IMPERIAL COLLEGE LONDON

Department of Theoretical Physics

The role of Axions in Baryogenesis

Claves H. do Amaral

Supervised by Prof. Kellogg Stelle

Imperial College London

Submitted in partial fulfilment of the requirements for the degree of Master of Science of Imperial College London

September 25, 2015

Fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza
(D.Alighieri)

Contents

1 Introduction

We are certain[1] that for any known particle there exists the antiparticle with exactly the same mass, $m = \bar{m}$ decay width, $\Gamma = \bar{\Gamma}$, and opposite signs of all the charges associated with this particle, $Q_j = -\bar{Q}_j$. Despite this impressive symmetry, which would naturally imply equal number densities of particles and antiparticles in the Universe, $n = \bar{n}$, the current situation is quite different: the Universe (at least in our neighborhood) is predominantly populated by particles: protons, neutrons, and electrons [12]. The evidence that the universe is has no or very few antimatter comes from a variety of different observations. On the very small scale, the fact that we do not observe proton-antiproton annihilations in our everyday life, is a strong evidence that our world is composed only of matter and no antimatter. Moving up in scale, the success of satellite launches, lunar landings, and planetary probes suggests that our solar system is made up of the same type of matter that we are, and that there is negligible antimatter on that scale. The first detection of antimatter outside particle accelerators comes from cosmic rays. Mixed in with the many protons present in radiation coming from space are a few antiprotons, present at a level of around 10^{-4} in comparison with the number of protons. Also, if matter and antimatter galaxies were to coexist in clusters of galaxies[2], then we would expect there to be a detectable background of γ-radiation coming from the reaction of annihilation of $p\bar{p}$ into π-mesons with the subsequent decay $\pi^0 \to 2\gamma$, which would take place in the boundary area between the world and anti-world. This background is not observed and so we conclude that there is negligible antimatter on the scale of clusters. Finally, if large domains of matter and antimatter exist, then annihilations would take place at the interfaces between them. If the typical size of such a domain was small enough, then the energy released by these annihilations would result in a diffuse γ-ray background and a distortion of the cosmic microwave radiation, neither of which is observed. Today there is general agreement that the universe consists entirely of matter on all scales up to the Hubble size[3]. It therefore seems that the universe is fundamentally matter-antimatter asymmetric. [32].

Baryogenesis is the hypothetical physical processes that produced an asymmetry (imbalance) between baryons[4] and antibaryons in the very early Universe. Understanding the generation of the matter-antimatter symmetry of the universe is one of the key motivations for physics beyond the Standard Model (SM). In this paper a mechanism is described, which explains the dominance of particles over antiparticles in the universe by the introduction of a very light pseudoscalar, pseudo-Goldstone boson, the "axion" [23].

The necessary conditions for the generation of the asymmetry, as formulated by Sakharov [25], are the following

1. It must be assumed that there are no antimatter bodies in nature i.e. the Universe is asymmetrical with respect to the number of particles and antiparticles (C asymmetry). Which implies different interactions for particles and antiparticles, or in other words, a violation of Charge (C) and Charge-Parity (CP) symmetries is required.

2. The absence of antibaryons implies non conservation of baryonic charge B.

3. Deviation from thermal equilibrium in the early universe is also required.

The first of these is required because, if C and CP are exact symmetries, then one can prove that the total rate for any process which produces an excess of baryons is equal to the rate of the

[1]The Dirac equation, formulated by Paul Dirac as part of the development of relativistic quantum mechanics, predicts the existence of antiparticles along with the expected solutions for the corresponding particles. Since then, it has been verified experimentally that every known kind of particle has a corresponding antiparticle. The CPT Theorem guarantees that a particle and its antiparticle have exactly the same mass and lifetime, and exactly opposite charge.

[2]Like Virgo which typically contain 10^{13} to $10^{14} M_\odot$ of material.

[3]If there is a significant amount of antimatter in the Universe, it must be segregated from matter on scales at least as large as $10^{12} M_\odot$, and probably larger than $10^{14} M_\odot$. On the face of it, this does not preclude a baryon symmetric Universe. However, in a locally-baryon-symmetric Universe nucleons and antinucleons remain in chemical equilibrium down to a temperature of $\sim 22\,MeV$, when $n_B/s = n_{\bar{B}}/s \simeq 7 \times 10^{-20}$, a number that is 9 orders of magnitude smaller than the observed value of n_B/s (see Section 4.1). In order to avoid the "annihilation catastrophe" an unknown physical mechanism would have to operate at a temperature greater than $38\,MeV$, the temperature when $n_B/s = n_{\bar{B}}/s \simeq 7 \times 10^{-11}$,and separate nucleons and antinucleons. However, the horizon at that time only contained $\sim 10^{-7}\,M_\odot$, and so causality precludes separating out chunks even approaching a solar mass, let alone $10^{14} M_\odot$ [17]

[4]A baryon is a composite subatomic particle made up of three quarks (as distinct from mesons, which are composed of one quark and one antiquark). Baryons and mesons belong to the hadron family of particles, which are the quark-based particles.

complementary process which produces an excess of anti-baryons and so no net baryon number can be created. That is to say that the thermal average of B, which is odd under both C and CP , is zero unless those discrete symmetries are violated. The second Sakharov criterion is self-explanatory. If no processes ever occur in which B is violated, then the total number of baryons in the universe must remain constant, and therefore no asymmetry can be generated from symmetric initial conditions. Finally, the interactions must be out of thermal equilibrium, since otherwise CPT symmetry would impose compensation between processes increasing and decreasing the baryon number (see section 4.2 for more details).

The SM fails to satisfy two of the three conditions above needed for baryogenesis, namely violation of CP symmetry and deviation from thermal equilibrium. CP violation in the SM appears to be too small to explain the value of the baryon number-to-photon ratio η [2]. There is no proof that CP violation in the SM is insufficient for baryogenesis to work. However, so far, all attempts to predict the value of η with SM CP violation have failed. In particular, CP violation from the Cabibbo-Kobayashi-Maskawa (CKM) matrix has been shown to be too small to play any role in electroweak baryogenesis.

We will turn our attention to the role for baryogenesis of the CP non conserving term in the SM Quantum Chromodynamics (QCD) Lagrangian [26]

$$\mathcal{L} = \bar{\theta}\frac{\alpha_s}{8\pi}G_{\mu\nu\alpha}\tilde{G}^{\mu\nu}_{\alpha} \qquad , \qquad \bar{\theta} = \theta + \arg\det\{\mathbf{M_q}\} \tag{1.1}$$

The CP-violating $\bar{\theta}$ term is constrained today to be smaller than 10^{-11} from the absence of a measurable electric dipole moment for the neutron [6]. The θ parameter characterises the non-trivial nature of the QCD vacuum. Because chiral transformations change the θ vacuum once we include weak interactions and the quark mass matrix, the only physical observable angle is $\bar{\theta} = \theta + \arg\det\{\mathbf{M_q}\}$ where M_q is the quark mass matrix. The QCD angle θ, which is required to solve the $U(1)_A$ problem [30], and $\arg\det\{\mathbf{M_q}\}$ have nothing to do which each other and there is no reason why they should be tuned so that $|\bar{\theta}| < 10^{-11}$. This is the so-called strong CP problem [21].

The QCD vacuum energy depends on $\bar{\theta}$ and is minimised at $\bar{\theta} = 0$. Therefore the puzzle is solved if $\bar{\theta}$ is promoted to a dynamical field which relaxes naturally to zero, as postulated by Peccei and Quinn (PQ) [23]. This solution introduces a new additional chiral asymmetry $U(1)_{PQ}$ which allows to rotate the $\bar{\theta}$ parameter to zero. The symmetry is spontaneously broken by a scalar field

$$\Phi = \frac{(f_a + \rho(x))e^{i\frac{a(x)}{f_a}}}{\sqrt{2}} \tag{1.2}$$

where the Goldstone boson $a(x)$ is the *axion*. New heavy colored quarks with coupling to Φ generate a $G\tilde{G}$ term

$$\frac{\alpha_s}{8\pi}\frac{a(x)}{f_a}G_{\mu\nu\alpha}\tilde{G}^{\mu\nu}_{\alpha} \tag{1.3}$$

The axion couples to gluons, mixes with pions and couples to photons. Its couplings are all suppressed by the factor $1/f_a$ while its mass today satisfies

$$m_a f_a = m_\pi f_\pi \frac{\sqrt{m_u m_d}}{m_u + m_d} \tag{1.4}$$

where m_π is the pion mass and m_u, m_d are the up and down quark masses.

The axion $a(x)$ relaxes towards the minimum of its potential, at $\langle a \rangle = 0$, this explaining why $\bar{\theta}$ is very small today. However, in the early universe, just after $U(1)_{PQ}$ breaking, $\bar{\theta} = a(x)/f_a$ is large and frozen to a value of order 1 as long as the axion is massless. One question we want to address is whether $\bar{\theta}$ could have played any role at the time of the Electroweak (EW) phase transition (EWPT). Given that the physical effects of $\bar{\theta}$ are testified by the absence in strong interactions of the isosinglet axial symmetry and its associated Goldstone, a light pseudo scalar meson with mass comparable to pions [30], and that PQ solution is essentially the only solution to the strong CP problem, there are strong motivations for considering the role of $\bar{\theta}$ in the early universe.

This paper is organized as follows. In Section 2 we review the $U(1)$ problem: the approximate axial symmetry is known to be broken, so where is the corresponding quasi-Nambu-Goldstone boson? The answer to this question can be traced to the fact that $U(1)_A$ is broken, by virtue of the axial anomaly, to the discrete symmetry Z_{2N_f}, where N_f is the number of fermionic generations

[13]. This incomplete breaking of the axial symmetry opens the possibility that this breaking is not accompanied by a Nambu-Goldstone boson. This can be understood as follows: the rich structure of Yang-Mills vacuum corresponding to tunnelling between states of different winding number gives rise to an effective Lagrangian term proportional to θ times the Chern-Pontryagin density, which violates P and CP conservation. This implies that there is an infinite set of degenerate vacuum states, each labeled by its N_{CS}. Instanton solutions provide a mechanism of "vacuum tunneling" between topological inequivalent n-vacua. So that the "true" vacuum (the so-called θ vacuum) is a linear superposition of n-vacua and as a consequence the theory possesses a hidden parameter, the vacuum angle θ.

In Section 3, we will review the strong CP problem: we will describe the Peccei-Quinn (PQ) mechanism which introduces a new additional chiral $U(1)$ symmetry which allows to rotate the θ parameter to zero. We will see that a consequence of this mechanism is the generation of the axion, which has eluded detection so far.

Section 4 is an exposition of the basic principles of Baryogenesis. We will review the thermodynamics and all the quantities relevant for the discussion. The Sakharov conditions are exposed in detail, together with a particular solution of the electroweak field equations, the sphaleron, which play an important role in the baryon asymmetry. We will see in particular how such an asymmetry can be preserved only if the electroweak phase transition, characterised by electroweak symmetry breaking, was first-order. Finally, we will analyse a possible model for axion-induced CP violation.

In Section 5 a summary of the current state of axion search is given. Unfortunately, so far the axion has eluded detection.

We conclude in Section 6, and reserve for the Appendix a detailed discussion of topics relevant to the main subject.

2 The $U(1)$ Problem

The $U(1)$ problem is the general absence of Nambu-Goldstone particles associated with spontaneous symmetry breaking (SSB) when the relevant symmetry of the prevalent theory of the strong interactions, Quantum Chromodynamics (QCD), is spoiled by the quantum anomaly and instanton effects[5]. In fact, the Lagrangian of QCD, see Eq. (C.2),

$$\begin{aligned}\mathcal{L}_{QCD} &= \sum_q \left[-\bar{q}(\not{\partial} + m_q)q + \frac{ig_3}{2} G^\alpha_\mu \bar{q}\gamma^\mu \lambda_\alpha q \right] \\ &\quad -\frac{1}{4} G^\alpha_{\mu\nu} G^{\mu\nu}_\alpha - \frac{g_3^2 \theta}{64\pi^2} \epsilon_{\mu\nu\lambda\rho} G^{\alpha\mu\nu} G^{\alpha\lambda\rho} \end{aligned} \tag{2.1}$$

shows a chiral $U(1)$ symmetry, under which $q_i \to e^{i\theta\gamma^5} q_i$, which is not realised, or at least badly broken, in the real world: it does not seem to be reflected in the spectrum of light pseudoscalar mesons [34].

The successes of low-energy current algebra considerations strongly indicated that meson physics has an approximate chiral $SU(2)_L \times SU(2)_R$. The *pion* π^0 can then be regarded as being the Goldstone bosons associated with this symmetry. It is easy to incorporate this symmetry in the QCD Lagrangian, simply by postulating that the u and the d quarks must have very small mass terms here. The problem one then encounters is that, if this were the case, QCD should actually have an even larger symmetry: $U(2)_L \times U(2)_R$ which differs from the observed symmetries by an extra chiral $U(1)$ component, and this should be reflected in a (partially) conserved isoscalar axial vector current, j^A_μ. Thus, the symmetry held responsible for the relatively small value of the pion masses, should necessarily induce another symmetry in the model that would strongly reduce the mass of yet another particle: the pions should have had a pseudoscalar partner, somewhat like the η but composed predominantly of $u\bar{u}$ and $d\bar{d}$ quarks, in the combination $(u\bar{u} + d\bar{d})/\sqrt{2}$. Moreover, with masses included, chiral perturbation theory unambiguously predicts a neutral pseudoscalar meson whose mass is strictly less than $\sqrt{3}m_\pi$. However, the true hadron spectrum contains only the regular π^0 (140 MeV), the η (549 MeV), and the η' (957 MeV), so the chiral perturbation theory bound is clearly violated [31].

Now although it was soon established that the corresponding current conservation law is formally violated by quantum effects due to the chiral anomaly (see Appendix A.2) it was for some time a mystery how effective $U(1)$ violating interactions could take place to realise this violation, in particular because a less trivial variant of chiral $U(1)$ symmetry still seemed to exist [30]. Indeed, all perturbative calculations showed a persistence of the $U(1)$ invariance.

2.1 A proposed solution

With the discovery of instantons, see Appendix A.5, and the form the chiral anomaly takes in these non-perturbative field configurations, a solution to the so-called $U(1)$ problem was proposed [29]. It was now clear how entire units of axial $U(1)$ charge could appear or disappear into the vacuum without the need of (nearly) massless Goldstone bosons. In a world without instantons the η and η' particles would play the role of Goldstone bosons. Now the instantons provide them with an anomalous contribution to their masses.

