m_e^2 c^2} \right), \] 
(10.5.4)
is the square of the matrix element, \( s = (p + t)^2 \) and \( u = (p - t')^2 \) are invariants, \( t = (e_\gamma / c)(1, e_\gamma) \) and \( p = (e_\pm / c)(1, \beta_\pm e_\pm) \) are energy-momentum four vectors of photons and electrons, respectively, \( dp = d\epsilon_\pm do_\pm \beta_\pm / c^2, \) \( dk' = de_\gamma' do_\gamma' / c^3 \) and \( do = d\mu d\phi. \)

The energies of photon and positron (electron) after the scattering are
\[ \epsilon_\gamma' = \frac{\epsilon_\pm e_\gamma (1 - \beta_\pm b_\pm \cdot b_\gamma)}{\epsilon_\pm (1 - \beta_\pm b_\pm \cdot b_\gamma') + \epsilon_\gamma (1 - b_\gamma \cdot b_\gamma')}, \quad \epsilon_\gamma' = \epsilon_\pm + \epsilon_\gamma - \epsilon_\gamma', \] 
(10.5.5)
\[ b_i = p_i / p, \ b_i' = p_i' / p', \ b_\pm = (\beta_\pm e_\pm b_\gamma - \epsilon_\gamma' b_\gamma') / (\beta_\pm'^\epsilon). \]

For photons, the absorption coefficient (10.1.11) in the Boltzmann equations (10.1.4) is
\[ \chi_\gamma^{\epsilon \pm \rightarrow \epsilon' \pm} f_\gamma = -\frac{1}{c} \left( \frac{\partial f_\gamma}{\partial t} \right)_{\gamma \epsilon \pm \rightarrow \gamma' \epsilon'}^{abs} = \int dn_\pm do_\gamma' J_{cs} \frac{e_\gamma' |M_{fi}|^2 \hbar^2 c^2}{16\epsilon_\gamma e_\gamma e_\gamma'} f_\gamma, \] 
(10.5.6)
where \( dn_i = d\epsilon_i do_i e_i^2 \beta_i f_i / c^3 = d\epsilon_i do_i E_i / (2\pi \epsilon_i). \)

From equations (10.5.1) and (10.5.6), the absorption coefficient for photon energy density \( E_\gamma \) averaged over the \( \epsilon, \mu \)-grid with zone numbers \( \omega \) and \( k \) is
\[ (\chi E)^{\gamma \epsilon \pm \rightarrow \gamma' \epsilon'} = \frac{1}{\Delta \epsilon_\gamma, \omega} \int_{\epsilon_\gamma \in \Delta \epsilon_\gamma, \omega} d\epsilon_\gamma d\mu_\gamma (\chi E)^{\gamma \epsilon \pm \rightarrow \gamma' \epsilon'} = \frac{1}{\Delta \epsilon_\gamma, \omega} \int_{\epsilon_\gamma \in \Delta \epsilon_\gamma, \omega} dn_\gamma do_\gamma' J_{cs} \frac{e_\gamma' |M_{fi}|^2 \hbar^2 c^2}{16\epsilon_\gamma e_\gamma e_\gamma'}, \] 
(10.5.7)
where the Jacobian of the transformation is
\[ J_{cs} = \frac{e_\gamma' e_\gamma'}{\epsilon_\gamma e_\gamma (1 - \beta_\pm b_\gamma \cdot b_\pm)}. \] 
(10.5.8)

Similar integrations can be performed for the other terms of equations (10.5.1),
10.5 Binary interactions

(10.5.2), and

\[ \eta_{\gamma\omega}^{\gamma'\omega'} = \frac{1}{\Delta e_{\gamma\omega}} \int_{\epsilon'_{\gamma} \in \Delta e_{\gamma\omega}} d\epsilon'_{\gamma} \prod_{\pm, \omega} e_{\gamma}^{2} |M_{fi}|^{2} \hbar^{2} c^{2} 16e_{\epsilon_{\gamma}} \epsilon'_{\gamma}, \]  

(10.5.9)

\[ \eta_{\pm, \omega}^{\gamma'\omega'} = \frac{1}{\Delta e_{\pm, \omega}} \int_{\epsilon'_{\pm} \in \Delta e_{\pm, \omega}} d\epsilon'_{\pm} \prod_{\pm, \omega} e_{\gamma}^{2} |M_{fi}|^{2} \hbar^{2} c^{2} 16e_{\epsilon_{\gamma}} \epsilon'_{\gamma}, \]  

(10.5.10)

\[ (\chi E)^{\gamma'\omega'} \prod_{\pm, \omega} e_{\gamma}^{2} |M_{fi}|^{2} \hbar^{2} c^{2} \]  

(10.5.11)

In order to perform integrals (10.5.7)-(10.5.11) numerically over \( \phi \) \( 0 \leq \phi \leq 2\pi \) a uniform grid \( \phi_{i+1/2} \) is introduced with \( 1 \leq l \leq l_{\text{max}} \) and \( \Delta \phi_{i} = (\phi_{i+1/2} - \phi_{i-1/2}) / 2 = 2\pi / l_{\text{max}} \). It is assumed that any function of \( \phi \) in equations (10.5.7)-(10.5.11) in the interval \( \Delta \phi_{j} \) is equal to its value at \( \phi = \phi_{j} = (\phi_{j-1/2} + \phi_{j+1/2}) / 2 \). It is necessary to integrate over \( \phi \) only once at the beginning of calculations. The number of intervals of the \( \phi \)-grid depends on the average energy of particles and is typically taken as \( l_{\text{max}} = 2k_{\text{max}} = 64 \).

10.5.2 Pair creation and annihilation

The rates of change of the distribution function due to pair creation and annihilation are

\[ \left( \frac{\partial f_{\gamma_{i}}(k_{i}, t)}{\partial t} \right)_{\gamma_{1} \gamma_{2} \rightarrow \gamma' \omega} = - \int d\mathbf{p}_{-} d\mathbf{p}_{+} V w_{p_{-}p_{+}k_{i}, k_{2}f_{\gamma_{1}}(k_{i}, t)f_{\gamma_{2}}(k_{2}, t)} \]  

(10.5.12)

\[ \left( \frac{\partial f_{\gamma_{i}}(k_{i}, t)}{\partial t} \right)_{\gamma \rightarrow \gamma' \omega} = \int d\mathbf{p}_{-} d\mathbf{p}_{+} V w_{k_{1}k_{2}p_{-}p_{+}f_{-}(p_{-}, t)f_{+}(p_{+}, t)} \]  

(10.5.13)

for \( i = 1, j = 2 \), and for \( j = 1, i = 2 \).

\[ \left( \frac{\partial f_{\pm}(p_{\pm}, t)}{\partial t} \right)_{\gamma_{1}\gamma_{2} \rightarrow \gamma' \omega} = \int d\mathbf{p}_{\pm} d\mathbf{k}_{1} d\mathbf{k}_{2} V w_{p_{-}p_{+}k_{i}k_{2}f_{\gamma}(k_{1}, t)f_{\gamma}(k_{2}, t)} \]  

(10.5.14)

\[ \left( \frac{\partial f_{\pm}(p_{\pm}, t)}{\partial t} \right)_{\gamma \rightarrow \gamma' \omega} = - \int d\mathbf{p}_{\pm} d\mathbf{k}_{1} d\mathbf{k}_{2} V w_{k_{1}k_{2}p_{-}p_{+}f_{-}(p_{-}, t)f_{+}(p_{+}, t)} \]  

(10.5.15)

where

\[ w_{p_{-}p_{+}k_{1}k_{2}} = \frac{\hbar^{2} c^{6}}{(2\pi)^{2} V} \delta(\epsilon_{-} + \epsilon_{+} - \epsilon_{1} - \epsilon_{2}) \delta(p_{-} + p_{+} - k_{1} - k_{2}) |M_{fi}|^{2} \]  

(10.5.16)
Here, the matrix element $|M_{fi}|^2$ is given by equation (10.5.4) with the new invariants $s = (p_--t_1)^2$ and $u = (p_--t_2)^2$, see [90].