The starting point is the solution of classical field equations in four-dimensional ($4D$) Euclidean gauge-field theories. The solution is obtained from the vacuum by mapping $SU(2)$ gauge transformations onto a large sphere in Euclidean space. Taking the new, gauge-rotated, vacuum as a boundary condition, one obtains a nontrivial solution inside the sphere, characterised by a topological quantum number. If the Lagrangian is

$$\begin{aligned}&\mathcal{L}_{YM} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}, \quad a = 1,2,3 \\ &G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\epsilon_{abc} A^b_\mu A^c_\nu \end{aligned} \tag{2.2}$$

then the topological quantum number (compare with Eq. (A.62)) is

[5] For an overview of Spontaneous Symmetry Breaking and the Goldstone's Theorem see [27]

$$N_{CS} = \frac{g^2}{32\pi^2}\int \mathrm{d}^4x\, G^a_{\mu\nu}\tilde{G}^{a\mu\nu} \tag{2.3}$$

with

$$\tilde{G}^a_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}G^{a\alpha\beta} \tag{2.4}$$

is an integer for all field configurations in Euclidean space that have the vacuum (or a gauge transformation thereof) at the boundary. In Minkowsky space n would be i times an integer. The solution (instanton) with $n = 1$ in Euclidean space is

$$A^a_\mu(x)^{cl} = \frac{2}{g}\frac{\eta_{a\mu\nu}(x-x_0)^\nu}{(x-x_0)^2+\lambda^2} \tag{2.5}$$

Here, x_0 is free because of translation invariance and λ is a free scale parameter; η is a tensor that maps antisymmetric representations of $SO(4)$ onto vectors of one of its two invariant subgroups $SO(4)$:

$$\begin{array}{rcllll} \eta_{a\mu\nu} & = & \epsilon_{a\mu\nu} & \text{for} & a,\mu,\nu & = 1,2,3 \\ \eta_{a4\nu} & = & -\delta_{a\nu} & \text{for} & a,\nu & = 1,2,3 \\ \eta_{a\mu 4} & = & \delta_{a\mu} & \text{for} & a,\mu & = 1,2,3 \\ \eta_{a44} & = & 0 & & & \end{array}$$

Thus isospin is linked to one of the $SO(3)$ subgroups of $SO(4)$. The solution has

$$S = \int \mathrm{d}^4x\,\mathcal{L}\left(A^{cl}\right) = -\frac{8\pi^2}{g^2} \tag{2.6}$$

Since we have a 4D rotational symmetry, the solution is not only localised in three-space, but also instantaneous in time. These solutions of the Euclidean field equations are relevant for describing a tunnelling mechanism in real (Minkowsky) space-time, from one vacuum state to a gauge-rotated vacuum (a gauge rotation that cannot be obtained via a series of infinitesimal gauge rotations).

Suppose now that we have in addition N massless fermion doublets coupled to the gauge field:

$$\mathcal{L}_{fermion} = -\sum_{t=1}^{N}\bar{\psi}^t \not{D}\psi^t \tag{2.7}$$

The axial vector current j^5_μ produced by the QCD Lagrangian has an anomaly, see section A.4

$$\partial^\mu j^5_\mu = -i\frac{N_f g^2}{16\pi^2}G^a_{\mu\nu}\tilde{G}^{a\mu\nu} \tag{2.8}$$

Let us compare this with Eq. (2.3). We see that in a volume V_3

$$\int_{V_3}\mathrm{d}^4x\,\partial^\mu j^5_\mu = -2iN_f N_{CS} \tag{2.9}$$

Defining the charge Q_5 in a volume V_3 by

$$Q_5 = \int_{V_3}\mathrm{d}^3x\, j^5_0 = i\int_{V_3}\mathrm{d}^3x\, j^5_4 \tag{2.10}$$

we can see that a configuration in Minkowsky space with $n = 1$ would be associated with a violation of axial charge conservation:

$$\Delta Q^5 = 2N_f \tag{2.11}$$

Let us now write the vacuum to vacuum amplitude in QCD. To calculate the amplitude for such an event directly in Minkowsky space one needs more understanding of the quantum mechanical tunnelling from one vacuum to the gauge-rotated vacuum. In practice it is much easier to make use of the explicit solution in Euclidean space. Let us assume then that all Green's functions in Minkowsky space can simply be obtained from the Euclidean ones by analytic continuation. After some manipulation, we obtain

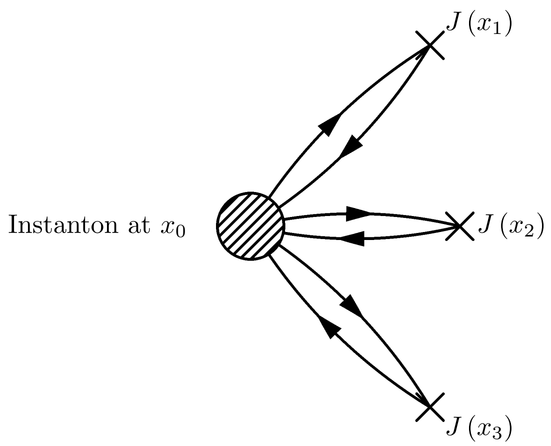

Figure 1: The sources J turn the axial charge Q^5 into $-Q^5$ for each flavor. The amplitude goes like J^3

$$\langle 0|0\rangle = \int \mathrm{d}^4x_0 \int \mathrm{d}\lambda \,(\det J)(\det M_1)^{-\frac{1}{2}}(\det M_2)(\det M_3)\exp \int \mathrm{d}^4x\, \mathcal{L}\big(A^{cl}\big) \tag{2.12}$$

where $(\det J)$ is the Jacobian following from the transition to some convenient coordinate system and the matrices M_1, M_2 and M_3 are connected to the gauge, fermion and Faddeev-Popov ghost term respectively[6]. M_1 and M_3 have some zero eigenvalues that neatly cancel. But M_2 has zero eigenvalues that are not cancelled by anything, so $\det M_2 = 0$ and the amplitude (2.12) vanishes. From Eq. (2.6) it follows that the exponent equals $\exp\left(-8\pi^2/g^2\right)$ which is a term that is unobtainable through ordinary perturbation expansion.

Let us now insert a source term $J\bar{\phi}\phi$ into the Lagrangian, where $J(x)$ may contain flavor indices and γ matrices, but must be gauge invariant. Now the lowest eigenvalues will be come different from zero. Now we find that

$$\det\left(M_2 + J\right) \simeq \prod_{i=1}^{N} \left(x_i - x_0\right)^{-6} J(x_i) \tag{2.13}$$

for large distances. The factor $\prod_{i=1}^{N}\left(x_i - x_0\right)^{-6}$ is exactly reproduced by the $2N$ propagators that connect the sources with the instanton (See Fig. 1). Note that the sources have to switch chirality. This explains why the instanton gives Eq. (2.11).

It is now possible to build an effective vertex that could mimic the same amplitude

$$\mathcal{L}^{eff}(x) = \kappa e^{i\theta_{inst}} \det\left[-\bar{\psi}_R(x)\psi_L(x)\right] \quad + \quad \text{h.c} \tag{2.14}$$

Here, κ is a constant that should be in principle computable, and it contains the factor $e^{-8\pi^2/g^2}$. The subscripts L and R refer to the left- and right handed helicities, obtained by means of the projection operators $\frac{1}{2}\left(1 \pm \gamma^5\right)$. The determinant is the determinant of the matrix $\bar{\psi}_R^b\psi_L^a$ here a and b are flavor indices only. We see that the interaction (2.14) has exactly the right quantum numbers for absorbing N_{flavor} left helicity fermions and creating an equal number of right handed ones (or vice-versa). In particular, the determinant is easily seen to be the simplest possible interaction that conserves $SU(N_f) \times SU(N_f)$ symmetry, while breaking $U(N_f) \times U(N_f)$. θ_{inst} is a new phase angle which emerges in the description of interference between the instanton and mass quark symmetry breaking terms. The cases $N_f = 0$ and $N_f = 1$ are rather special. If $N_f = 0$ while $\theta_{inst} \neq 0$, we see the appearance of an explicit P and CP violation term in the QCD Lagrangian. This will be discussed at length in the next section.

In a color gauge theory for strong interactions with two massless quark triplets, Eq. (2.14) is an effective four-fermion interaction with exactly the chiral quantum numbers of a mass term for

[6]For the details of the derivation, see [28]

the η particle. For simplicity, we can study the case $N_f = 2$, which gives us the physics of QCD if we allow ourselves to neglect the effects of the strange quarks. A low energy effective meson model for QCD with instanton effects included is described in [30, 31]. The effective meson fields q_{ij} basically correspond to the composite operators $\bar{q}_{Rj} q_{Li}$, and this 2×2 matrix is decomposed into eight real mesonic fields: a scalar isoscalar σ , a pseudoscalar isoscalar η, a scalar isovector $\vec{\alpha}$, and a pseudoscalar isovector $\vec{\pi}$. It can be shown that if the u and the d quark masses may be neglected then θ_{inst} will be aligned to zero, and the effective coupling goes as

$$\mathcal{L}^{eff} \to 2\kappa\left(\sigma^2 + \vec{\pi}^2 - \eta^2 - \vec{\alpha}^2\right) \tag{2.15}$$

Now, since this is the only effect that splits the pion from the eta, and since the pion continues to behave as a massless Goldstone boson, one can deduce from Eq. (2.15) that the eta mass becomes

$$m_\eta^2 = 8\kappa \tag{2.16}$$

This analysis shows that the instanton interaction bares exactly the quantum numbers required for the eta mass term.

3 The Strong CP Problem

The possible resolution for the $U(1)$ problem given in section 2.1 came through the realisation that the QCD vacuum has a quite richer structure then expected. The more complex nature of the QCD vacuum, in effect, makes $U(1)_A$ not a true symmetry of QCD, even though it is an apparent symmetry of the QCD Lagrangian in the limit of vanishing quark masses. However, associated with this more complicated QCD vacuum there is a phase parameter θ and only if this parameter is very small is CP not very badly broken in the strong interactions. So the solution of the $U(1)_A$ problem begets a different problem: why is CP not badly broken in QCD? This is known as the strong CP problem [21].

Consider any of the field configurations described in Eq. (A.59). Neglecting tunnelling effect we might expect the vacuum to be of the form

$$\psi_n[A] = \phi[A - A_n] \tag{3.1}$$

where the wave functional ϕ is peaked about zero and has a spread due to quantum fluctuations and any A_n be chosen as representative of the classical vacuum, i.e., the classical zero-energy configuration. But the pseudoparticle solution connects A_n with A_{n+1} giving origin to tunnelling between the different ψ_n [16]. The true quantal vacuum state will therefore be a superposition of the form

$$\Psi[A] = \sum_n c_n \psi_n[A] + \mathcal{O}\left[\exp\left(-\frac{8\pi^2}{g^2}\right)\right] \tag{3.2}$$

To determine the coefficients c_n in this equation let us observe that the finite gauge transformation $g(x)$ defined in Eq. (A.67), changes ψ_n into ψ_{n+1}. Requiring the vacuum state to be stable against gauge transformations determines the coefficients to be

$$c_n = e^{in\theta} \tag{3.3}$$

Thus we find a family of vacua, parametrised by an angle θ, where under the instanton gauge transformation we have

$$\Psi_\theta[A] \underset{g(x)}{\rightarrow} e^{-i\theta}\Psi_\theta[A] \tag{3.4}$$

So that the true vacuum is a superposition of these, so-called, n-vacua and is called the θ-vacuum:

$$|\theta\rangle = \sum_n e^{-in\theta}\,|n\rangle \tag{3.5}$$

One can write for the vacuum to vacuum transition amplitude

$$\langle\theta|\theta\rangle = \sum_{m,n} e^{im\theta}e^{-in\theta}\,\langle m|n\rangle = \sum_k e^{ik\theta}\sum_n \langle n+k|n\rangle \tag{3.6}$$

Now, one can write (see Eq. (A.62))

$$\begin{aligned} \frac{g_3^2}{64\pi^2}\epsilon^{\mu\nu\lambda\beta}G^\alpha_{\mu\nu}G^\alpha_{\lambda\beta} &= \partial_\mu K^\mu \\ K^\mu &\equiv \frac{g_3^2}{64\pi^2}\epsilon^{\mu\nu\lambda\beta}\left(G^\alpha_\nu G^\alpha_{\lambda\beta} - \frac{g}{3}f_{\alpha\beta\gamma}G^\alpha_\nu G^\beta_\alpha G^\gamma_\beta\right) \end{aligned} \tag{3.7}$$

Because of this property, the space-time integral of the term multiplying θ is determined by the change in the charge $\int \mathrm{d}^3x\, K^0 \equiv N_{CS}$ between initial and final surfaces (see Eq. (2.3))

$$\begin{aligned} \frac{g_3^2}{64\pi^2}\int \mathrm{d}^3x\int_{t_i}^{t_f}\mathrm{d}t\,\epsilon^{\mu\nu\lambda\beta}G^\alpha_{\mu\nu}G^\alpha_{\lambda\beta} &= \int \mathrm{d}^3x\,K^0(x,t_f) - \int \mathrm{d}^3x\,K^0(x,t_i) \\ &\equiv N_{CS}(t_f) - N_{CS}(t_i) \end{aligned} \tag{3.8}$$

where we assume boundary conditions which ensure the vanishing of contributions from spatial infinity.

Now, we can write Eq. (3.6) as

$$\langle\theta|\theta\rangle = \sum_k \int \delta A e^{iS_{eff}[A]} \delta\left[k - \frac{g_3^2}{64\pi^2}\int \mathrm{d}^4x\, \epsilon^{\mu\nu\lambda\beta} G^\alpha_{\mu\nu} G^\alpha_{\lambda\beta}\right] \tag{3.9}$$

where

$$\begin{aligned} S_{eff}[A] &= S_{QCD}[A] + \theta \frac{g_3^2}{64\pi^2}\int \mathrm{d}^4x\, \epsilon^{\mu\nu\lambda\beta} G^\alpha_{\mu\nu} G^\alpha_{\lambda\beta} \\ &= S_{QCD}[A] + \theta N_{CS} \end{aligned} \tag{3.10}$$

Naive perturbation theory includes only the $k = 0$ term of the infinite sum in (3.9).

The resolution of $U(1)$ problem, by recognising the complicated nature of the QCDs vacuum, effectively adds and extra term to the QCD Lagrangian (see Eq. (C.2)):

$$\mathcal{L}_\theta = \theta \frac{g_3^2}{64\pi^2} \epsilon^{\mu\nu\lambda\beta} G^\alpha_{\mu\nu} G^\alpha_{\lambda\beta} \tag{3.11}$$

This term violates Parity and Time reversal invariance, but conserves Charge conjugation invariance, so it violates CP (see Appendix B). In fact, from Eq. (3.10) we see that the S-matrix element is periodic in θ with modulus 2π

$$\langle\theta + 2\pi|\theta + 2\pi\rangle = \langle\theta|\theta\rangle \tag{3.12}$$

Since the θ vacuum is no longer P- and T-invariant, one can choose the phase of P and T operators so that

$$P\,|\theta\rangle = |-\theta\rangle\,, \quad T\,|\theta\rangle = |-\theta\rangle \tag{3.13}$$

Setting $\theta = \pi$ and using Eq. (3.12), it follows that

$$\langle\pi|\pi\rangle = \langle\pi|T^{-1}T|\pi\rangle\,, \quad \langle\pi|\pi\rangle = \langle\pi|P^{-1}P|\pi\rangle \tag{3.14}$$

Consequently, for pure gauge theory $\theta = \pi$ also describes a P-and T-conserving world owing to the periodicity of θ. Thus if strong CP violation is very weak, a priori the world $\theta \approx \pi$ is as possible as the world $\theta \approx 0$. Nevertheless, as we shall see later, $\theta \approx \pi$ is ruled out based on current-algebra arguments [7]. Also, present constraints on the size of the neutron electric dipole moment imply that it must be smaller than $\sim 10^{-10}$. This extremely small magnitude ensures that low-energy QCD is approximately P and CP conserving: the puzzle of why the value should be so small is related to the strong CP problem.