The energies of photons created via annihilation of a $e^\pm$ pair are

$$e_1(b_1) = \frac{m^2c^4 + \epsilon_-\epsilon_- (1 - \beta_-\beta_+ b_-\cdot b_+)}{\epsilon_- (1 - \beta_-\beta_+ b_-\cdot b_1) + \epsilon_+(1 - \beta_+\beta_- b_+\cdot b_1)}, \quad e_2(b_1) = \epsilon_- + \epsilon_+ - \epsilon_1,$$

(10.5.17)

while the energies of pair particles created by two photons are found from

$$\epsilon_- (b_-) = \frac{B + \sqrt{B^2 - AC}}{A}, \quad \epsilon_+ (b_-) = \epsilon_1 + \epsilon_2 - \epsilon_-,$$

(10.5.18)

where $A = (\epsilon_1 + \epsilon_2)^2 - [(\epsilon_1 b_1 + \epsilon_2 b_2)\cdot b_-]^2$, $B = (\epsilon_1 + \epsilon_2)\epsilon_1\epsilon_2 (1 - b_1\cdot b_2)$, $C = m^2c^4[(\epsilon_1 b_1 + \epsilon_2 b_2)\cdot b_-]^2 + \epsilon_1^2\epsilon_2^2 (1 - b_1\cdot b_2)^2$. Only one root in equation (10.5.18) has to be chosen. From energy-momentum conservation

$$t_1 + t_2 - p_- = p_+,$$

(10.5.19)

taking square from the energy part leads to

$$\epsilon_1^2 + \epsilon_2^2 + \epsilon_-^2 + 2\epsilon_1\epsilon_2 - 2\epsilon_1\epsilon_- - 2\epsilon_2\epsilon_- = \epsilon_+^2,$$

(10.5.20)

and taking square from the momentum part

$$\epsilon_1\epsilon_2 (1 - b_1\cdot b_2) - \epsilon_1\epsilon_- (1 - \beta_- b_-\cdot b_1) - \epsilon_2\epsilon_- (1 - \beta_+ b_+\cdot b_-) = 0,$$

(10.5.22)

$$\epsilon_- \epsilon_- (b_1 + b_2)\cdot b_- = \epsilon_-(\epsilon_1 + \epsilon_2) - \epsilon_1\epsilon_2 (1 - b_1\cdot b_2).$$

Eliminating $\beta$ it is obtained

$$\epsilon_1^2\epsilon_2^2 (1 - b_1\cdot b_2)^2 - 2\epsilon_1\epsilon_2 (1 - b_1\cdot b_2)(\epsilon_1 + \epsilon_2)\epsilon_- +$$

$$+ \left\{ (\epsilon_1 + \epsilon_2)^2 - [(\epsilon_1 b_1 + \epsilon_2 b_2)\cdot b_-]^2 \right\} \epsilon_-^2 =$$

$$= [(\epsilon_1 b_1 + \epsilon_2 b_2)\cdot b_-](-m^2),$$

(10.5.23)

Therefore, the condition to be checked reads

$$\epsilon_- \epsilon_- [(\epsilon_1 b_1 + \epsilon_2 b_2)\cdot b_-]^2 = [\epsilon_- (\epsilon_1 + \epsilon_2) - (\epsilon_1\epsilon_2) (1 - b_1\cdot b_2)] \times$$

$$\times [(\epsilon_1 b_1 + \epsilon_2 b_2)\cdot b_-] \geq 0.$$

(10.5.24)
Finally, integration of equations (10.5.12)-(10.5.15) yields

\[
\eta_{\gamma \omega}^{e^+e^- \rightarrow \gamma \gamma_2} = \frac{1}{\Delta e_{\gamma \omega}} \left( \int_{e_1 \in \Delta e_{\gamma \omega}} d^2n_\pm J_{ca} \frac{e_1^2 |M_{fi}|^2 \hbar^2 c^2}{16 e_- e_+ e_2} \right) + \\
+ \frac{1}{\Delta e_{\gamma \omega}} \left( \int_{e_2 \in \Delta e_{\gamma \omega}} d^2n_\pm J_{ca} \frac{e_1 |M_{fi}|^2 \hbar^2 c^2}{16 e_- e_+ e_2} \right), \tag{10.5.25}
\]

\[
(\chi E)_{\gamma \omega}^{e^+e^- \rightarrow \gamma \gamma_2} = \frac{1}{\Delta e_{\gamma \omega}} \left( \int_{e_- \in \Delta e_{\gamma \omega}} d^2n_\pm J_{ca} \frac{e_1 |M_{fi}|^2 \hbar^2 c^2}{16 e_+ e_- e_2} \right) + \\
+ \frac{1}{\Delta e_{\gamma \omega}} \left( \int_{e_+ \in \Delta e_{\gamma \omega}} d^2n_\pm J_{ca} \frac{e_1 |M_{fi}|^2 \hbar^2 c^2}{16 e_+ e_- e_2} \right), \tag{10.5.26}
\]

\[
(\chi E)_{\gamma \omega}^{\gamma \gamma_2 \rightarrow e^- e^+} = \frac{1}{\Delta e_{\gamma \omega}} \left( \int_{e_1 \in \Delta e_{\gamma \omega}} d^2n_\gamma J_{ca} \frac{e_1 |M_{fi}|^2 \hbar^2 c^2}{16 e_+ e_+ e_2} \right) + \\
+ \frac{1}{\Delta e_{\gamma \omega}} \left( \int_{e_2 \in \Delta e_{\gamma \omega}} d^2n_\gamma J_{ca} \frac{e_1 |M_{fi}|^2 \hbar^2 c^2}{16 e_+ e_+ e_2} \right), \tag{10.5.27}
\]

\[
\eta_{\gamma \omega}^{\gamma \gamma_2 \rightarrow e^- e^+} = \frac{1}{\Delta e_{\gamma \omega}} \left( \int_{e_- \in \Delta e_{\gamma \omega}} d^2n_\gamma J_{ca} \frac{e_1 |M_{fi}|^2 \hbar^2 c^2}{16 e_+ e_+ e_2} \right) + \\
+ \frac{1}{\Delta e_{\gamma \omega}} \left( \int_{e_+ \in \Delta e_{\gamma \omega}} d^2n_\gamma J_{ca} \frac{e_1 |M_{fi}|^2 \hbar^2 c^2}{16 e_+ e_+ e_2} \right), \tag{10.5.28}
\]

where \(d^2n_\pm = dn_- dn_+ do_1, d^2n_\gamma = dn_\gamma_1 dn_\gamma_2 do_-, dn_\pm = de_\pm do_\pm e_\pm \beta_\pm f_\pm,\)

\(dn_\gamma_1 = de_1 do_1 e_1^2 f_\gamma_1,\) and the Jacobian is

\[
J_{ca} = \frac{e_+ \beta_-}{(e_+ + e_-) \beta_- - (e_1 b_1 + e_2 b_2) \cdot b_-}. \tag{10.5.29}
\]

### 10.5.3 Møller scattering of electrons and positrons

The time evolution of the distribution functions of electrons (or positrons) is described by

\[
\left( \frac{\partial f_i(p_i, t)}{\partial t} \right)_{e_1 e_2 \rightarrow e_1' e_2'} = \int dp'_1 dp'_2 dp_2' V w_{p'_1 p'_2 p_1 p_2} \times \\
\times [f_1(p'_1, t) f_2(p'_2, t) - f_1(p_1, t) f_2(p_2, t)], \tag{10.5.30}
\]
with \( i = 1, j = 2 \), and with \( j = 1, i = 2 \), and where

\[
\omega_{p_1'p_2'p_1p_2} = \frac{\eta c^6}{(2\pi)^2 V} \delta(e_1 + e_2 - e_1' - e_2') \delta(p_1 + p_2 - p_1' - p_2') \frac{|M_{fi}|^2}{16\epsilon_1\epsilon_2 e_1' e_2'}, \tag{10.5.31}
\]

\[
|M_{fi}|^2 = 2^6\pi^2 \alpha^2 \left\{ \frac{1}{t^2} \left[ \frac{s^2 + u^2}{2} + 4m_c^2c^2(t - m_c^2c^2) \right] + \frac{1}{u^2} \left[ \frac{s^2 + t^2}{2} + 4m_c^2c^2(u - m_c^2c^2) \right] + \frac{4}{tu} \left( \frac{s}{2} - m_c^2c^2 \right) \left( \frac{s}{2} - 3m_c^2c^2 \right) \right\}, \tag{10.5.32}
\]

\[
with \ s = (p_1 + p_2)^2 = 2(m_c^2c^2 + p_1p_2), \ t = (p_1 - p_1')^2 = 2(m_c^2c^2 - p_1p_1'), \text{ and} \ u = (p_1 - p_2')^2 = 2(m_c^2c^2 - p_1p_2) \ [90].
\]

The energies of final state particles are given by (10.5.18) with new coefficients \( \hat{A} = (e_1 + e_2)^2 - (e_1\beta_1 b_1 \cdot b_1' + e_2\beta_2 b_2 \cdot b_2')^2, \ \hat{B} = (e_1 + e_2)[m_c^2c^4 + \epsilon_1\epsilon_2(1 - \beta_1\beta_2 b_1 \cdot b_2)], \text{ and} \ \hat{C} = m_c^2c^4(e_1\beta_1 b_1 \cdot b_1' + e_2\beta_2 b_2 \cdot b_2')^2 + [m_c^2c^4 + \epsilon_1\epsilon_2(1 - \beta_1\beta_2 b_1 \cdot b_2)]^2. \) The condition to be checked is

\[
\left[ e_1'(e_1 + e_2) - m_c^2c^4 - (e_1\epsilon_2)(1 - \beta_1\beta_2 b_1 \cdot b_2) \right] \left[ (e_1\beta_1 b_1 + e_2\beta_2 b_2) \cdot b_1' \right] \geq 0. \tag{10.5.35}
\]

Integration of equations (10.5.30), similar to the case of Compton scattering in Section 10.5.1 yields

\[
\eta_{\epsilon_1 \epsilon_2 \epsilon_1' \epsilon_2'}^{e_1 e_2 \rightarrow e_1' e_2'} = \frac{1}{\Delta \epsilon_{\epsilon, \omega}} \left( \int d^2n_{ms} \frac{e_1'^2 b_1' |M_{fi}|^2 h c^2}{16\epsilon_1 e_2 e_1' e_2'} \right) + \frac{1}{\Delta \epsilon_{\epsilon, \omega}} \left( \int d^2n_{ms} \frac{e_1'^2 b_1' |M_{fi}|^2 h c^2}{16\epsilon_1 e_2 e_1' e_2'} \right), \tag{10.5.36}
\]

\[
(\chi E)_{\epsilon_1 \epsilon_2 \epsilon_1' \epsilon_2'}^{e_1 e_2 \rightarrow e_1' e_2'} = \frac{1}{\Delta \epsilon_{\epsilon, \omega}} \left( \int d^2n_{ms} \frac{e_1'^2 b_1' |M_{fi}|^2 h c^2}{16\epsilon_1 e_2 e_1' e_2'} \right) + \frac{1}{\Delta \epsilon_{\epsilon, \omega}} \left( \int d^2n_{ms} \frac{e_1'^2 b_1' |M_{fi}|^2 h c^2}{16\epsilon_1 e_2 e_1' e_2'} \right), \tag{10.5.37}
\]

where \( d^2n = dn_1 dn_2 d\epsilon_1', \ dn_{1,2} = d\epsilon_1 d\epsilon_2 \epsilon_1'^2 \beta_1, f_{1,2}, \) and the Jacobian is

\[
f_{ms} = \frac{e_1'^2 b_1'}{(e_1' + e_2') \beta_1' - (e_1\beta_1 b_1 + e_2\beta_2 b_2) \cdot b_1'}. \tag{10.5.38}
\]
10.5.4 Bhaba scattering of electrons on positrons

The time evolution of the distribution functions of electrons and positrons due to Bhaba scattering is described by

\[
\frac{\partial f_{\pm}(p_{\pm}, t)}{\partial t} = \int dp_{\pm} dp'_{-} dp'_{+} V w_{p_{-}, p'_{-}, p'_{+}} \times
\]
\[
\left[ f_{-}(p'_{-}, t) f_{+}(p'_{+}, t) - f_{-}(p_{-}, t) f_{+}(p_{+}, t) \right],
\]
(10.5.39)

where

\[
w_{p'_{-}, p'_{+}} = \frac{\hbar^2 e^6}{(2\pi)^2 V} \delta(\epsilon_{-} + \epsilon_{+} - \epsilon'_{-} - \epsilon'_{+}) \delta(p_{-} + p'_{+} - p'_{-} - p_{+}) \left| M_{fi} \right|^2
\]
\[
\frac{16 e - e_{+} + e'_{-} e'_{+}},
\]
(10.5.40)

and \( |M_{fi}| \) is given by the equation (10.5.34), but the invariants are \( s = (p_{-} - p'_{+})^2 \), \( t = (p_{+} - p'_{-})^2 \) and \( u = (p_{-} + p_{+})^2 \). The final energies \( \epsilon'_{-}, \epsilon'_{+} \) are functions of the outgoing particle directions in a way similar to that in Section 10.5.3, see also [90].