If, besides QCD, one includes the weak interactions, the quark mass matrix is in general complex, nondiagonal and may contain γ^5 terms:

$$L_{mass} = \bar{q}_{iR} M_{ij} q_{jL} \quad + \quad \text{h.c} \tag{3.15}$$

To go to a physical basis one must diagonalize this mass matrix and when one does so, in general, one performs a chiral rotation of the quark field $q \to \exp\left(i\alpha\gamma^5\right) q$ which changes the phase of $\det M \to \det M + 2\alpha$. Due to the chiral anomaly (see Appendix A), θ also changes: $\theta \to \theta - 2\alpha$. So, in the total theory, the physical and measurable strong CP violation parameter is

$$\bar{\theta} = \theta + \arg\det M \tag{3.16}$$

and this is invariant under chiral transformations of the quark fields. The strong CP problem is: why is this $\bar{\theta}$ angle, coming from the strong and weak interactions, so small?

What is wrong with simply choosing $\bar{\theta} = 0$? After all, it is consistent simply to choose the value of $\bar{\theta}$ to be small within the standard model, and that phenomenologically-small values for $\bar{\theta}$ do not pose a naturalness problem for the standard model itself. However, such a small value is rather surprising. This is mainly a theorist's problem since only a very few phenomena can manifest the elusive strong CP-odd effects, e.g. the electric dipole moment of the neutron and $\eta \to 2\pi$ decays. The strong CP predicament can be rephrased in several different ways. It is a CP hierarchy problem: Weak CP violation in $K^0 - \bar{K}^0$ systems is characterized by the ϵ parameter which is of order of 10^{-3},whereas strong CP nonconservation is measured by the enormously small parameter $\bar{\theta}$. It is very difficult to construct realistic models which can generate adequate weak CP violation responsible for $K_L \to 2\pi$ decays and which meanwhile have sufficiently small nonconservation of strong CP. It is a problem of naturalness: according to 't Hooft's principle of naturalness, a physical parameter can be naturally small if putting it equal to zero increases the symmetry. A well-known example is the smallness of the electron mass which is attributed to the approximate chiral symmetry. The strong CP problem comes about because in most models letting $\bar{\theta} = 0$ does not correspond to an enlarged symmetry of the theory since CP is still violated in the weak interactions; hence the smallness of $\bar{\theta}$ is not protected by a symmetry [7]. It is also a problem of fine-tuning in the sense that θ is a free parameter, and in QCD there is no reason why it should take the value $-\arg\det M$.

3.1 Proposed solution: a new Symmetry

A possible solution is that $\mathcal{L}$ must possess a chiral $U(1)$ invariance, such that changes in θ are equivalent to changes in the definitions of the various fields in $\mathcal{L}$ and have no physical consequences. Any such theory is equivalent to a $\theta = 0$ theory and this has no strong P and CP violations. This property is trivially true for theories where $\mathcal{L}$ represents a non-Abelian gauge field coupled only to massless fermions. In fact, the rotation of a fermion field by $\exp\left(i\gamma^5\sigma\right)$ induces a change in the effective action (3.10)

$$\delta S_{eff}[A] = -2i\int \mathrm{d}^4x\,(\partial^\mu\gamma^5_\mu)\sigma = -iN_{CS}\sigma \tag{3.17}$$

where we used Eq. (A.37). Thus in such a theory the net effect of such a rotation is

$$\theta \to \theta' = \theta - 2\sigma \tag{3.18}$$

If, however, all fermions are massive such a rotation will also change the fermion mass term in $\mathcal{L}$. Hence one can define inequivalent theories which have the same mass term and various choices of θ. Only one such class of theories, those in which $\theta \to 0$ when all fermion masses have been made real by a suitable $\exp\left(i\gamma^5\sigma\right)$ rotation, yield a CP- and P-invariant theory of the strong interactions. Which is the aforementioned strong CP problem, indeed.

A theory was proposed [23] such that the above invariance property remains true when some fermion masses are included in $\mathcal{L}$, or even when all strongly interacting fermions become massive, provided that at least one fermion gets its entire mass from a Yukawa coupling G_F to a scalar field, so that the full $\mathcal{L}$ can possess at least a single chiral $U(1)$ invariance. To exemplify the theory, a simplified model of the strong interactions is used, in which there is only a single fermion flavor and a single color-singlet complex scalar field ϕ.

$$\begin{aligned}\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + i\bar{\psi}\not{D}\psi + \bar{\psi}\left[G_F\phi\left(\frac{1+\gamma^5}{2}\right) + G_F^*\phi^*\left(\frac{1-\gamma^5}{2}\right)\right]\psi \\ - \left|\partial_\mu\phi\right|^2 - \mu^2\left|\phi\right|^2 - h\left|\phi\right|^4\end{aligned} \tag{3.19}$$

with $\mu^2 < 0$. We note that Eq. (3.19) is formally invariant under the chiral rotation[7]

$$\begin{aligned}\psi &\to \exp\left(i\sigma\gamma^5\right)\psi \\ \phi &\to \exp\left(-2i\sigma\right)\phi\end{aligned} \tag{3.20}$$

for which the same considerations as Eq. (3.17) apply: in this theory a chiral rotation redefines the θ parameter as Eq. (3.18). This demonstrates that for any θ value we obtain an equivalent theory. To show that these theories are CP conserving we also need

$$\alpha = \arg\left(e^{i\theta}G_F\left\langle\phi\right\rangle\right) = 0 \tag{3.21}$$

which, when $\langle\phi\rangle \neq 0$, corresponds to requiring that the fermion mass $G_F\langle\phi\rangle$ be real when the fields are defined so that $\theta = 0$. The proof that Eq. (3.21) is satisfied is quite convoluted. For the

[7]This is dubbed the Peccei-Quinn symmetry $U(1)_{PQ}$ from the names of the proposers

details see [22]. Here we give just a summary. We start examining the generating functional of the scalar Green's functions

$$Z_\theta(J,J^*) = \sum_q e^{i\theta_q} \int (\mathrm{d}A_\mu)_q \int \mathrm{d}\psi\, \mathrm{d}\bar{\psi}\, \mathrm{d}\phi\, \mathrm{d}\phi^* \exp\left[\int \mathrm{d}^4x\, (\mathcal{L} + J\phi + J^*\phi^*)\right] \tag{3.22}$$

The scalar vacuum expectation value is defined by

$$\langle\phi\rangle = \frac{1}{Z_\theta}\frac{\delta Z_\theta}{\delta J}\bigg|_{J=J^*=0} = \lambda e^{i\beta} \tag{3.23}$$

where λ and β are real constants. The proof proceeds as follows

1. We make the change of variables

$$\phi = e^{i\beta}(\lambda + \rho + i\sigma) \tag{3.24}$$

   where ρ and σ are real scalar fields.

2. We use the knowledge that only terms of chirality n contribute in each n sector to allow us to formally integrate out vector and fermion fields and obtain an expression for Z in terms of nonlocal polynomials of the scalar fields.

3. Using only known reality properties of the polynomials we can write the constraints $\langle\rho\rangle = \langle\sigma\rangle = 0$. We find that they require $\alpha = 0, \pi$. These are stationary points of the scalar potential. To find which is the true minimum we must examine the potential itself.

After some manipulation we can re-write Eq. (3.22) as

$$\begin{aligned} Z_\theta(J,J^*) = \int \mathrm{d}\rho\, \mathrm{d}\sigma \left\{ A_0(\rho,\sigma^2) + \sum_n \left[F_n(\rho,\sigma^2)\cos n\alpha - \sigma G_{Fn}(\rho,\sigma^2)\sin n\alpha\right]\right\} \\ \times \exp\left[Je^{i\beta}(\lambda+\rho+i\sigma) + J^* e^{-i\beta}(\lambda+\rho-i\sigma)\right] \end{aligned} \tag{3.25}$$

where F_n and G_{Fn} are the real and imaginary parts of $A_n(\phi\phi^*)\,|G_F|^n\,(\lambda+\rho+i\sigma)^n$, respectively. Now we impose the constraints that, by definition, the field ρ and σ have vanishing vacuum expectation values. This gives us

$$\begin{aligned} \langle\rho\rangle &= \int \mathrm{d}\rho\, \mathrm{d}\sigma\, \rho \left(A_0 + \sum_n F_n \cos n\alpha\right) = 0 \\ \langle\rho\rangle &= \int \mathrm{d}\rho\, \mathrm{d}\sigma\, \sigma^2 \sum_n G_{Fn} \sin n\alpha = 0 \end{aligned} \tag{3.26}$$

The first of these equations is satisfied for arbitrary α by appropriately choosing $\lambda = \lambda(\alpha)$. The second equation can then in general only be satisfied for $\alpha = 0, \pi$. These values are both stationary points of the potential. To find which is the true minimum we must examine $V(\phi)$. This we can only do to leading order in the Yukawa couplings G_F and h, for which we find

$$V(\phi) = \mu^2\phi^*\phi + h(\phi^*\phi)^2 - K\,|\phi|\cos\alpha \tag{3.27}$$

where K is a real positive constant. Thus if G and h are sufficiently small the minimum occurs at $\alpha = 0$. The interaction term in Eq. (3.19), see Fig. 2, becomes

$$\lambda\bar{\psi}\left[G_F e^{i\beta}\left(\frac{1+\gamma^5}{2}\right) + G_F^* e^{i\beta}\left(\frac{1-\gamma^5}{2}\right)\right]\psi \tag{3.28}$$

This mass can be made real by rotating the fermion fields by $\exp\left(i\gamma^5\sigma/2\right)$. Such a rotation gives (see Eq. (3.18))

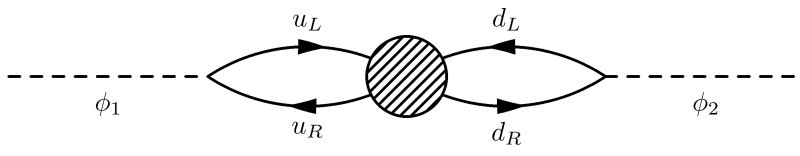

Figure 2: Instanton interaction generating the axion mass.

$$\theta \to \theta' = \theta - \theta = 0 \tag{3.29}$$

which is the sought result.

There are possible solutions to the string CP problem other than the Peccei-Quinn symmetry [7]. One possibility is a global $U(1)$ symmetry implemented by a vanishing mass of any light quark. The most natural choice, $M_u = 0$, is disfavoured since it is inconsistent with the observed meson and baryon masses. Another possibility is based on supersymmetry. However, this is problematic since a logarithmic divergence which in general occurs in the higher loop correction cannot prevent $\bar{\theta}$ from being infinite.

On the other hand the PQ symmetry solution is appealing since it is compatible with any mechanism of weak CP violation: a QCD theory can be built that includes weak and electromagnetic interactions. Obviously, one must arrange things so that there is no possibility of obtaining strong CP violations arising through the (strong) pseudoparticle sectors. This requires that one violate CP in the weak Lagrangian in such a way that the PQ symmetry remains effective. An example is the Weinberg model of CP violation which needs at least three different Higgs doublets to have CP-violation. However, this theory does not have strong P and CP conservation because it also lacks a chiral $U(1)$ symmetry that allows us to change θ. The situation is, however, easily remedied. By adding a fourth scalar doublet, PQ invariance may be imposed in the model, which would guarantee strong CP conservation, while Weinberg's theory is still rich enough in parameters to cause weak CP violations [22].

3.2 The Axion

We have seen in the previous section that the solution to the strong CP problem was brough by the realisation that the quark-mass matrix is a function $G_F \langle\phi\rangle$ of the vacuum expectation values of a set of weakly coupled scalar fields ϕ_i, see Eq. (3.21). Although θ is arbitrary, $\langle\phi\rangle$ is not: it is determined by the minimization of the potential $V(\phi)$ of Eq. (3.27). The phase of $G_F \langle\phi\rangle$ at the minimum of $V(\phi)$ is then undetermined in any finite order of perturbation theory, and is fixed only by instanton effects which break the $U(1)_{PQ}$ symmetry. Hence the spontaneous breakdown of the chiral $U(1)_{PQ}$ symmetry associated with the appearance of nonzero vacuum expectation values $\langle\phi\rangle$ leads to a very light pseudoscalar pseudo-Goldstone boson, the "axion", with m_a^2 proportional to the Fermi coupling G_F [35].

So, the introduction in the theory of the $U(1)_{PQ}$ symmetry effectively replaces the static CP-violating angle $\bar{\theta}$ with a dynamical CP-conserving field, the axion. As a result, under a $U(1)_{PQ}$ transformation, the axion field $a(x)$ translates

$$a(x) \to a(x) + \alpha f_a \tag{3.30}$$

where f_a is the order parameter associated with the breaking of $U(1)_{PQ}$. Formally, to make the Lagrangian of the Standard Model $U(1)_{PQ}$-invariant this Lagrangian, Eq. (C.1), must be augmented by axion interactions [21]

$$\mathcal{L}_{total} = \mathcal{L}_{SM} + \mathcal{L}_a \tag{3.31}$$

with

$$\mathcal{L}_a = -\frac{1}{2}\partial_\mu a \partial^\mu a + \mathcal{L}_{int}[\partial^\mu a/f_a; \Psi] + \xi \frac{a}{f_a}\frac{g_3^2}{32\pi^2}\epsilon^{\mu\nu\lambda\beta}G^\alpha_{\mu\nu}G^\alpha_{\lambda\beta} \tag{3.32}$$

The last term above is needed to ensure that the $U(1)_{PQ}$ current indeed has a chiral anomaly:

$$\partial_\mu J^\mu_{PQ} = \xi \frac{g_3^2}{32\pi^2}\epsilon^{\mu\nu\lambda\beta}G^\alpha_{\mu\nu}G^\alpha_{\lambda\beta} \tag{3.33}$$

where ξ is the color anomaly of the PQ symmetry. This term also represents an effective potential for the axion field, and its minimum occurs at $\langle a \rangle = -f_a\bar{\theta}/\xi$

$$\left\langle \frac{\partial V_{eff}}{\partial a} \right\rangle = -\frac{\xi}{f_a}\frac{g_3^2}{32\pi^2}\left\langle \epsilon^{\mu\nu\lambda\beta}G^\alpha_{\mu\nu}G^\alpha_{\lambda\beta} \right\rangle \Big|_{\langle a \rangle = -\frac{f_a}{\xi}\bar{\theta}} = 0 \tag{3.34}$$

Since at the minimun the $\bar{\theta}$-term is cancelled out, this provides a dynamical solution to the strong CP problem.