Integration of equations (10.5.39) yields

\[
\eta^{e^{-}e^{+} \rightarrow e^{-}'e^{+}'} = \frac{1}{\Delta \epsilon_{\pm, \omega}} \left( \int_{\epsilon_{\omega} \Delta \epsilon_{\omega}} d^2 n'_{\pm} I_{bs} \frac{\epsilon'^{2} \beta'_{-} |M_{fi}|^2 \hbar^2 c^2}{16 e - e_{+} + e'_{-} e'_{+}} \right) +
\]
\[
+ \frac{1}{\Delta \epsilon_{\pm, \omega}} \left( \int_{\epsilon'_{\omega} \Delta \epsilon_{\omega}} d^2 n'_{\pm} I_{bs} \frac{\epsilon'^{2} \beta'_{-} |M_{fi}|^2 \hbar^2 c^2}{16 e - e_{+} + e'_{-} e'_{+}} \right),
\]
(10.5.41)

\[
(\chi E)^{e^{-}e^{+} \rightarrow e^{-}'e^{+}'} = \frac{1}{\Delta \epsilon_{\pm, \omega}} \left( \int_{\epsilon_{\omega} \Delta \epsilon_{\omega}} d^2 n'_{\pm} I_{bs} \frac{\epsilon'^{2} \beta'_{-} |M_{fi}|^2 \hbar^2 c^2}{16 e - e_{+} + e'_{-} e'_{+}} \right) +
\]
\[
+ \frac{1}{\Delta \epsilon_{\pm, \omega}} \left( \int_{\epsilon'_{\omega} \Delta \epsilon_{\omega}} d^2 n'_{\pm} I_{bs} \frac{\epsilon'^{2} \beta'_{-} |M_{fi}|^2 \hbar^2 c^2}{16 e - e_{+} + e'_{-} e'_{+}} \right),
\]
(10.5.42)

where \( d^2 n'_{\pm} = dn dn_{+} do'_{-}, dn_{\pm} = de_{\pm} do_{\pm} e_{\pm}^{2} \beta_{\pm} f_{\pm} \), and the Jacobian is

\[
I_{bs} = \frac{\epsilon'^{2} \beta'_{-}}{(\epsilon'_{-} + \epsilon'_{+}) \beta'_{-} - (\epsilon_{-} \beta_{-} b_{-} + \epsilon_{+} \beta_{+} b_{+}) \cdot b'_{-}}.
\]
(10.5.43)

Analogously to the case of pair creation and annihilation in Section (10.5.2) the energies of final state particles are given by (10.5.18) with the coefficients \( \tilde{A} = (\epsilon_{-} + \epsilon_{+})^{2} - (\epsilon_{-} \beta_{-} b_{-} \cdot b'_{-} + \epsilon_{+} \beta_{+} b_{+} \cdot b'_{+})^{2}, \tilde{B} = (\epsilon_{-} + \epsilon_{+}) [m_{e}^{2} c^{4} + e_{-} e_{+} (1 - \beta_{-} \beta_{+} b_{-} \cdot b_{+})]^{2}, \tilde{C} = [m_{e}^{2} c^{4} + e_{-} e_{+} (1 - \beta_{-} \beta_{+} b_{-} \cdot b_{+})]^{2} + m_{e}^{2} c^{4} [e_{-} \beta_{-} b_{-} \cdot b'_{-} + e_{+} \beta_{+} b_{+} \cdot b'_{+}]^{2} \]

In order to select the correct root one has to check the condition (10.5.35) changing the subscripts \( 1 \rightarrow -, 2 \rightarrow + \).
10.6 Three-body processes

As we discussed above, for the collisional integrals for three-body interactions we assume that particles already reached kinetic equilibrium. In that case one can use the corresponding expressions, obtained in the literature for the thermal equilibrium case, and multiply the collisional integrals by the exponents, containing the chemical potentials of particles.

Emission coefficients for triple interactions in thermal equilibrium may be computed by averaging of the differential cross-sections given in Section 5 of the corresponding processes over the thermal distributions of particles. Analytic results as a rule exist only for nonrelativistic and/or ultrarelativistic cases. The only way to get approximate analytical expressions is then to find the fitting formulas, connecting the two limiting cases with reasonable accuracy. This work has been done by Svensson [497], also for reactions with protons, and in what follows we adopt the emission and absorption coefficient for triple interactions given in that paper.

Bremsstrahlung

\[
\eta_{\gamma}^{\mu \nu} \rightarrow e^{\mu} e^{\nu} \gamma = (n_+^2 + n_-^2) \frac{16}{3} \frac{a c}{\varepsilon} \left( \frac{e^2}{m_c c^2} \right)^2 \times 
\]

\[
\times \ln \left[ 4 \zeta (11.2 + 10.4 \theta^2) \frac{\theta}{\varepsilon} \right] \frac{3}{2} \sqrt{2 \theta + 2 \theta^2} \exp \left( \frac{1}{\theta} \right) K_2 \left( \frac{1}{\theta} \right), \quad (10.6.1)
\]

\[
\eta_{\gamma}^{\mu +} \rightarrow e^{\mu} e^{+} \gamma = n_+ n_- \frac{16}{3} \frac{a c}{\varepsilon} \left( \frac{e^2}{m_c c^2} \right)^2 \times 
\]

\[
\times \ln \left[ 4 \zeta (1 + 10.4 \theta^2) \frac{\theta}{\varepsilon} \right] \frac{\sqrt{2} + 2 \theta + 2 \theta^2}{\exp \left( \frac{1}{\theta} \right) K_2 \left( \frac{1}{\theta} \right)}, \quad (10.6.2)
\]

where \( \zeta = e^{-0.5772}, \) and \( K_2 \left( \frac{1}{\theta} \right) \) is the modified Bessel function of the second kind of order 2.

Double Compton scattering

\[
\eta_{\gamma}^{\mu - \nu} \rightarrow e^{\mu} e^{\nu} \gamma = (n_+ + n_-) n_+ \frac{128}{3} \frac{a c}{\varepsilon} \times 
\]

\[
\times \left( \frac{e^2}{m_c c^2} \right)^2 \frac{\theta^2}{1 + 13.91 \theta + 11.05 \theta^2 + 19.92 \theta^3}, \quad (10.6.3)
\]

Three photon annihilation
\[ \eta_{\gamma}^{e^{\pm}e^{\mp} \rightarrow \gamma \gamma} = n_+ n_- \alpha c \left( \frac{e^2}{m_e c^2} \right)^2 \frac{1}{\epsilon} \left( \frac{2 \ln^2 2 \xi \theta + \frac{\pi^2}{6} - \frac{1}{2}}{4 \theta + \frac{1}{\epsilon^2}} \right), \] (10.6.4)

where two limiting approximations [497] are joined together.

Radiative pair production

\[ \eta_{\gamma}^{e^{+}e^{+} \rightarrow \gamma \gamma} = \frac{n_+^2}{n_+ n_-} \left[ \frac{K_2 (1/\theta)}{2 \theta^2} \right]^2. \] (10.6.5)

Electron-photon pair production

\[ \eta_{\gamma}^{e^{+}e^{-} \rightarrow \gamma \gamma} = \begin{cases} (n_+ + n_-) n_\gamma \alpha c \left( \frac{e^2}{m_e c^2} \right)^2 \exp \left( -\frac{2}{\epsilon^2} \right) 16.1 \theta^{0.541}, & \theta \leq 2, \\ (n_+ + n_-) n_\gamma \alpha c \left( \frac{e^2}{m_e c^2} \right)^2 \left( \frac{5 \theta}{9} \ln 2 \xi \theta - \frac{8}{27} \right) \frac{1}{1 + 0.5/\theta}, & \theta > 2. \end{cases} \] (10.6.6)

The absorption coefficient for three-body processes is written as

\[ \chi_{\gamma}^{3p} = \eta_{\gamma}^{3p} / E_{\gamma}^{eq}, \] (10.6.7)

where \( \eta_{\gamma}^{3p} \) is the sum of the emission coefficients of photons in the three-particle processes, \( E_{\gamma}^{eq} = 2 \pi e^3 f_{\gamma}^{eq} / c^3 \), where \( f_{\gamma}^{eq} \) is given by (10.1.18).

From equation (10.2.4), the law of energy conservation in the three-body processes is

\[ \int \sum_i (\eta_i^{3p} - \chi_i^{3p} E_i) d\mu d\epsilon = 0. \] (10.6.8)

For exact conservation of energy in these processes the following coefficients of emission and absorption for electrons are introduced:

\[ \chi_{\gamma}^{3p} = \int (\eta_{\gamma}^{3p} - \chi_{\gamma}^{3p} E_{\gamma}) d\mu d\epsilon, \quad \eta_{\gamma}^{3p} = 0, \quad \int (\eta_{\gamma}^{3p} - \chi_{\gamma}^{3p} E_{\gamma}) d\mu > 0, \] (10.6.9)

or

\[ \frac{\eta_{\gamma}^{3p}}{E_{\gamma}} = -\int (\eta_{\gamma}^{3p} - \chi_{\gamma}^{3p} E_{\gamma}) d\mu / \int E_{\gamma} d\mu, \quad \chi_{\gamma}^{3p} = 0, \quad \int (\eta_{\gamma}^{3p} - \chi_{\gamma}^{3p} E_{\gamma}) d\mu < 0. \] (10.6.10)
10.7 Cutoff in the Coulomb scattering

Denote quantities in the center of mass (CM) frame with index 0, and with prime after interaction. Suppose there are two particles with masses $m_1$ and $m_2$. The change of the angle of the first particle in CM system is

$$\theta_{10} = \arccos(b_{10} \cdot b'_{10}), \quad (10.7.1)$$

the numerical grid size is $\Delta \theta_g$, the minimal angle at the scattering is $\theta_{\text{min}}$.

By definition in the CM frame

$$p_{10} + p_{20} = 0, \quad (10.7.2)$$

where

$$p_{i0} = p_i + \left[ (\Gamma - 1)(Np_i) - \Gamma \frac{V}{c} \epsilon_i \right] N, \quad i = 1, 2, \quad (10.7.3)$$

and

$$\epsilon_i = \Gamma(\epsilon_{i0} + Vp_{i0}). \quad (10.7.4)$$

Then for the velocity of the CM frame

$$\frac{V}{c} = c \frac{p_1 + p_2}{\epsilon_1 + \epsilon_2}, \quad N = \frac{V}{V'}, \quad \Gamma = \frac{1}{\sqrt{1 - \left(\frac{V}{c}\right)^2}}. \quad (10.7.5)$$

By definition

$$b_{10} = b_{20}, \quad b'_{10} = b'_{20}, \quad (10.7.6)$$

and then

$$|p_{10}| = |p_{20}| = p_0 \equiv \frac{1}{c} \sqrt{\epsilon_{10}^2 - m_1^2 c^4} = \frac{1}{c} \sqrt{\epsilon_{20}^2 - m_2^2 c^4}, \quad (10.7.7)$$

where

$$\epsilon_{10} = \frac{(\epsilon_1 + \epsilon_2)^2 - \Gamma^2(m_2^2 - m_1^2)c^4}{2(\epsilon_1 + \epsilon_2)\Gamma}, \quad (10.7.8)$$

$$\epsilon_{20} = \frac{(\epsilon_1 + \epsilon_2)^2 + \Gamma^2(m_2^2 - m_1^2)c^4}{2(\epsilon_1 + \epsilon_2)\Gamma}. \quad (10.7.9)$$

Haug [506] gives the minimal scattering angle in the center of mass system

$$\theta_{\text{min}} = \frac{2\hbar}{m_e cD} \frac{\gamma_r}{(\gamma_r + 1)\sqrt{2(\gamma_r - 1)}}, \quad (10.7.10)$$
where the maximum impact parameter (neglecting the effect of protons) is

\[ D = \frac{c^2 p_0}{\omega \epsilon_{10}}, \]  

(10.7.11)

and the invariant Lorentz factor of relative motion (e.g. [506]) is

\[ \gamma_r = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} = \frac{\epsilon_1 \epsilon_2 - \mathbf{p}_1 \cdot \mathbf{p}_2 c^2}{m_1 m_2 c^4}. \]  

(10.7.12)

Finally, in the CM frame

\[ t_{\text{min}} = 2 \left[ (m_e c)^2 - \left(\frac{\epsilon_{10}}{c}\right)^2 \left(1 - \beta_{10}^2 \cos \theta_{\text{min}}\right)\right]. \]

Since it is invariant, \( t \) in the denominator of \(|M_{fi}|^2\) in (10.5.34) is replaced by the value \( t \sqrt{1 + t_{\text{min}}^2 / t^2} \) to implement the cutoff scheme. Also at the scattering of equivalent particles \( u \) in the denominator of \(|M_{fi}|^2\) in (10.5.34) is changed to the value \( u \sqrt{1 + t_{\text{min}}^2 / u^2} \).

### 10.8 Numerical results

The results of numerical simulations are reported below. Two limiting initial conditions with flat spectra are chosen: (i) electron–positron pairs with a \( 10^{-5} \) energy fraction of photons and (ii) the reverse case, i.e., photons with a \( 10^{-5} \) energy fraction of pairs. The grid consists of 60 energy intervals and \( 16 \times 32 \) intervals for two angles characterizing the direction of the particle momenta. In both cases the total energy density is \( \rho = 10^{24} \text{ erg} \cdot \text{cm}^{-3} \). In the first case initial concentration of pairs is \( 3.1 \cdot 10^{29} \text{ cm}^{-3} \), in the second case the concentration of photons is \( 7.2 \cdot 10^{29} \text{ cm}^{-3} \).

In Fig. 10.1, panel A concentrations of photons and pairs as well as their sum for both initial conditions are shown. After calculations begin, concentrations and energy density of photons (pairs) increase rapidly with time, due to annihilation (creation) of pairs. Then, in the kinetic equilibrium phase, concentrations of each component stay almost constant, and the sum of concentrations of photons and pairs remains unchanged. Finally, both individual components and their sum reach stationary values. If one compares and contrasts both cases as reproduced in Fig. 10.1 A one can see that, although the initial conditions are drastically different, in both cases the same asymptotic values of the concentration are reached.

One can see in Fig. 10.1, panel C that the spectral density of photons and pairs can be fitted already at \( t_k \approx 20 t_{\text{Cs}} \approx 7 \cdot 10^{-15} \text{ sec} \) by distribution func-
Figure 10.1: A: Dependence on time of concentrations of pairs (black), photons (red) and both (thick) when all interactions take place (solid). Upper (lower) figure corresponds to the case when initially there are mainly pairs (photons). Dotted curves on the upper figure show concentrations when inverse triple interactions are neglected. In this case an enhancement of the pairs occurs with the corresponding increase in photon number and thermal equilibrium is never reached. B: Time dependence of temperatures, measured on the left axis (solid), and chemical potentials, measured on the right axis (dotted), of electrons (black) and photons (red). The dashed lines correspond to the reaching of the kinetic ($\sim 10^{-14}$ sec) and the thermal ($\sim 10^{-12}$ sec) equilibria. Upper (lower) figure corresponds to the case when initially there are mainly pairs (photons). C: Spectra of pairs (upper figure) and photons (lower figure) when initially only pairs are present. The black curve represents the results of numerical calculations obtained successively at $t = 0$, $t = t_k$ and $t = t_{th}$ (see the text). Both spectra of photons and pairs are initially taken to be flat. The yellow curves indicate the spectra obtained from (10.1.18) at $t = t_k$. The perfect fit of the two curves is most evident in the entire energy range leading to the first determination of the temperature and chemical potential both for pairs and photons. The orange curves indicate the final spectra as thermal equilibrium is reached.