It is easy to understand the physics of this solution to the strong CP problem. If one neglects the effects of QCD then the extra $U(1)_{PQ}$ symmetry introduced allows all values for $\langle a \rangle$ to exist: $0 \leq \langle a \rangle \leq 2\pi$. However, including the effects of the QCD anomaly serves to generates a potential for the axion field which is periodic in the effective vacuum angle $\bar{\theta} + \xi \langle a \rangle / f_a$

$$V_{eff} \equiv \cos\left(\bar{\theta} + \xi\frac{\langle a \rangle}{f_a}\right) \tag{3.35}$$

the Lagrangian (3.32) written in terms of $a_{phys} = a - \langle a \rangle$ no longer has a CP violating $\bar{\theta}$-term. Expanding V_{eff} at the minimum gives the axion a mass, which is generated entirely by processes such as shown in Fig. 2

$$m_a^2 = \left\langle \frac{\partial^2 V_{eff}}{\partial a^2} \right\rangle = -\frac{\xi}{f_a}\frac{g_3^2}{32\pi^2}\frac{\partial}{\partial a}\left\langle \epsilon^{\mu\nu\lambda\beta}G^\alpha_{\mu\nu}G^\alpha_{\lambda\beta} \right\rangle \Big|_{\langle a \rangle = -\frac{f_a}{\xi}\bar{\theta}} \tag{3.36}$$

In [36], counting the visible coupling constants and supposing that instanton effects are characterized by a typical strong-interaction scale $\mu = 200\, MeV$, an order-of-magnitude estimate is given of

$$m_a \approx G_F^{1/2}\mu^2 \approx 100\, keV \times 10^{\pm 1} \tag{3.37}$$

4 Baryogenesis

The interactions that violate baryon number (B) are today very weak as evidenced by the longevity of the proton. They can be characterised by a coupling constant analogous to the Fermi coupling constant, but at least 25 orders of magnitude smaller:

$$G_{\Delta B} \sim M^{-2} \lesssim 10^{-30}\, GeV^{-2} \tag{4.1}$$

where M is of the order of the energy scale of unification, which is likely to be $10^{14}\, GeV$ or greater. However, at temperatures comparable to, or greater than, M, B-violating forces (if they exist) should have strength comparable to all the other interactions of Nature. These interactions can allow a baryon-symmetric Universe to evolve a baryon asymmetry of the magnitude required to explain the present baryon-to-photon ratio [17].

4.1 Thermodynamics

The Universe has for much of its history been very nearly in thermal equilibrium. However, the departures from equilibrium have been very important. Without them, the past history of the Universe would be irrelevant, as the present state would be merely that of a system at $2.75\, K$. To properly understanding the thermal history of the Universe we can compare the particle interaction rates and the expansion rate. Ignoring the temperature variation of g_* (see Eq. (4.14)), for this discussion, $T \propto R^{-1}$ and the rate of change of the temperature $\dot{T}/T$ is just set by the Hubble expansion rate: $\dot{T}/T = -H$. So long as the interactions necessary for particle distribution functions to adjust to the changing temperature are rapid compared to the expansion rate, the Universe will, to a good approximation, evolve through a succession of nearly thermal states with temperature decreasing as R^{-1}. A useful rule of thumb is that a reaction is occurring rapidly enough to maintain thermal distributions when $\Gamma \gtrsim H$, where Γ is the interaction rate per particle, $\Gamma \equiv n\sigma|v|$. Here n is the number density of target particles and $\sigma|v|$ is the cross section for interaction times relative velocity (appropriately averaged). We will use

$$\begin{aligned} \Gamma &\gtrsim H\,(\text{coupled}) \\ \Gamma &\lesssim H\,(\text{decoupled}) \end{aligned} \tag{4.2}$$

as the criterion for whether or not a species is coupled to (decoupled from) the thermal plasma in the Universe. So, since throughout most of the history of the Universe (in particular the early Universe) the reaction rates of particles in the thermal bath, Γ_{int}, were much greater than the Hubble expansion rate, H, and local thermal equilibrium (LTE) should have been maintained. In this case the entropy per comoving volume element remains constant. The entropy in a comoving volume provides a very useful and reliable quantity during the expansion of the Universe [17, 15].

In the expanding Universe, the second law of thermodynamics, as applied to a comoving volume element of unit coordinate volume[8] and physical volume $V = R^3$, implies that

$$T\,\mathrm{d}S = \mathrm{d}(\rho V) + p\,\mathrm{d}V = \mathrm{d}[(\rho + p)V] - V\,\mathrm{d}p \tag{4.3}$$

where ρ and p are the equilibrium energy density and pressure. Now, from

$$\mathrm{d}S = \frac{\partial S}{\partial T}\,\mathrm{d}T + \frac{\partial S}{\partial V}\,\mathrm{d}V \tag{4.4}$$

and comparing with (4.3), we can see that

$$\frac{\partial}{\partial V}\left(\frac{\partial S}{\partial T}\right) = \frac{1}{T}\frac{\mathrm{d}\rho}{\mathrm{d}T} \tag{4.5}$$

and

$$\frac{\partial}{\partial T}\left(\frac{\partial S}{\partial V}\right) = \frac{\mathrm{d}}{\mathrm{d}T}\left(\frac{\rho + p}{T}\right) = -\frac{1}{T^2}(\rho + p) + \frac{1}{T}\frac{\mathrm{d}}{\mathrm{d}T}(\rho + p) \tag{4.6}$$

imposing the integrability condition[9]

[8] For simplicity we will take the comoving volume to be of unit coordinate volume.

[9] Equivalent to requiring that $\mathrm{d}S$ is an exact differential

$$\frac{\partial^2 S}{\partial T \partial V} = \frac{\partial^2 S}{\partial V \partial T} \tag{4.7}$$

we get

$$\mathrm{d}p = \frac{\rho + p}{T}\,\mathrm{d}T \tag{4.8}$$

Finally substitute into Eq. (4.3) to obtain

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}[(\rho+p)V] - (\rho+p)V\frac{\mathrm{d}T}{T^2} = \mathrm{d}\left[\frac{(\rho+p)V}{T} + \mathrm{const}\right] \tag{4.9}$$

That is, up to an additive constant, the entropy per comoving volume is $S = R^3(\rho+p)/T$. Recall that the first law (energy conservation) can be written as

$$\mathrm{d}[(\rho+p)V] = V\,\mathrm{d}p \tag{4.10}$$

Substituting Eq. (4.8) into Eq. (4.10), it follows that

$$\mathrm{d}\left[\frac{(\rho+p)V}{T}\right] = 0 \tag{4.11}$$

This result implies that in thermal equilibrium, the entropy per comoving volume, S, is conserved[10]. It is useful to define the entropy density s

$$s = \frac{S}{V} = \frac{\rho+p}{T} \tag{4.12}$$

The entropy density is dominated by the contribution of relativistic particles, so that to a very good approximation

$$s = \frac{2\pi^2}{45} g_{*S} T^3 \tag{4.13}$$

where

$$g_{*S} = \sum_{i=bosons} g_i \left(\frac{T_i}{T}\right)^3 + \frac{7}{8} \sum_{i=fermions} g_i \left(\frac{T_i}{T}\right)^3 \tag{4.14}$$

is the effective number of massless degrees of freedom. The relative factor of 7/8 accounts for the difference in Fermi and Bose statistics. We also note that s is proportional to the number density of relativistic particles, and that in particular s is related to the photon number density, $s = 1.80\, g_{*S}\, n_\gamma$, where n_γ is the number density of photons. Today $s = 7.04\, n_\gamma$. Since g_{*S} is a function of temperature, s and n_γ cannot always be used interchangeably.

Conservation of S implies that $s \propto R^{-3}$, and therefore that $g_{*S}\, T^3 R^3$ remains constant as the Universe expands. The first fact, that $s \propto R^{-3}$, implies that the physical size of a comoving volume element $\propto R^3 \propto s^{-1}$. Thus the number of some species in a comoving volume, $N \equiv R^3\, n$, is equal to the number density of that species divided by s (see Fig. 3):

$$N \equiv \frac{n}{s} \tag{4.15}$$

If the number of a given species in a comoving volume is not changing, i.e., particles of that species are not being created or destroyed, then N remains constant.

As an example of the utility of the ratio n/s, consider baryon number. Define n_b to be the number density of baryons in the universe. Similarly define $n_{\bar{b}}$ to be the number density of antibaryons, and the difference between the two, the baryon number in a comoving volume, to be

$$B = \frac{n_B}{s} = \frac{n_b - n_{\bar{b}}}{s} \tag{4.16}$$

So long as baryon number non-conserving interactions (if such exist in nature) are occurring very slowly, the baryon number in a comoving volume, n_B/s, is conserved. Although $\eta = n_B/n_\gamma =$

[10] Here we have assumed that all chemical potentials are zero. It is a very good approximation, as all evidence indicates that $|\mu| \ll T$

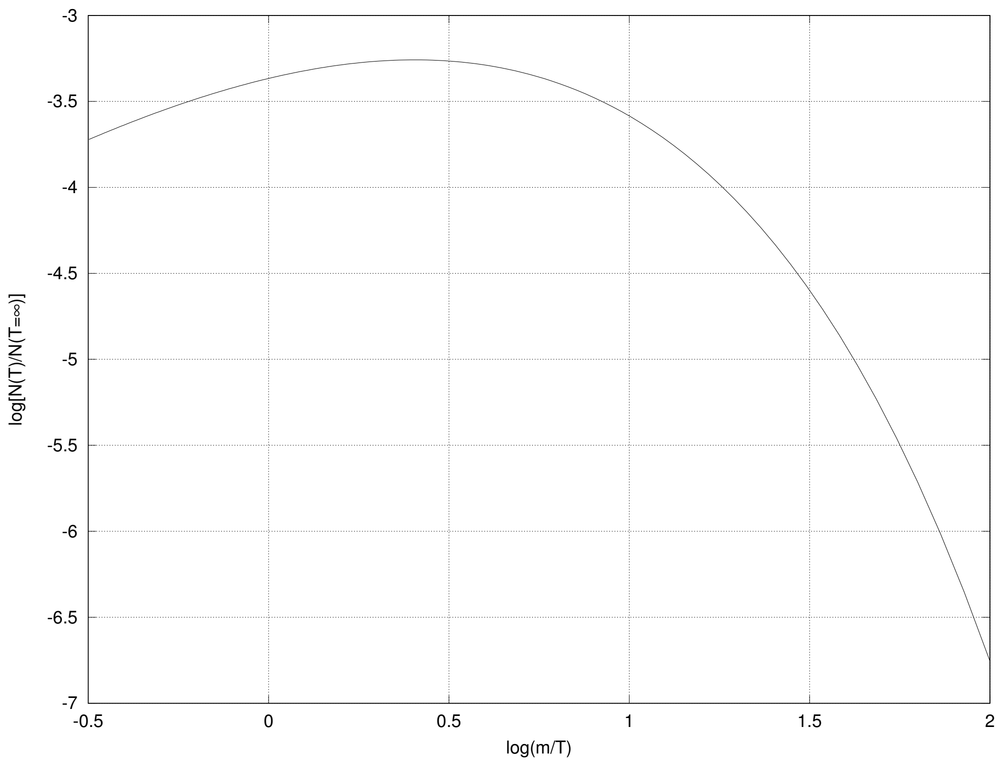

Figure 3: The equilibrium abundance of a species in a comoving volume element, $N = n/s$

$1.8\, g_{*S}(n_B/s)$, the baryon number-to-photon ratio does not remain constant with time because g_{*S} changes. During the era of $e^{\pm}$ annihilation, the number of photons per comoving volume, $N_\gamma = R^3\, n_\gamma$ increases by a factor of 11/4, so that η decreases by the same factor. After the time of $e^{\pm}$ annihilations, however, g_*, is constant, and $\eta \simeq 7\, n_B/s$ and n_B/s can be used interchangeably.

4.2 The Sakharov conditions

We will now review in more detail the Sakharov's conditions [25], necessary to generate a non-zero baryon number from an initially baryon symmetric state:

- Baryon Number Violation. This is rather obvious: there must be a violation of baryon number. If baryon number is conserved in all interactions, the present baryon asymmetry can only reflect asymmetric initial conditions.
- C and CP Violation: Even in the presence of B-non-conserving interactions a baryon asymmetry will not develop unless both C (charge conjugation) and CP (charge conjugation combined with parity) are violated: In the absence of a preference for matter or antimatter, B-non-conserving reactions will produce baryon and antibaryon excesses at the same rate, thereby maintaining zero net baryon number. Both C and CP violation are necessary to supply such an arrow. Put concisely, baryon number is odd under both C and CP. For further details about CP violation see Appendix B.
- Non-Equilibrium Conditions: In chemical equilibrium the entropy is maximal when the chemical potentials associated with all non-conserved quantum numbers vanish. Further, particle and antiparticle masses are guaranteed to be equal by CPT invariance. Thus, in thermal equilibrium the phase space density of baryons and antibaryons, given by the Bose-Einstein distribution

$$f(\vec{\mathbf{p}})_{BE} = \frac{1}{1 + \exp\left(\frac{p^2+m^2}{T}\right)} \tag{4.17}$$

are necessarily identical, implying that $n_b = n_{\bar{b}}$.

particle		final state	branching ratio	B
X	$\rightarrow$	qq	r	2/3
X	$\rightarrow$	$\bar{q}\bar{l}$	$1-r$	$-1/3$
$\bar{X}$	$\rightarrow$	$\bar{q}\bar{q}$	$\bar{r}$	$-2/3$
$\bar{X}$	$\rightarrow$	$q\bar{l}$	$1-\bar{r}$	1/3

Table 1: Final states and branching ratios for X, $\bar{X}$ decay.

To illustrate the mechanics of Baryogenesis, consider a particle X that decays to quark/lepton final states qq (B = 2/3) and $\bar{q}\bar{l}$ (B = −1/3) (see Table 1). Since the two final states have different baryon number, the decays of X, $\bar{X}$ violate B. Note that CPT invariance requires the equality of the decay rates of the X and $\bar{X}$ bosons; C and CP are violated if the branching ratio of the X to the qq final state ($= r$) is unequal to the branching ratio of the $\bar{X}$ to the $\bar{q}\bar{q}$ final state ($= \bar{r}$); that is, $r \neq \bar{r}$ [17].

Imagine a system with symmetric initial conditions: equal numbers of X and $\bar{X}$ bosons. The mean net baryon number produced by the decay of an X is equal to

$$B_X = r\left(\frac{2}{3}\right) + (1-r)\left(-\frac{1}{3}\right) \tag{4.18}$$

and that produced by the decay of an $\bar{X}$ is equal to

$$B_{\bar{X}} = \bar{r}\left(-\frac{2}{3}\right) + (1-\bar{r})\left(\frac{1}{3}\right) \tag{4.19}$$

The mean net baryon number produced by the decay of an X, $\bar{X}$ pair is just

$$\epsilon = B_X + B_{\bar{X}} = r - \bar{r} \tag{4.20}$$

The baryon number produced vanishes, of course, if C or CP is conserved $r = \bar{r}$. If there are no further baryon number violating reactions, then a net baryon asymmetry will persist after all the X, $\bar{X}$ bosons decay.

Let's see the Sakharov conditions in more detail.