From (10.1.18) with definite values of temperature $\theta_k(t_{CS}) \approx 1.2$ and chemical potential $\phi_k(t_k) \approx -4.5$, common for pairs and photons. As expected, after $t_k$ the distribution functions preserve their form (10.1.18) with the values of temperature and chemical potential changing in time, as shown in Fig. 10.1, panel B. As one can see from this figure, the chemical potential evolves with time and reaches zero at the moment $t_{th} \approx \frac{1}{\alpha} t_k \approx 7 \times 10^{-13}$ sec, corresponding to the final stationary solution. Condition (10.4.6) is satisfied in kinetic equilibrium.
The necessary condition for thermal equilibrium in the pair plasma is the detailed balance between direct and inverse triple interactions. This point is usually neglected in the literature where there are claims that the thermal equilibrium may be established with only binary interactions [491]. In order to demonstrate it explicitly in Fig. 10.1, panel A the dependence of concentrations of pairs and photons when inverse triple interactions are artificially switched off is also shown. In this case (see dotted curves in the upper part of Fig. 10.1, panel A), after kinetic equilibrium is reached concentrations of pairs decrease monotonically with time, and thermal equilibrium is never reached.

The existence of a non-null chemical potential for photons indicates the departure of the distribution function from the one corresponding to thermal equilibrium. Negative (positive) value of the chemical potential generates an increase (decrease) of the number of particles in order to approach the one corresponding to the thermal equilibrium state. Then, since the total number of particles increases (decreases), the energy is shared between more (less) particles and the temperature decreases (increases). Clearly, as thermal equilibrium is approached, the chemical potential of photons is zero.

In this example with the energy density \(10^{24} \text{ erg \cdot cm}^{-3}\) the thermal equilibrium is reached at \(\sim 7 \cdot 10^{-13} \text{ sec}\) with the final temperature \(T_{\text{th}} = 0.26 \text{ MeV}\). For a larger energy density the duration of the kinetic equilibrium phase, as well as of the thermalization timescale, is smaller. Recall, that in entire temperature range the plasma is nondegenerate.

The results, obtained for the case of an uniform plasma, can only be adopted for a description of a physical system with dimensions \(R_0 > 4.3 \times 10^{-5} \text{ cm}\).

The assumption of the constancy of the energy density is only valid if the dynamical timescale \(t_{\text{dyn}} = \left(\frac{1}{R} \frac{dR}{dt}\right)^{-1}\) of the plasma is much larger than the above timescale \(t_{\text{th}}\) which is indeed true in all the cases of astrophysical interest.

Since thermal equilibrium is obtained already on the timescale \(t_{\text{th}} < 10^{-12} \text{ sec}\), and such a state is independent of the initial distribution functions for electrons, positrons and photons, the sufficient condition to obtain an isothermal distribution on a causally disconnected spatial scale \(R > ct_{\text{th}} = 10^{-2} \text{ cm}\) is the request of constancy of the energy density on such a scale as well as, of course, the invariance of the physical laws.

To summarize, the evolution of an initially nonequilibrium optically thick electron–positron-photon plasma is considered up to reaching thermal equilibrium. Starting from arbitrary initial conditions kinetic equilibrium is obtained from first principles, directly solving the relativistic Boltzmann equations with collisional integrals computed from QED matrix elements. The essential role of direct and inverse triple interactions in reaching thermal equilibrium is demonstrated. These results can be applied in the theories of the early Universe and of astrophysical sources, where thermal equilibrium is postulated at the very early stages. These results can in principle be tested in
laboratory experiments in the generation of electron–positron pairs.
11 Concluding remarks

We have reviewed three fundamental quantum processes which have highlighted some of the greatest effort in theoretical and experimental physics in last seventy years. They all deal with creation and annihilation of electron–positron pairs. We have followed the original path starting from the classical works of Dirac, on the process $e^+e^- \rightarrow 2\gamma$, and the inverse process, $2\gamma \rightarrow e^+e^-$, considered by Breit–Wheeler. We have then reviewed the $e^+e^-$ pair creation in a critical electric field $E_c = m^2e^3/(\hbar c)$ and the Sauter-Heisenberg-Euler-Schwinger description of this process both in Quantum Mechanics and Quantum Electro-Dynamics. We have also taken this occasion to reconstruct the exciting conceptual developments, initiated by the Sauter work, enlarged by the Born-Infeld nonlinear electrodynamical approach, finally leading to the Euler and Euler-Heisenberg results. We were guided in this reconstruction by the memories of many discussions of one of us (RR) with Werner Heisenberg. We have then reviewed the latest theoretical developments deriving the general formula for pair production rate in electric fields varying in space and in time, compared with one in a constant electric field approximation originally studied by Schwinger within QED. We also reviewed recent studies of pair production rates in selected electric fields varying both in space and in time, obtained in the literature using instanton and JWKB methods. Special attention has been given to review the pair production rate in electric fields alternating periodically in time, early derived by Brezin, Itzykson and Popov, and the nonlinear Compton effect in the processes of electrons and photons colliding with laser beams, studied by Nikishov and Narozhny. These theoretical results play an essential role in Laboratory experiments to observe the pair production phenomenon using laser technologies.

We then reviewed the different level of verification of these three processes in experiments carried all over the world. We stressed the success of experimental verification of the Dirac process, by far one of the most prolific and best tested process in the field of physics. We also recalled the study of the hadronic branch in addition to the pure electrodynamical branch originally studied by Dirac, made possible by the introduction of $e^+e^-$ storage rings technology. We then turned to the very exciting current situation which sees possibly the Breit–Wheeler formula reaching its first experimental verification. This result is made possible thanks to the current great developments of laser physics. We reviewed as well the somewhat traumatic situation in the last forty years of the heavy-ion collisions in Darmstadt and Brookhaven, yet unsuccessfully attempting to observe the creation of electron–positron pairs.
We also reviewed how this vast experimental program was rooted in the theoretical ideas of Zeldovich, Popov, Greiner and their schools.

We have also recalled the novelty in the field of relativistic astrophysics where we are daily observing the phenomenon of Gamma Ray Bursts [404, 477, 516–518]. These bursts of photons occur in energy range keV to MeV, last about one second and come from astrophysical sources located at a cosmological distance [394, 519–522]. The energy released is up to \( \sim 10^{55} \text{ ergs} \), equivalent to all the energy emitted by all the stars of all the galaxies of the entire visible Universe during that second. It is generally agreed that the energetics of these GRB sources is dominated by a dense plasma of electrons, positrons and photons created during the process of gravitational collapse leading to a Black Hole, see e.g. [68, 398–400] and references therein. The Sauter-Heisenberg-Euler-Schwinger vacuum polarization process, we have considered in the first part of the report, is a classic theoretical model to study the creation of an electron–positron optically thick plasma. Similarly the Breit–Wheeler and the Dirac processes we have discussed, are essential in describing the further evolution of such an optically thick electron–positron plasma. The GRBs present an unique opportunity to test new unexplored regime of ultrahigh energy physics with Lorentz factor \( \gamma \sim 100 – 1000 \) and relativistic field theories in the strongest general relativistic domain.