4.2.1 B violation

The existence of baryon number violation seems to be a generic feature of a Grand Unified Theory[11] (GUT). When the strong and electroweak interactions are unified, quarks and leptons typically appear as members of a common irreducible representation of the gauge group. Thus gauge bosons mediate interactions that transform quarks into leptons or antiquarks, and thereby violate B. The lifetime of the proton should be $\tau_p \gtrsim 10^{31}$ to 10^{32} years, which implies that such additional gauge bosons must be very massive: $M \gtrsim 10^{14}\,GeV$ or so. In addition, there are also likely to be Higgs bosons with B-non-conserving interactions. The typically weaker couplings of Higgs bosons that mediate baryon-number violation allow them to have somewhat smaller masses, perhaps as low as $10^{10}\,GeV$. In both cases the large mass of the intermediate boson is responsible for the feebleness of baryon-number violation today. This suppression (relative to the familiar interactions) is of course overcome at the extremely high temperatures that should have existed shortly after the big bang, and interactions that violate baryon number should have been just as potent as all other interactions. The requirement of B violation arises naturally in GUTs [17].

Almost a decade before the advent of GUTs, Sakharov in [25] proposed a concrete model, of an interaction that violates baryon charge in the super dense state of the initial Universe.

Both C and CP are observed to be violated microscopically in Nature, in the interactions of K^0 and $\bar{K}^0$ mesons.

[11] A Grand Unified Theory is a model in particle physics in which at high energy, the three gauge interactions of the Standard Model which define the electromagnetic, weak, and strong interactions or forces, are merged into one single force. This unified interaction is characterised by one larger gauge symmetry and thus several force carriers, but one unified coupling constant. If Grand Unification is realised in nature, there is the possibility of a grand unification epoch in the early universe in which the fundamental forces are not yet distinct.

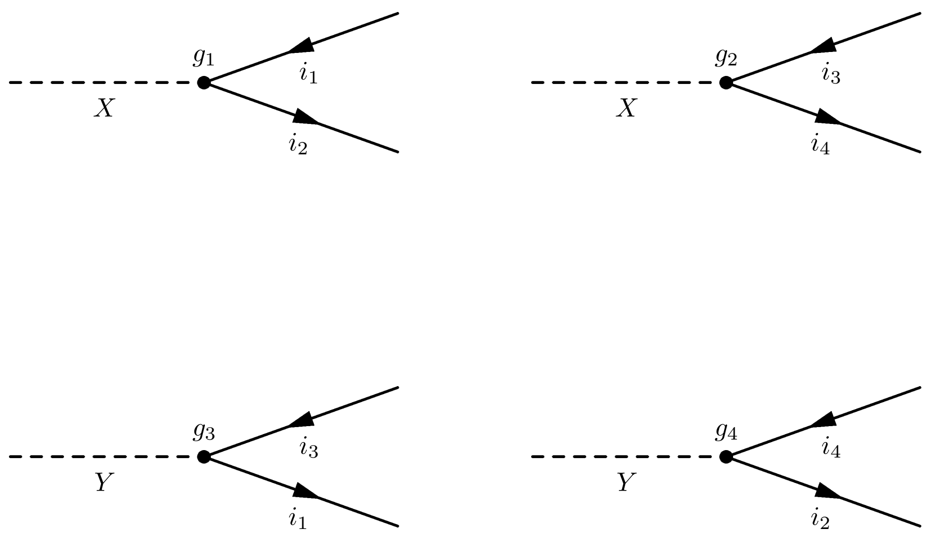

Figure 4: Lowest-order Feynman diagrams for X and Y decay.

4.2.2 C and CP violation

C is maximally violated in the weak interactions, so C violation in the decay of the X boson should not be a fundamental problem. CP violation is observed in the neutral kaon system, with dimensionless strength of 10^{-3}. Since its origin is not well understood, it is easy to imagine that C and CP violation manifest themselves in all sectors of the theory, including the super-heavy boson sector and at some level C and CP violation must occur in the super-heavy sector due to loop corrections involving the light quarks. It can be shown that a C, CP violation of only $\epsilon \sim 10^{-8}$ or so is required to produce the observed value of n_b/s. To explicitly see how C, CP violation enters, consider a system with two super-heavy bosons, X and V, with baryon number violating decays. The generalisation of ϵ defined above is

$$\begin{aligned} \epsilon_X &= \sum_f B_f \frac{\Gamma(X \to f) - \Gamma\big(\bar{X} \to \bar{f}\big)}{\Gamma_X} \\ \epsilon_Y &= \sum_f B_f \frac{\Gamma(Y \to f) - \Gamma\big(\bar{Y} \to \bar{f}\big)}{\Gamma_Y} \end{aligned} \tag{4.21}$$

where the sum runs over all final states f, state f has baryon number B_f, and Γ_X (Γ_Y) is the total X (Y) decay width. For simplicity, assume there are but two final states for X and Y decay, and that the interaction Lagrangian is given by

$$\mathcal{L} = g_1 X i_2^\dagger i_1 + g_2 X i_4^\dagger i_3 + g_3 X i_1^\dagger i_3 + g_4 X i_2^\dagger i_4 \quad + \quad \text{h.c.} \tag{4.22}$$

where i_1, i_2, i_3, i_4 are fermion states (quarks and leptons), and the g_i are coupling strengths (which can be complex). This Lagrangian leads to the decay processes: $X \to \bar{i}_1 + i_2$, $\bar{i}_3 + i_4$, and $Y \to i_1 + \bar{i}_3$, $i_3 + \bar{i}_4$. The lowest order diagrams for X and Y decay are shown in Fig. 4

The lowest order processes cannot contribute to ϵ, as $\Gamma(X \to \bar{i}_1 i_2) = |g_1|^2 I_X = \Gamma\big(\bar{X} \to i_1 \bar{i}_2\big) = |g_1^*|^2 I_{\bar{X}}$ where the kinematic factors $I_X = I_{\bar{X}}$ arise from the phase space integrals. The first non-zero contribution to ϵ comes from the interference of the lowest order graphs in Fig. 4 with the one-loop corrections shown in Fig. 5.

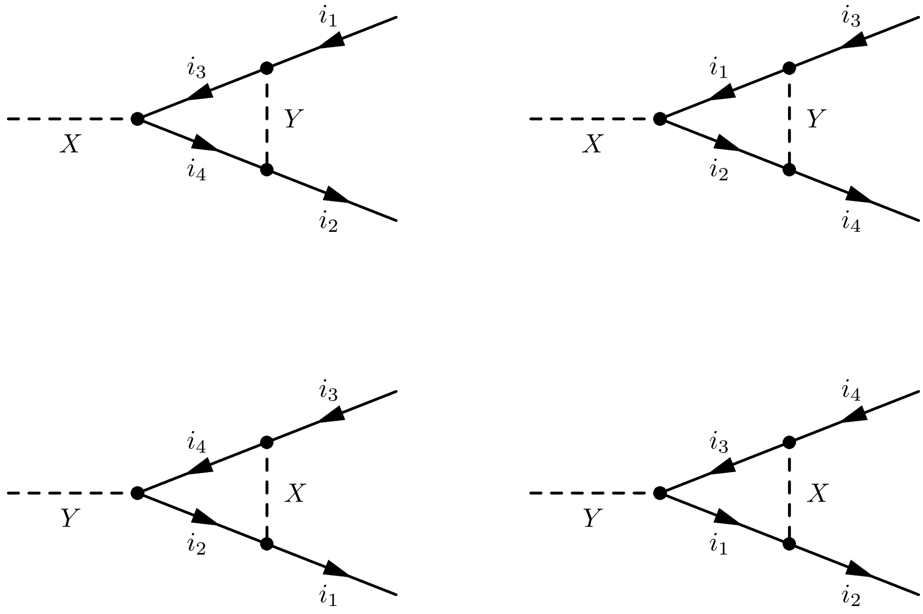

Figure 5: One-loop corrections to X and Y decay.

These interference terms are given by

$$\begin{aligned}\Gamma(X \to \bar{i}_1 i_2) &= g_1 g_2^* g_3 g_4^* I_{XY} + (g_1 g_2^* g_3 g_4^* I_{XY})^* \\ \Gamma(\bar{X} \to i_1 \bar{i}_2) &= g_1^* g_2 g_3^* g_4 I_{XY} + (g_1^* g_2 g_3^* g_4 I_{XY})^* \end{aligned} \tag{4.23}$$

where the phase-space factors I_{IJ} (I, $J = X, Y$) now also include the kinematic factors arising from integrating over the internal momentum loop due to J exchange in I decay. If the intermediate particles (i_1, i_2, i_3, i_4) in the loop are kinematically allowed to propagate on shell, as will be the case if X, Y are super-heavy bosons and i_{1-4} are light quarks and leptons, the quantity I_{IJ} will be complex. The complexity of I_{IJ} is crucial to Baryogenesis. The difference between $X \to \bar{i}_1 i_2$ and $\bar{X} \to i_1 \bar{i}_2$ is given by

$$\begin{aligned}\Gamma(X \to \bar{i}_1 i_2) - \Gamma(\bar{X} \to i_1 \bar{i}_2) &= 2iI_{XY} \operatorname{Im}\{g_1 g_2^* g_3 g_4^*\} + 2iI_{XY}^* \operatorname{Im}\{g_1^* g_2 g_3^* g_4\} \\ &= 4 \operatorname{Im}\{I_{XY}\} \operatorname{Im}\{g_1^* g_2 g_3^* g_4\} \end{aligned} \tag{4.24}$$

After a similar calculation for the other decay mode, we find that

$$\epsilon_X = \frac{4}{\Gamma_X} \operatorname{Im}\{I_{XY}\} \operatorname{Im}\{g_1^* g_2 g_3^* g_4\}[(B_{i_4} - B_{i_3}) - (B_{i_2} - B_{i_1})] \tag{4.25}$$

Repeating the same calculations for the Y decay modes, we find $\epsilon_Y = -\epsilon_X$. So, we see that 3 things are required in order to have $\epsilon \neq 0$.

- There must be *two* baryon number violating bosons, each with mass greater than the sum of the (fermion) masses in the internal loops; otherwise, $\operatorname{Im}\{I_{XY}\} = 0$.
- C, CP violation arises from the interference of loop diagrams with the tree graph, and manifests itself in complex coupling constants. Thus, in general we expect ϵ to be of the order of α^N, where α characterises the coupling constant of the loop particle(s) and N is the number of loops in the lowest-order diagram that interferes with the tree graph to give rise to $\epsilon \neq 0$. Note that for this reason, $\epsilon \sim \alpha^N$ is expected to be small.

- The X and Y particles in the above example must not be degenerate in mass, or the baryon number produced by X decays will precisely cancel that produced by Y decays.

4.2.3 Departure from Thermal Equilibrium

The interactions must be out of thermal equilibrium, since otherwise CPT symmetry would assure compensation between processes increasing and decreasing the baryon number. In fact [32]

$$\begin{aligned} \langle B \rangle_T &= \mathrm{Tr}\left(e^{-\beta H} B\right) \\ &= \mathrm{Tr}\left[(CPT)(CPT)^{-1} e^{-\beta H} B\right] \\ &= \mathrm{Tr}\left[e^{-\beta H}(CPT)B(CPT)^{-1}\right] \\ &= -\,\mathrm{Tr}\left(e^{-\beta H} B\right) \end{aligned} \tag{4.26}$$

where in the third step I have used the requirement that the Hamiltonian H commutes with CPT, and in the last step used the properties of B that it is odd under C and even under P and T symmetries. Thus $\langle B \rangle_T = 0$ in equilibrium and there is no generation of net baryon number.

The necessary non-equilibrium condition is provided by the expansion of the Universe. Recall that if the expansion rate is faster than key particle interaction rates, departures from equilibrium can result. Here the departure from equilibrium will be the overabundance of X, $\bar{X}$ bosons. Assume that at some very early time when $T \gg m_X$ (e.g., the Planck time), X, $\bar{X}$ bosons are present in equilibrium numbers, $n_X = n_{\bar{X}} \simeq n_\gamma$. In LTE

$$\begin{aligned} n_X = \;& n_{\bar{X}} & \simeq n_\gamma \quad \text{for} \quad T \gtrsim m_X \\ n_X = \;& n_{\bar{X}} \simeq (m_X T)^{\frac{3}{2}} \exp\left(-\tfrac{m_X}{T}\right) & \ll n_\gamma \quad \text{for} \quad T \lesssim m_X \end{aligned} \tag{4.27}$$

Equilibrium numbers of X, $\bar{X}$ bosons will only be present provided that the interactions which create and destroy X, $\bar{X}$ bosons (decay, annihilation, and their inverse processes) are occurring rapidly on the expansion time scale: $\Gamma \gtrsim H$. The annihilation process is "self-quenching" since $\Gamma_{ANN} \propto n_X$, and the decay process is most important for maintaining equilibrium numbers of X, $\bar{X}$ bosons. For simplicity then, we will ignore the annihilation process.

4.3 The Boltzmann equation

For much of its early history, most of the constituents of the Universe were in thermal equilibrium, thereby making an equilibrium description a good approximation. However, there have been a number of very notable departures from thermal equilibrium: neutrino decoupling, decoupling of the background radiation, primordial nucleosynthesis, and on the more speculative side, inflation, Baryogenesis, decoupling of relic WIMPs (weakly-interacting massive particles), etc. As we have previously emphasised, if not for such departures from thermal equilibrium, the present state of the Universe would be completely specified by the present temperature. The departures from equilibrium have led to important relicsthe light elements, the neutrino backgrounds, a net baryon number, relic WIMPs , relic cosmologists, and so on.