The aim in this report, in addition to describe the above mentioned three basic quantum processes, has been to identify and review three basic relativistic regimes dealing with an optically thin and optically thick electron–positron plasma.

The first topic contains the basic results of the physics of black holes, of their energetics and of the associated process of vacuum polarization. We reviewed the procedures to generalize in a Kerr–Newman geometry the QED treatment of Schwinger and the creation of enormous number of \( 10^{60} \) electron–positron pairs in such a process.

The second topic is the back reaction of a newly created electron–positron plasma on an overcritical electric field. Again we reviewed the Breit–Wheeler and Dirac processes applied in the wider context of the Vlasov–Boltzmann–Maxwell equations. To discuss the back reaction of electron–positron pair on external electric fields, we reviewed semi-classical and kinetic theories describing the plasma oscillations using respectively the Dirac-Maxwell equations and the Boltzmann–Vlasov equations. We also reviewed the discussions of plasma oscillations damping due to quantum decoherence and collisions, described by respectively the quantum Boltzmann–Vlasov equation and Boltzmann–Vlasov equation with particle collisions terms. We particularly addressed the study of the influence of the collision processes \( e^+e^- \leftrightarrow \gamma \gamma \) on the plasma oscillations in supercritical electric field \( E > E_c \). After \( 10^3–4 \) Compton times, the oscillating electric field is damped to its critical value with a large number of photons created. An equipartition of number and energy between electron–positron pairs and photons is reached. For the
plasma oscillation with undercritical electric field $E \lesssim E_c$, we recalled that
electron–positron pairs, created by the vacuum polarization process, move as
charged particles in external electric field reaching a maximum Lorentz fac-
tor at finite length of oscillations, instead of arbitrary large Lorentz factors,
as traditionally assumed. Finally we point out some recent results which dif-
ferentiate the case $E > E_c$ from the one $E < E_c$ with respect to the creation
of the rest mass of the pair versus its kinetic energy. For $E > E_c$ the vac-
umum polarization process transforms the electromagnetic energy of the field
mainly in the rest mass of pairs, with moderate contribution to their kinetic
energy. Such phenomena, certainly fundamental on astrophysical scales, may
become soon directly testable in the verification of the Breit–Wheeler process
tested in laser experiments in the laboratory.

As the third topic we have reviewed the recent progress in the understand-
ing of thermalization process of an optically thick electron–positron-photon
plasma. Numerical integration of relativistic Boltzmann equation with colli-
sional integrals for binary and triple interactions is used to follow the time
evolution of such a plasma, in the range of energies per particle between
0.1 and 10 MeV, starting from arbitrary nonequilibrium configuration. It is
recalled that there exist two types of equilibria in such a plasma: kinetic equi-
librium, when all particles are at the same temperature, but have different
nonzero chemical potential of photons, and thermal equilibrium, when chem-
ical potentials vanish. The crucial role of direct and inverse binary and triple
interactions in reaching thermal equilibrium is emphasized.

In a forthcoming report we will address how the above mentioned three
relativistic processes can be applied to a variety of astrophysical systems in-
cluding neutron stars formation and gravitational collapse, supernovae ex-
plosions and GRBs.

This report is dedicated to the progress of theoretical physics in extreme
regimes of relativistic field theories which are on the verge of finding their
experimental and observational verification in physics and astrophysics. It is
then possible from our review and the many references we have given to gain
a basic understanding of this new field of research. The three topics which
we have reviewed are closely linked to the three quantum processes currently
being tested in precision measurement in the laboratories. The experiments
in the laboratories and the astrophysical observations cover complementary
aspects which may facilitate a deeper and wider understanding of the nuclear
and laser physics processes, of heavy-ion collisions as well as neutron stars
formation and gravitational collapse, supernovae and GRBs phenomena. We
shall return on such an astrophysical and observational topics in a dedicated
forthcoming report.

* * *
We are witnessing in these times some enormous experimental and observational successes which are going to be the natural ground to test some of the theoretical works which we have reviewed in this report. Among the many experimental progresses being done in particle accelerators worldwide we like to give special mention to two outstanding experimental facilities which are expected to give results in the forthcoming years. We refer here to the National Ignition Facility at the Lawrence Livermore National Laboratory to be soon becoming operational, see e.g. [523] as well as the corresponding facility in France, the Mega Joule project [524].

In astrophysics these results will be tested in galactic and extragalactic black holes observed in binary X-ray sources, active galactic nuclei, microquasars and in the process of gravitational collapse to a neutron star and also of two neutron stars to a black hole in GRBs. The progress there is equally remarkable. In the last few days after the completion of this report thanks to the tremendous progress in observational technology for the first time a massive hypergiant star has been identified as the progenitor of the supernova SN 2005gl [525]. In parallel the joint success of observations of the flotilla of X-ray observatories and ground-based large telescopes [177] have allowed to identify the first object ever observed at $z \approx 8$ the GRB090423 [526].

To follow the progress of this field we are planning a new report which will be directed to the astrophysical nature of the progenitors and the initial physical conditions leading to the process of the gravitational collapse. There the electrodynamical structure of neutron stars, the phenomenon of the supernova explosion as well as theories of Gamma-Ray Bursts (GRBs) will be discussed. Both in the case of neutron stars and the case of black holes there are fundamental issues still to be understood about the process of gravitational collapse especially with the electrodynamical conditions at the onset of the process. The major difficulties appear to be connected with the fact that all fundamental interactions, the gravitational, the electromagnetic, the strong, the weak interactions appear to participate in essential way to this process which appear to be therefore one of the most interesting fundamental process of theoretical physics. Current progress is presented in the following works [527–538]. What is important to recall at this stage is only that both the supernovae and GRBs processes are among the most energetic and transient phenomena ever observed in the Universe: a supernova can reach energy of $\sim 10^{52} \text{ ergs}$ (hypernovae) on a time scale of a few months and GRBs can have emission of up to $\sim 10^{55} \text{ ergs}$ [539] in a time scale as short as of a few seconds. The central role in their description of neutron stars, for supernova as well as of black holes and the electron–positron plasma discussed in this report, for GRBs, are widely recognized. The reason which makes this last research so important can be seen in historical prospective: the Sun has been the arena to understand the thermonuclear evolution of stars [540], Cyg X-1 has evidenced the gravitational energy role in explaining an astrophysical system [541], the GRBs are promising to prove the existence for the first
time of the “blackholic energy”. These three quantum processes described in our report reveal the basic phenomena in the process of gravitational collapse predicted by Einstein theory of General Relativity [177].
12 Notes added in proof

- In Ref. [258] two terms related to the initial and final frequencies are missing in the effective action (real part and imaginary part). These have been corrected in [542, 543].