Once a species totally decouples from the plasma its evolution is very simple: particle number density decreasing as R^{-3} and particle momenta decreasing as R^{-1} [17]. The evolution of the phase space distribution of a species which is in LTE or is completely decoupled is simple. It is the evolution of particle distributions around the epoch of decoupling that is challenging. Recall that the rough criterion for a particle species to be either coupled or decoupled involves the comparison of the interaction rate of the particle, T, with the expansion rate of the Universe, H, as seen in Eq. (4.2). In order to properly treat decoupling one must follow the microscopic evolution of the particle's phase space distribution function $f(p^\mu, x^\mu)$. This of course is governed by the Boltzmann equation, which can be written as

$$\hat{\mathbf{L}}[f] = \mathbf{C}[f] \tag{4.28}$$

where $\mathbf{C}$ is the collision operator and $\hat{\mathbf{L}}$ is the Liouville operator. The non-relativistic Liouville operator for the phase space density $f(\vec{\mathbf{v}}, \vec{\mathbf{x}})$ of a particle species of mass m subject to a force $\vec{\mathbf{F}} = \mathrm{d}\vec{\mathbf{p}}/\mathrm{d}t$ is

$$\hat{\mathbf{L}}_{NR} = \frac{\mathrm{d}}{\mathrm{d}t} + \frac{\mathrm{d}\vec{\mathbf{x}}}{\mathrm{d}t}\cdot\boldsymbol{\nabla}_x + \frac{\mathrm{d}\vec{\mathbf{v}}}{\mathrm{d}t}\cdot\boldsymbol{\nabla}_v = \frac{\mathrm{d}}{\mathrm{d}t} + \vec{\mathbf{v}}\cdot\boldsymbol{\nabla}_x + \frac{\vec{\mathbf{F}}}{m}\cdot\boldsymbol{\nabla}_v \tag{4.29}$$

The covariant, relativistic generalisation of the Liouville operator is

$$\hat{\mathbf{L}} = p^\alpha \frac{\partial}{\partial x^\alpha} - \Gamma^\alpha_{\beta\gamma} p^\beta p^\gamma \frac{\partial}{\partial p^\alpha} \tag{4.30}$$

For the Friedmann-Robertson-Walker (FRW) metric[12] the Liouville operator is

$$\hat{\mathbf{L}}[f(E,t)] = E\frac{\partial f}{\partial t} - \frac{\dot{R}}{R}|\vec{\mathbf{p}}|^2\frac{\partial f}{\partial E} \tag{4.31}$$

Using the definition of the number density in terms of the phase space density

$$n(t) = \frac{g}{(2\pi)^3}\int \mathrm{d}^3p\, f(E,t) \tag{4.32}$$

and upon integration by parts, the Boltzmann equation can be written in the form

$$\frac{\mathrm{d}n}{\mathrm{d}t} + 3\frac{\dot{R}}{R}n = \frac{g}{(2\pi)^3}\int \mathbf{C}[f]\frac{\mathrm{d}^3p}{E} \tag{4.33}$$

The collision term for the process $\psi + a + b + \ldots \leftrightarrow i + j + \ldots$ is given by

$$\begin{aligned}\frac{g}{(2\pi)^3}\int \mathbf{C}[f]\frac{\mathrm{d}^3p_\psi}{E_\psi} = &-\int \mathrm{d}\Pi_\psi\,\mathrm{d}\Pi_a\,\mathrm{d}\Pi_b\ldots\mathrm{d}\Pi_i\,\mathrm{d}\Pi_j\ldots \\ &\times (2\pi)^4\delta^4(p_\psi + p_a + p_b\ldots - p_i - p_j\ldots) \\ &\times \Big[|\mathcal{M}|^2_{\psi+a+b+\ldots\to i+j+\ldots} f_a f_b\ldots f_\psi(1\pm f_i)(1\pm f_j)\ldots \\ &\quad - |\mathcal{M}|^2_{i+j+\ldots\to\psi+a+b+\ldots} f_i f_j\ldots(1\pm f_a)(1\pm f_b)\ldots(1\pm f_\psi)\Big]\end{aligned} \tag{4.34}$$

where $f_i, f_j, f_a, f_b, \ldots$ are the phase space densities of species $i, j, \ldots, a, b, \ldots$; f_ψ is the phase space density of ψ (the species whose evolution we are focusing on); $(+)$ applies to bosons; $(-)$ applies to fermions; and

$$\mathrm{d}\Pi \equiv \frac{g}{(2\pi)^3}\frac{\mathrm{d}^3p}{2E} \tag{4.35}$$

where g counts the internal degrees of freedom. The 4-dimensional delta function enforces energy and momentum conservation, and the scattering matrix element squared $|\mathcal{M}|^2$, which includes the appropriate symmetry factors for identical particles in the initial or final states, relates the initial state and the final state of the process $\psi + a + b + \ldots \to i + j + \ldots$.

The significance of the individual terms is manifest: the $3\dot{R}/Rn = 3Hn$ term accounts for the dilution effect of the expansion of the Universe, and the right hand side of Eq. (4.33) accounts for interactions that change the number of ψ's present. In the absence of interactions the solution to Eq. (4.33) is $n_\psi \propto R^{-3}$.

In the most general case, the Boltzmann equations are a coupled set of integral-partial differential equations for the phase space distributions of all the species present. Fortunately, in problems of interest to us, all but one (or two) species will have equilibrium phase space distribution functions because of their rapid interactions with other species, reducing the problem to a single integral-partial differential equation for the one species of interest, denoted by ψ. We can do some approximations that greatly simplify Eq. (4.34). We can use the Maxwell-Boltzmann statistics for all species instead of Fermi-Dirac for fermions and Bose-Einstein for bosons[13]. In the

[12]The FRW metric is an exact solution of Einstein's field equations of general relativity; it describes a homogeneous, isotropic expanding or contracting universe. It is the metric we use to model our Universe

[13]In the absence of a degenerate ($\mu_i \gtrsim T$) Fermi species or a Bose condensate,the use of Maxwell-Boltzmann statistics introduces only a small quantitative change, as all three distribution functions are very similar (and much less than one) for momenta near the peak of the distribution. Moreover, for any non-relativistic species, Maxwell-Boltzmann statistics becomes exact in the limit $(m_i - \mu_i)/T \gg 1$

absence of Bose condensation or Fermi degeneracy, the blocking and stimulated emission factors can be ignored, $(1 \pm f) \simeq 1$, and $f_i(E_i) = \exp[-(E_i - \mu_i)/T]$ for all species in kinetic equilibrium.

We can use the Boltzmann equation to study the precise evolution of the baryon asymmetry. In [17] a Boltzmann equation is derived for a simplified model, which serves to illustrate most of the salient features of Baryogenesis, and which can be used to mimic realistic GUTs. The simple model consists of a massive boson that is self conjugate ($X = \bar{X}$) and has interactions that violate B, C, and CP,"light" (i.e., highly relativistic) particles b and $\bar{b}$ that carry baryon numbers $+1/2$ and $-1/2$, and light particles that carry no baryon number and represent the thermal bath of radiation, and collectively have g_* degrees of freedom.

4.4 The Sphaleron

As we have seen in section 2, a result of non-trivial vacuum gauge configurations, the vacuum structure of non-Abelian gauge theories is very rich. The θ-vacuum structure in the electroweak theory leads to the anomalous non-conservation of baryon number. Transitions between the different vacua are accompanied by a change in the baryon number. And baryon-number non-conservation (together with the lack of CP non-conservation and thermodynamic equilibrium) can give rise to the baryon asymmetry of the universe (BAU). This process is usually associated with the instanton solutions that describe the tunnelling between different θ-vacua (see section A.5) [17]. In the weakly coupled theories, the probabilities of these transitions are exponentially suppressed; in particular, the corresponding suppression factor in the standard electroweak theory is $\exp(-4\pi/\alpha_W)$, where $\alpha_W = g_W^2/4\pi$ and g_W is the electroweak coupling constant. However, if the energy of the system is large enough, the system can, in principle, pass over the barrier between the different vacua instead of penetrating through the barrier (see Fig. 6)

In [18] a model is proposed to estimate the equilibrium rate of the anomalous fermion-number non-conserving processes in the cosmic plasma in the standard Weinberg-Salam electroweak theory, and the possibility of the BAU generation at the electroweak temperatures in the standard theory with light fermions. The authors find that at sufficiently high temperature this rate exceeds the expansion rate of the universe, but that BAU is not generated by anomalous electroweak processes if the EWPT is of second order (see section 4.5). The Lagrangian of the theory is

$$\mathcal{L} = \frac{1}{2g_W^2}\,\mathrm{tr}\,F_{\mu\nu}^2 + (D_\mu\phi)^\dagger(D^\mu\phi) - \lambda\left(\phi^\dagger\phi - \frac{v^2}{2}\right)^2 + \mathcal{L}_F \tag{4.36}$$

where $\mathcal{L}_F$ is the fermionic part. The gauge $A_0 = 0$ is used. To estimate the rate of the fermion-number non-conservation (at thermodynamic equilibrium) we have to study the number of level crossings per unit time per unit volume. In the picture proposed in section 2 we have seen that the vacua with, say, $n = 1$ and $n = 0$ differ by a topologically non-trivial gauge transformation, and the level crossing occurs at the top of the barrier, as seen if Fig. 6. Therefore, the rate of level crossings coincides with the rate of the transitions from the regions left to the barrier maximum to that right to it, into the adjacent θ-vacuum. This can be evaluated with the use of the theory of the"vacuum"decay at finite temperature which states that this rate is proportional to $\exp(-S)$ where S is the action for the appropriate periodic (up to twist) solution to the Euclidean field equations. At sufficiently high temperatures (including those under discussion) two approximations can be made: the first is to ignore the effects of fermions on the fluctuations of the bosonic fields and to consider A_μ^a and ϕ^a as the only dynamical fields. The second is to replace the finite temperature action S by the free energy at the maximum of the barrier between adjacent θ-vacua. In that case the transition rate is proportional to $\exp[-F(A_{cl}, \phi_{cl})/T]$.

The sphaleron solution is a saddle point in field configuration space: the lowest barrier between two θ-vacua. Moreover, the sphaleron configuration is classically unstable[14]. At temperatures under discussion, the free energy functional is approximately equal to the static energy

$$F = \int \mathrm{d}^3x\,\frac{1}{2g_W^2}\,\mathrm{tr}\,F_{\mu\nu}^2 + (D_\mu\phi)^\dagger(D^\mu\phi) - \lambda\left(\phi^\dagger\phi - \frac{v^2}{2}\right)^2 \tag{4.37}$$

The configuration that extremists (4.37) for this transition is indeed the sphaleron, given in the $A_0 = 0$ gauge by

[14] As its name implies: *sphaleron*, Greek for "ready to fall"

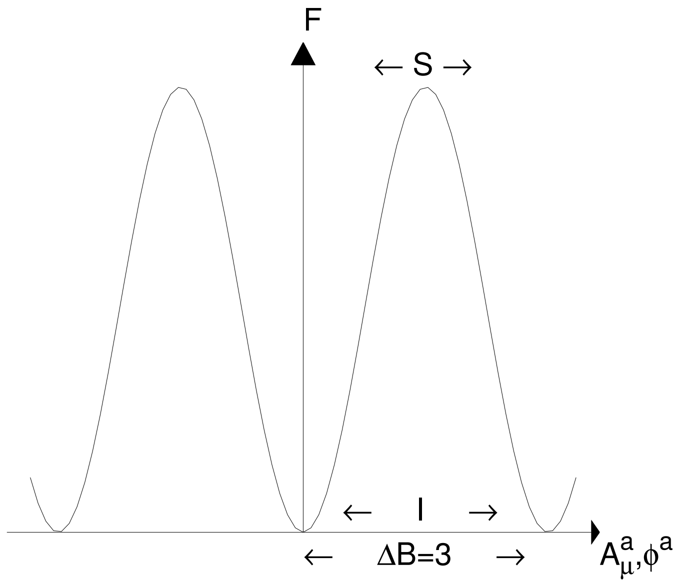

Figure 6: Schematic illustration of the barriers separating different θ-vacua, the instanton (I) tunnelling path through the barrier, and the sphaleron (S) path over the barrier. The vertical axis is the free energy (F), and the horizontal axis represents gauge (A^a_μ) and Higgs (ϕ^a) field configurations.

$$\begin{aligned} A^i_{cl} &= \frac{i}{g_W}\frac{\epsilon_{ijk}x_j\tau_k}{r^2}f(\xi) \\ \phi^a_{cl} &= \frac{v}{\sqrt{2}}\frac{i\vec{\tau}\cdot\vec{\mathbf{x}}}{r}\begin{bmatrix}0\\1\end{bmatrix}h(\xi) \end{aligned} \tag{4.38}$$

where $\vec{\tau}$ are Pauli matrices, $r^2 = |\vec{\mathbf{x}}|^2$ and $\xi = g_W(T)v(T)r$. The functions $f(\xi)$ and $h(\xi)$ have the following asymptotics

$$f(0) = h(0) = 0, \quad f(\infty) = h(\infty) = 1 \tag{4.39}$$

With the assumption that the action for the tunnelling rate at finite temperature is given by F, the rate of vacuum transitions should be proportional to $\exp(F/T)$. Heuristically, this corresponds to the probability of having a sufficiently large thermal fluctuation to take the fields over a barrier of height F. The rate at which the baryon number per comoving volume is violated should be proportional to this rate times the net baryon number per comoving volume, i.e.,

$$\dot{B} = CBT\exp\left(-\frac{F}{T}\right) \tag{4.40}$$

where C is a dimensionless constant expected to be of order unity, and the overall factor of T is assumed on dimensional grounds. Since CP is conserved by the gauge and Higgs interactions in this model, a baryon asymmetry cannot be generated by the vacuum transitions. In the early Universe g_W, λ, and v (and hence m_W) are all functions of temperature. The temperature dependence of g_W and v is only proportional to $\ln T$ and can be ignored. The temperature dependence of around the critical temperature is much more important. At $T > T_c$, $v = 0$, and as the temperature passes below the critical temperature, $v^2 \simeq v_0^2\big(1 - T^2/T_c^2\big)$ ($v_0 = 246\,GeV$).

At lower temperatures, $T \ll T_C$, $F/T \gg 1$ and the rate for sphaleron-induced B violation is exponentially suppressed. However as $T \to T_C$, $F/T \propto v(T)/T \to 0$, and the rate for B violation is no longer exponentially suppressed. For a range of temperatures just below T_C, $\Gamma_{\Delta B}/H = \left(\dot{B}/B\right)/H \sim 10^{17}C$, corresponding to very rapid B violation. The sphaleron solution has also been interpreted as a massive, unstable particle whose interactions violate B.

4.5 Phase Transitions

One of the most important concepts in modern particle theory is that of spontaneous symmetry breaking[15]. The idea that there are underlying symmetries of Nature that are not manifest in the structure of the vacuum appears to play a crucial role in the unification of the forces. In all unified gauge theories (including the Standard Model of particle physics, see Appendix C) the underlying gauge symmetry is larger than that of our vacuum, whose symmetry is that of $SU_c(3) \times SU_L(2) \times U_Y(1)$ (see Appendix C). Of particular interest for cosmology is the theoretical expectation that at high temperatures, symmetries that are spontaneously broken today were restored, and that during the evolution of the Universe there were phase transitions, perhaps many, associated with the spontaneous breakdown of gauge (and perhaps global) symmetries. In particular, we can be reasonably confident that there was such a phase transition at a temperature of order 300 GeV and a time of order 10^{-11} Sec, associated with the breakdown of $SU_L(2) \times U_Y(1) \to U_{EM}(1)$. Moreover, the vacuum structure in many spontaneously broken gauge theories is very rich: Topologically stable configurations of gauge and Higgs fields exist as domain walls, cosmic strings, and monopoles[16].

A simple model to illustrate high-temperature symmetry restoration and phase transitions is proposed in [17]. Consider a real scalar field described by the Lagrangian density

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi) \\ V(\phi) &= -\frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\phi^4 \end{aligned} \tag{4.41}$$

The potential $V(\phi)$, as given in Eq. (4.45), is shown in Fig. 7. The key point is that the Lagrangian is invariant under the discrete symmetry transformation $\phi \leftrightarrow -\phi$. The minima of the potential, denoted by $\sigma_\pm$, and the value of the potential and its second derivative at the minima, are given by $\sigma_\pm = \pm\sqrt{m^2/\lambda}$, $V(\sigma_\pm) = -m^4/4\lambda$ and $V''(\sigma_\pm) = 2m^2$ respectively.

Since $V(+\sigma) = V(-\sigma)$, both $\sigma_\pm = \pm\sigma$ are equivalent minima of the potential. On the other hand, although $V'(0) = 0$, $\langle\phi\rangle = 0$ is an unstable extremum of the potential because $V''(0) < 0$. Since the quantum theory must be constructed about a stable extremum of the classical potential, the ground state of the system is either $\langle\phi\rangle = \sigma$ or $\langle\phi\rangle = -\sigma$, and the reflection symmetry present in the Lagrangian is broken by the choice of a vacuum state. A symmetry of the Lagrangian not respected by the vacuum is said to be *spontaneously broken*. The mass of the physical boson of the theory is determined by the curvature of the potential about the true ground state

$$M^2 = V''(\sigma_\pm) = 2m^2 = 2\lambda\sigma_\pm^2 \tag{4.42}$$

The stress tensor for a scalar field ϕ is given by [24]

$$T_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - \mathcal{L}g_{\mu\nu} \tag{4.43}$$

Taking ϕ to be constant, $\phi = \langle\phi\rangle$, we find that $T^{\mu\nu} = V(\langle\phi\rangle)g^{\mu\nu}$, so that the energy density of the vacuum is

$$\langle T_0^0\rangle \equiv \rho_V = -\frac{m^4}{4\lambda} \tag{4.44}$$

The contribution of the vacuum energy to the energy density of the Universe today can at most be comparable to the critical density $\rho_C \simeq 10^{-46}\,GeV$. A larger vacuum energy would lead to a present expansion rate greater than that observed. Therefore, we can set $\rho_V = 0$. This can be accomplished by adding to the Lagrangian the constant $+m^4/4\lambda = \lambda\langle\phi\rangle^4/4$. This constant term

[15] For an overview of Spontaneous Symmetry Breaking see [27]

[16] In addition, classical configurations that are not topologically stable, so-called non-topological solitons, may exist and be stable for dynamical reasons. Interesting examples include soliton stars, Q-balls, non topological cosmic strings, and so on.

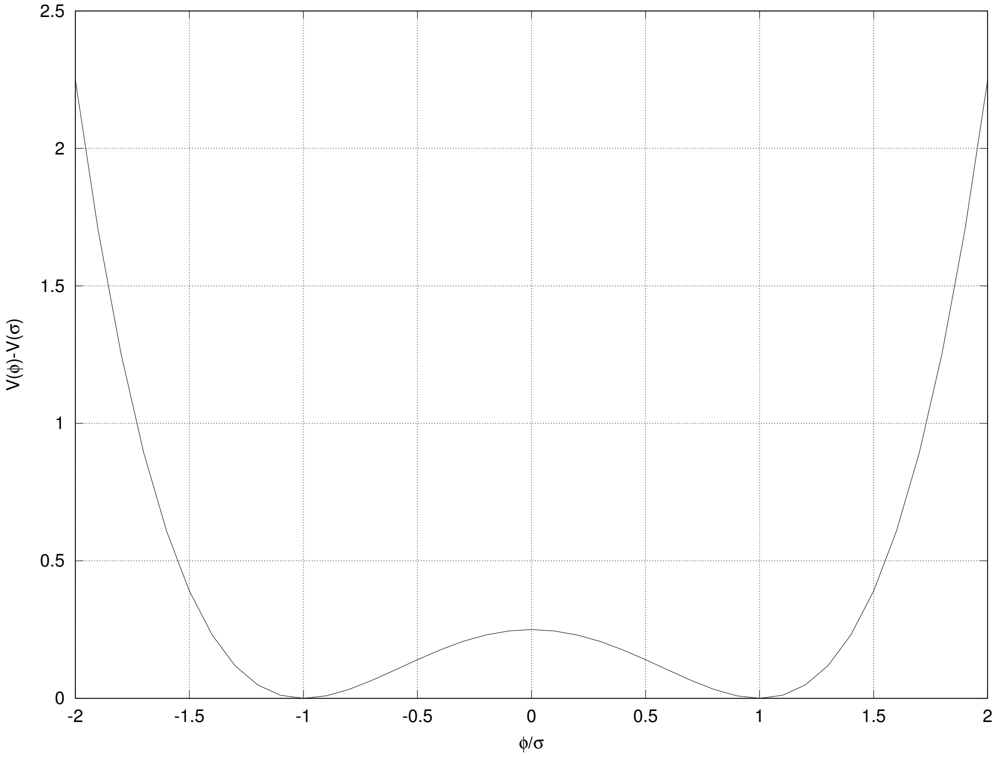

Figure 7: An example of a potential that implements SSB.

will not affect the equations of motion or the quantum theory; its sole effect is to make the present vacuum energy vanish. By adding this constant, we can write the potential in the form

$$V(\phi) = -\frac{\lambda}{4}\left(\phi^2 - \sigma^2\right)^2, \quad \sigma^2 = \sigma_\pm^2 \tag{4.45}$$

We can understand the phenomenon of high-temperature symmetry restoration considering the effect of finite temperature (or density) upon the propagation of a particle. If one naively attempts to construct a theory about $\langle\phi\rangle = 0$ using the potential of Eq. (4.45), one finds that the mass squared of the scalar field is negative: $V''(0) = -\lambda\sigma^2$. The imaginary mass results in exponentially growing solutions, with ϕ growing until it finds the true ground state. However, if the ϕ field is in contact with a heat bath, the interaction of ϕ particles with particles in the thermal bath will, in general, damp this exponential growth. A way to quantify this damping is to assign to the ϕ a temperature-dependent "plasma mass", which on dimensional grounds must be of the form $m^2_{plasma} = a\lambda T^2$, where a is a numerical constant of order unity. At finite temperature the effective mass of the scalar field about the classical solution $\langle\phi\rangle = 0$ is $m_T^2 = -\lambda\sigma^2 + m^2_{plasma}$. At temperatures where $m_T^2 < 0$, $\langle\phi\rangle = 0$ will be an unstable point, signalling SSB; while at temperatures where $m_T^2 > 0$, the effective mass will be real, and $\langle\phi\rangle = 0$ becomes a stable, classical minimum of the potential. Clearly, there is some critical temperature, $T_C \simeq \sigma/\sqrt{a}$, where $m_T^2 = 0$; above this critical temperature $\langle\phi\rangle = 0$ is a stable minimum and the symmetry is restored.

The phase transition from the symmetric phase to the broken phase in this model is second order. In general, a symmetry-breaking phase transition can be first or second order. The temperature dependence of $V(\phi)$ for a first-order phase transition is shown in Fig. 8. For $T \gg T_C$ the potential is quadratic, with only one minimum at $\phi = 0$. When $T > T_C$, a local minimum develops at $\phi \neq 0$. For $T = T_C$, the two minima become degenerate, and below T_C, the $\phi \neq 0$ minimum becomes the global minimum. If for $T \leq T_C$ the extremum at $\phi = 0$ remains a local minimum, there must be a barrier between the minima at $\phi = 0$ and $\phi \neq 0$. Therefore, the change in ϕ in going from one phase to the other must be discontinuous, indicating a first-order phase transition. Moreover, the transition cannot take place classically, but must proceed either through quantum or thermal tunnelling. Finally, when $T < T_C$ the barrier disappears and the transition may proceed classically. For a second-order transition there is no barrier at the critical temperature, and the transition occurs smoothly.

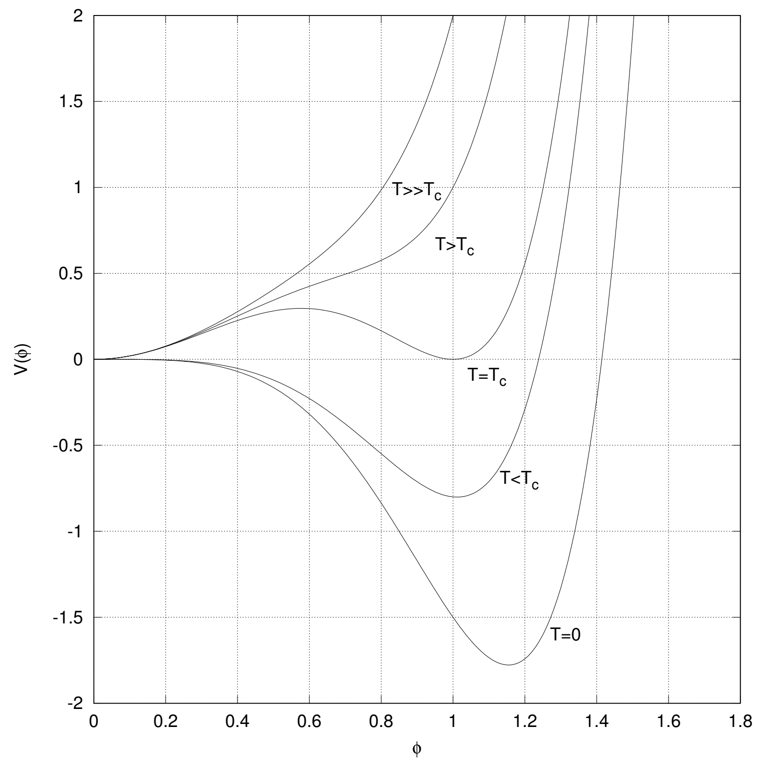

Figure 8: The temperature dependence of $V_T(\phi)$ for a first-order phase transition.

Transition to the true vacuum state by quantum tunnelling occurs through the nucleation of bubbles of the energetically favored asymmetric phase ($\phi = \sigma$), which then expand outward at the speed of light [17]. The first step in the calculation of the probability for bubble nucleation is solving the Euclidean equation of motion for $\phi(t_E, \vec{\mathbf{x}})$, where $t_E = it$

$$\Box_E \phi - V'(\phi) = \frac{\mathrm{d}^2\phi}{\mathrm{d}t_E^2} + \nabla^2\phi - V'(\phi) = 0 \tag{4.46}$$

The probability of bubble nucleation per unit volume per unit time is given by

$$\Gamma = A \exp\left(-S_E\right) \tag{4.47}$$

where S_E is the Euclidean action for the solution of Eq. (4.46)

$$S_E(\phi) = \int \mathrm{d}^3x \, \mathrm{d}t_E \left[\frac{1}{2}\left(\frac{\mathrm{d}\phi}{\mathrm{d}t_E}\right)^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi) \right] \tag{4.48}$$

For most applications an estimate for A based on dimensional grounds will suffice. All possible solutions to Eq. (4.46) contribute to the tunnelling rate; and the one with least action makes the largest contribution to the tunnelling rate. In flat space, at zero temperature, the least-action Euclidean solution has $O(4)$ symmetry, in which case ϕ is only a function of r, where $r^2 = t_E^2 + |\vec{\mathbf{x}}|^2$ and the $O(4)$-Euclidean equation of motion for $\phi(r)$ is

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}r^2} + \frac{3}{r}\frac{\mathrm{d}\phi}{\mathrm{d}r} - V'(\phi) = 0 \tag{4.49}$$

Once the solution $\phi(r)$ is known, the Euclidean action is obtained from Eq. (4.48):

$$S_E = 2\pi^2 \int_0^\infty \mathrm{d}r \, r^3 \left[\frac{1}{2}\left(\frac{\mathrm{d}\phi}{\mathrm{d}r}\right)^2 + V(\phi) \right] \tag{4.50}$$

The Euclidean solution $\phi(r)$ can be interpreted if, for simplicity, we suppose that the transition of $\phi(r)$ from $\phi(r=0) = \phi_e$ to $\phi(r=\infty) = 0$ occurs suddenly at $r = R$, so that very crudely:

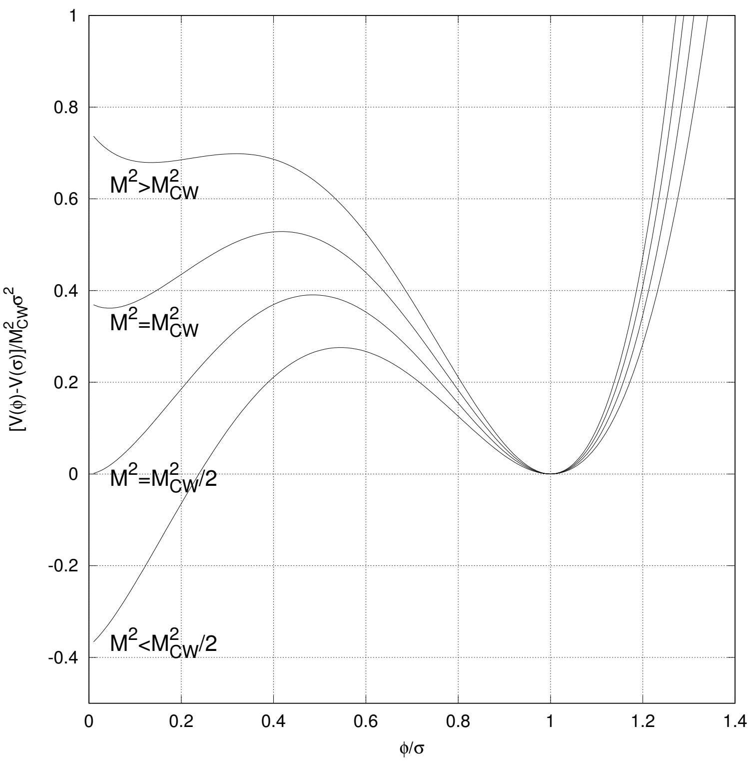

Figure 9: The one-loop Higgs potential in the electroweak model for different values of the Higgs mass M.

$\phi(r) \simeq \phi_e$ for $r \lesssim R$ and $\phi(r) \simeq 0$ for $r \gtrsim R$. When $\phi(r)$ is expressed in terms of Lorentzian time, $\phi(r) = \phi\left(\sqrt{|\vec{\mathbf{x}}|^2 - t^2}\right)$, one can see the Minkowski-space interpretation of the $O(4)$-Euclidean solution as a spherical bubble of radius R, within which $\phi(r) = \phi_e$, is nucleated at time $t_E = t = 0$, and expands outward at the speed of light. Outside the bubble, space is still in the false vacuum: $\phi = 0$. In general, a closed-form analytic solution to Eq. (4.49) cannot be found. However in the"thin-wall" approximation, where the difference in energy between the meta-stable and true vacua are small compared to the height of the barrier, it is possible to find a simple, approximate analytic expression for S_E.

The final step in the evaluation of the Euclidean action is the determination of R by the minimisation of S_E ($\mathrm{d}S_E/\mathrm{d}R = 0$).

The tunnelling rate at finite temperature is computed by following the same procedure above, remembering that field theory at finite temperature is equivalent to Euclidean field theory, periodic in imaginary time with period T^{-1} [17]. Thus, the finite-temperature tunnelling rate is found by solving Eq. (4.46) subject to the additional condition that $\phi(t_E, \vec{\mathbf{x}}) = \phi\big(t_E + T^{-1}, \vec{\mathbf{x}}\big)$, and computing S_E according to Eq. (4.48).

4.5.1 Electroweak Phase Transition

During EWPT, the full gauge symmetry of the Electroweak theory, $SU_L(2) \times U_Y(1)$ is broken down to $U_{EM}(1)$: it is when in our Universe evolution, the Electric and Weak forces differentiate.