- Some new results were obtained for QED in quasi constant in time electric field, in particular, the distributions of particles created is discussed in [544, 545]; consistency restrictions on maximal electric field strength in QFT are discussed in [546]. One-loop energy-momentum tensor in QED is obtained in [547]. The exact rate of pair production by a smooth potential proportional to \( \tanh(kz) \) in three dimensions is obtained in [548].

- A recent review on the muon anomalous magnetic moment (muon g-2), offers the possibility, by making most precise measurement of muon g-2 in low-energies, to infer virtual hadronic vacuum polarization and light-by-light scattering effects due to virtual quark-pairs production in high-energies. In recent BNL E821 experiment\(^1\), the muon anomalous magnetic moment can be rather accurately measured. In the theory of Standard Model for elementary particles, the muon anomalous magnetic moment \( a_\mu \) receive leptonic QED-contributions \((e, \mu, \tau)\) to \( a_\mu^{\text{QED}} \), which has been calculated, see for a review [549], up to 5-loop contributions \( \mathcal{O}(\alpha^5) \)

\[
a_\mu^{\text{QED}} \sim 663(20)(4.6) \left( \frac{\alpha}{\pi} \right)^5.
\]

While hadronic \((u, d, s, c, \cdots)\) contributions to the muon anomalous magnetic moment \( a_\mu^{\text{had}} \) contain, due to the strong interaction, both perturbative and non-perturbative parts, the \( \mathcal{O}(\alpha^2) \) contribution to \( a_\mu^{\text{had}} \)

\[
a_\mu^{(4)}(\text{vap, had}) = \left( \frac{\alpha m_\mu}{3\pi} \right)^2 \left( \int_{m_{\pi}^2}^{E_{\text{cut}}^2} ds \frac{R_{\text{had}}^\text{data}(s) \tilde{K}(s)}{s^2} \right) + \int_{E_{\text{cut}}^2}^{\infty} ds \frac{R_{\text{had}}^\text{PQCD}(s) \tilde{K}(s)}{s^2},
\]

where \( R(s) \) is given by

\[
R_{\text{had}}(s) = \sigma(e^+e^- \rightarrow \text{hadrons}) / \frac{4\pi \alpha(s)^2}{3s}.
\]

\(^1\)http://www.g-2.bnl.gov
and $K(s)$ is the vacuum polarization contribution,

$$K(s) = \int_0^1 dx \frac{x^2(1-x)}{x^2 + (s/m^2_k)(1-x)},$$

and a cut $E_{\text{cut}}$ in the energy, separating the non-perturbative part to be evaluated from data and the perturbative high energy tail to be calculated using perturbative QCD (pQCD), analogously to QED-calculations. The pQCD calculations may only be trusted above 2 GeV and away from threshold and resonances. In the report [549], authors resort to a semiphenomenological approach using dispersion relations together with the optical theorem and experimental data. The “measurements of $R_{\text{had}}(s)$” get more difficult as increasing energies more and more channels open for meson-resonances. In addition, the most problematic set of hadronic corrections is that related to hadronic light-by-light scattering, which for the first time show up at order $O(\alpha^3)$ via the diagrams with insertion of a box with four photon lines. As a contribution to the anomalous magnetic moment three of four photons are virtual and to be integrated over all four-momentum space, such that a direct experimental input for the non-perturbative dressed four-photon correlator is not available. In this case one has to resort the low energy effective description of QCD like chiral perturbation theory (CHPT) extended to include vector-mesons, which is reviewed in detail [549]. Furthermore, the Electroweak corrections of weak virtual process including intermediate bosons $W^\pm$ and $Z$ to $g - 2$ are important and now almost three standard deviations, and without it the deviation between theory and experiment would be the $6\sigma$ level. The test of the weak contribution is actually one of the milestones achieved by Brookhaven experiment E821 (see the footnote above). These studies and experiments are crucial to include the contributions from all known particles and interactions such that from a possible deviation between theory and experiment we may get a hint of the yet unknown physics.

- Some recent results on $e^+e^-$ annihilation cross sections in the GeV region are obtained in the following experiments: KLOE [550–552], CMD-2 [553, 554] and SND [554].

- We would like to point out that in our report, one-loop vacuum polarization and light-by-light scattering effects (effective Euler-Heisenberg Lagrangian), as well as their high-order corrections in low-energies are considered in Section 5.7; non-linear Compton effect is discussed in both Section 5.10 (theory) and Sections 7.2,7.3 (experiments); the Breit-Wheeler cutoff in high-energy $\gamma$-rays for astrophysics in Section 7.4. All these discussions are limited in the leptonic sector for low-energies.
The reason of recording here these hadronic contributions in the high-energy region is motivated by our expectation that these effects will be possibly soon detected by direct measurements of gamma ray emission from high-energy astrophysical processes. We shall return on this topic in our forthcoming report already mentioned above.

Acknowledgment.


[31] V. S. Popov, Pair Production in a Variable External Field (Quasiclassical Approximation), Soviet Journal of Experimental and Theoretical Physics 34 (1972) 709–64.


[40] A. I. Nikishov, V. I. Ritus, Quantum processes in the field of a plane electromagnetic wave and in a constant field. i, Sov. Phys. JETP 19 (1964) 529–541.


URL http://link.aps.org/abstract/PRD/v78/e025011


Bibliography


URL http://users.physik.fu-berlin.de/ kleinert/b6/psfiles/


[148] The L3 experiment at cern.


[198] C. e. a. Bacci, Gamma-gamma interaction processes at anode $e^+e^-$ storage ring — measurement of reaction $e^\pm e^\mp \rightarrow e^\pm e^\mp e^\pm e^\mp$, Lett. Nuovo Cimento 3 (1972) 709.


\texttt{arXiv:hep-th/0406216}.

\texttt{arXiv:hep-th/9907190}.

[210] M. S. Marinov, V. S. Popov, Electron-positron pair creation from vacuum induced by variable electric field, Fortschritte der Physik 25 (1977) 373–400.


Bibliography


[244] V. S. Popov, Resonant pair production in a strong electric field, ZhETF Pis ma Redaktsiiu 18 (1973) 435–


URL http://www.hep.lu.se/workshop/lincol/proceedings/gmaterlik.ps


URL http://www.hep.lu.se/workshop/lincol/proceedings/indau.ps

[253] The linac coherent light source.
URL http://www-ssrl.slac.stanford.edu/lcls/


[277] SLAC Experiment 144.
URL http://www.slac.stanford.edu/exp/e144/e144.html


Bibliography


[312] V. S. Popov, Spontaneous positron production in collisions between heavy nuclei, Zhurnal Eksperimental’noi i Teoreticheskoi Fiziki 65 (1973) 35–53.


Bibliography


[338] J. S. Greenberg, C. K. Davis, P. Vincent, Evidence for Quasimolecular K X-Ray Emission in Heavy-Ion Collisions from the Observation of


[351] C. Kozhuharov, P. Kienle, E. Berdermann, H. Bokemeyer, J. S. Greenberg, Y. Nakayama, P. Vincent, H. Backe, L. Handschug,
Bibliography


[369] B. Carter, Has the Black Hole Equilibrium Problem Been Solved?, in: T. Piran, R. Ruffini (Eds.), Recent Developments in Theoretical and


[424] S. W. Hawking, Particle creation by black holes, Communications in Mathematical Physics 43 (1975) 199–220.


doi:10.1016/0370-2693(74)90505-X.