Considering the case where the Higgs mass M is small ($M \lesssim 100\,GeV$), we can see the Higgs potential in Fig. 9, where M_{CW} is the Coleman-Weinberg mass [17]. The extrema of the potential are at $\phi = 0$ and $\phi = \sigma$. Of particular interest for the development of the phase transition is the question of whether, at zero temperature, $\phi = 0$ is also a minimum of the potential. The condition for $\phi = 0$ to be a minimum is $V''(0) \geq 0$, which is satisfied for $M^2 \leq M^2_{CW}$. Therefore, if $M \leq M_{CW}$, $\phi = 0$ is also a minimum, separated from $\phi = \sigma$ by a barrier, and the transition is first order. In the event that $M^2 \leq M^2_{CW}$ and $\phi = 0$ is a local minimum, we must require that the SSB minimum $\phi = \sigma$ have a lower free energy than the minimum at $\phi = 0$; otherwise the Universe would have remained in the symmetric minimum, and electroweak symmetry breaking would not have occurred. $V(0) \geq V(\sigma)$ is satisfied for $M^2 \geq M_{CW}^2/2$. This bound is referred to as the *Linde-Weinberg bound.*

Now consider the potential at finite temperature: the finite-temperature potential will have a temperature-dependent piece in addition to the zero-temperature part. The temperature-dependent part receives a contribution from all particles that couple to the scalar field, including the scalar field itself. It is convenient to study the phase transition in different regimes, determined by the Higgs mass.

- $1\,TeV \gtrsim M \gtrsim 100\,Gev$. In this limit the transition is second order with $T_C \simeq 500\,GeV$.
- $100\,GeV \gtrsim M > M_{CW}$. The transition is (weakly) first order. Determination of T_C is difficult. Thermal and quantum tunnelling rates are slow, and the transition does not proceed until $T = T_2 < T_C$, when the barrier vanishes and the transition can proceed classically. The transition temperature T_2, the effective temperature for the phase transition, is determined by the condition that $V_T''(0) = 0$. From a temperature of Tc until a temperature of about T_2, the Universe is trapped in the meta-stable, false-vacuum state $\phi = 0$, during which time it is said to"supercool" (just as supercooling liquid water below 0°C). When the barrier disappears ($T \simeq T_2$), the phase transition to the SSB vacuum takes place, and the vacuum energy is released, heating the Universe to a temperature of order T_C, and thereby increasing the entropy density by a factor of about $(T_C/T_2)^3$. Since for Higgs in this weight class, T_C is less than twice T_2, there is only a small amount of supercooling and associated entropy production; hence the transition is weakly first order.
- $M_{CW} \gtrsim M \gtrsim M_{CW}/\sqrt{2}$. In this range there will always be a barrier between the meta-stable and true vacuum ($T_2 \to 0$ as $M \to M_{CW}$), and the transition must proceed via tunnelling. The thin-wall approximation in this case is not valid and a numerical calculation of the bounce action is necessary. Numerical calculations give a quantum tunnelling action $S_E \geq 10^4$ for $M \lesssim M_{CW}$, and $S_E \to \infty$ for $M \to M_{CW}/\sqrt{2}$ where the two minima are exactly degenerate. Since the tunnelling rate is proportional to the exponential of the action, the transition is never completed for Higgs masses in this interval. However, there is an additional physical phenomenon that must be considered. The extreme supercooling will eventually be terminated when QCD interactions become strong, at a temperature $T \sim 200\,MeV$. Below this temperature, chiral symmetry is dynamically broken and a quark condensate forms, which is signalled by $\langle\bar{q}q\rangle$ developing a non-zero vacuum expectation value, $\langle\bar{q}q\rangle \simeq \Lambda_{QCD}^2$ (q is the quark field). Since the Higgs Yukawa coupling makes a contribution to the potential $\Delta V = -h_q\phi\bar{q}q$, the $\langle\bar{q}q\rangle$ condensate will result in an effective linear term in the Higgs potential which destabilises the symmetric minimum, thereby driving the transition. However, the large amount of supercooling will result in a big entropy release, $S_{after}/S_{before} \sim (T_C/200\,MeV)^3 \sim 10^9$, which dilutes the baryon number present before the phase transition. Such a large reduction in the baryon-to-entropy ratio is probably unacceptable, and thus cosmological considerations constrain the Higgs mass to be $\gtrsim 1.1M_{CW}$.
- $M \leq M_{CW}/\sqrt{2}$. The true ground state at zero temperature is $\phi = 0$. Since this is also the ground state at high temperature, the Universe will remain in the symmetric state, and the electroweak symmetry will never be broken, which is not the case. In sum, cosmological considerations of the electroweak phase transition result in the limit $M \gtrsim 1.1M_{CW}$

While the preceding discussion applies specifically to electroweak symmetry breaking, it provides an excellent paradigm for a generic SSB phase transition, e.g., GUT symmetry breaking.

4.6 Domain Walls

How can we tell if the Universe underwent a series of SSB phase transitions? One possibility is that symmetry-breaking transitions were not "perfect", and that false vacuum remnants were left behind, frozen in the form of topological defects: domain walls, strings, and monopoles [17]. Given the Lagrangian for a real scalar field that undergoes SSB, Eq. (4.41), we see that invariance under $\phi \to -\phi$ is spontaneously broken when ϕ takes on the vacuum expectation value $\langle\phi\rangle = +\sigma$ or $\langle\phi\rangle = -\sigma$. If space is divided into two regions, so that in one region of space $\langle\phi\rangle = +\sigma$, and in the other region of space $\langle\phi\rangle = -\sigma$, since the scalar field must make the transition from $\phi = -\sigma$ to

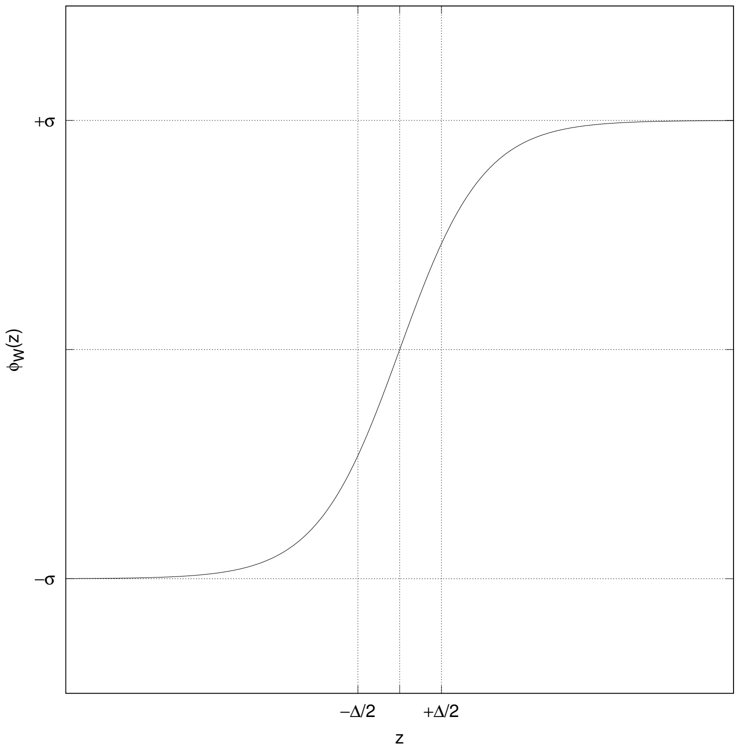

Figure 10: The solution for an infinite wall in the $x - y$ plane.

$\phi = +\sigma$ smoothly, there must be a region where $\phi = 0$, a region of false vacuum. This transition region between the two vacua is called a domain wall: domain walls can arise whenever any discrete symmetry is broken.

Consider an infinite wall in the $x - y$ plane at $z = 0$. At $z = -\infty$, $\phi = -\sigma$, and at $z = +\infty$, $\phi = +\sigma$. The equation of motion for ϕ is

$$-\frac{\partial^2\phi}{\partial z^2} + \lambda\phi\big(\psi^2 - \sigma^2\big) = 0 \tag{4.51}$$

The solution to the equation of motion, subject to the boundary conditions above, is

$$\phi_W(z) = \sigma \tanh\left(\frac{z}{\Delta}\right) \tag{4.52}$$

where the "thickness" of the wall is characterised by $\Delta = (\lambda/2)^{-1/2}\sigma^{-1}$. This solution is illustrated in Fig. 10

The domain wall is topologically stable. The finite, but non-zero, thickness of the wall is easy to understand. The terms contributing to the surface energy density include a gradient term $\sim \sigma^2/\Delta$, and a potential energy term $\sim \Delta\lambda\sigma^4$. The gradient term is minimised by making the wall as thick as possible, while the potential term is minimised by making the wall as thin as possible. The balance between these terms results in a wall of thickness $\Delta \sim \lambda^{-1/2}\sigma^{-1}$. Domain walls are inherently relativistic, and their gravitational effects are inherently non-Newtonian. To understand the production mechanism, assume that at the time of the phase transition, the correlation length ξ of the scalar field vacuum expectation value (VEV) is finite. For any two points separated by a distance $D \gtrsim \xi$ the VEV is uncorrelated, and there is a 50% chance that the two points will be in different vacua. If they are, they must be separated by a domain wall. Therefore, a phase transition associated with the SSB of a discrete symmetry should lead to a network of domain walls, with walls separated by a typical distance ξ. In addition, a typical wall in the network will be curved, with curvature radius characterised by ξ. Within the network there will be both infinite walls and finite walls. Just as a membrane with a surface tension evolves to minimise its surface area, a curved wall will evolve so as to minimise its surface area. The evolution of a wall can, in principle, be followed by solving the equations of motion for the scalar field itself. It is easier to use the equations of motion for an idealised thin wall. One finds that the dynamics of the expansion and evolution of the wall system depend upon the average velocity of the wall gas.

For non-relativistic walls, the energy density is proportional to R^{-1}. In this limit the wall gas stretches conformally with the expansion, with the distance between walls increasing as R, the surface area of the walls growing as R^2, and the number density of walls decreasing as R^{-3}. Since the mass of a wall is proportional to its area, and the wall energy density is proportional to the product of the wall mass and the wall number density, one finds that in fact the energy density of the walls decreases as R^{-1}. One expects the velocities associated with any domain walls created in the early Universe to be rapidly red shifted away, so cosmological walls should quickly come to dominate both the radiation and the matter energy densities, and thereby drastically alter the standard cosmology. However, the existence of large-scale domain walls in the Universe today can be precluded simply based upon their contribution to the total mass density. Walls would also lead to large fluctuations in temperature of the Cosmic Microwave Background (CMBR) unless σ is very small.

4.7 Electroweak Baryogenesis

The standard electroweak theory possesses all the necessary ingredients for generating the BAU. First, anomalous B non conservation is strong enough at high temperatures, and second, there is CP non conservation coming from Kobayashi-Maskawa mixing. Finally, large deviations from thermal equilibrium occur during the first-order phase transition associated with the breaking of $SU_L(2) \times U_Y(1)$ symmetry (see section 4.5.1) [19].

Baryon number violation in the SM follows from the EW anomaly

$$\partial_\mu j_B^\mu = N_f \partial_\mu K^\mu = N_f \frac{\alpha_W}{8\pi} \operatorname{Tr} F\tilde{F} \tag{4.53}$$

where N_f is the number of fermionic generations and F is the EW field strength. Variations in the baryon number are related to variations in the Chern-Simons number by $\Delta B = N_F \Delta N_{CS}$, where $N_{CS} = \int \mathrm{d}^3x\, K^0$ is the Chern-Simons number (see sections 3 and A.5).

Assuming we have some free energy difference ΔF between the states on either side of the barrier (see section 4.4), and a rate Γ for sphaleron transitions between neighbouring minima, the ratio of the rates Γ_+, Γ_- to go over the barrier in the two different directions (see Fig. 6) is [11]

$$\frac{\Gamma_+}{\Gamma_-} = e^{-\frac{1}{T}\Delta F} \tag{4.54}$$

which leads to the master equation for Baryogenesis

$$\frac{\mathrm{d}n_B}{\mathrm{d}t} = \Gamma_+ - \Gamma_- = -\frac{\Gamma}{T}\Delta F \tag{4.55}$$

Using

$$\Delta F = N_f \frac{\partial F}{\partial B} \tag{4.56}$$

we get

$$\dot{n}_B = -\frac{\Gamma}{T}\frac{\partial \mathcal{F}}{\partial N_{CS}} = \frac{\Gamma}{T}\mu_{CS} \tag{4.57}$$

where μ_{CS} is the chemical potential from CP-violating sources inducing a non-vanishing B number. The generated Chern-Simons number asymmetry can then be written

$$\langle N_{CS}\rangle(t) = \frac{1}{T_{eff}} \int_0^t \mathrm{d}t'\, \Gamma_{CS}(t')\mu(t') \tag{4.58}$$

where T_{eff} characterises the temperature at which Chern-Simons transitions are operative at the same time as efficient CP-violation effects [26].

The procedure of calculating the baryon asymmetry is technically very complicated and involves several steps. First, one has to construct an effective action for the gauge and Higgs fields integrating out the fermionic degrees of freedom. Second, one should generate a number of equilibrium gauge-Higgs configurations just before the phase transition. Then, it is necessary to solve classical equations of motion arising from the P- and CP-non-invariant effective action with the use of relevant thermal equilibrium configurations as initial conditions. In some cases, however, it is possible to estimate the produced asymmetry [19]. Suppose that the effective action has the form

$$\mathcal{L}_{eff} = \frac{\alpha_W}{8\pi}\Phi(\phi)\,\mathrm{Tr}\,F\tilde{F} \tag{4.59}$$

where $\Phi(\phi)$ is some time-varying function of fields which depends on the underlying Baryogenesis model. It is then relatively straightforward to evaluate the contribution to CP violation in Baryogenesis for various forms of the functional $\Phi(\phi)$. The size of CP violation ultimately depends on both the size of the coefficient in $\Phi(\phi)$ and the degree of time variation during Baryogenesis. In general, if the only time-variation in $\Phi(\phi)$ comes from the time-variation of T, then μ_{CS} is suppressed by a factor encoding the rate of Hubble expansion. Its only during a first-order phase transition that we expect more rapid time variation of the quantity $\Phi(\phi)$, and hence an unsuppressed chemical potential for N_{CS} [10].

Although there are a variety of non-equilibrium phenomena that may lead to Baryogenesis, we will focus here on Baryogenesis during the electroweak phase transition, under the assumption that the phase transition is first-order and sufficiently strong. In principle, the phase transition proceeds through bubble nucleation, and the generation of baryon number asymmetry occurs only in the bubble wall when $\mu = \dot{\Phi}(\phi)$ is significant. In practice, however, it is difficult to make any precise quantitative statements about a nucleation phase transition without extensive numerical simulation. One must understand not only the details of the phase transition and bubble wall propagation, but also particle transport at the phase boundary in the presence of CP violation. Qualitative estimates may be made if we consider a spinodal decomposition phase transition, in which the scalar field rolls uniformly to the true vacuum; such a transition gives a spatially uniform, but time-varying, phase transition. The assumption is that the results should be similar to those of a nucleation phase transition, as Lorentz invariance in principle relates processes with time-varying fields to those with space-varying fields.

We have

$$\int \mathrm{d}^4x\,\frac{\alpha_W}{8\pi}\Phi(\phi)\;\mathrm{Tr}\,F\tilde{F} = \int \mathrm{d}^4x\,\Phi(\phi)\,\partial_\mu K^\mu = -\int \mathrm{d}t\,\partial_t\Phi(\phi)\int \mathrm{d}^3x\,K^0 \tag{4.60}$$

where we made an approximation in which $\Phi(\phi)$ is replaced by its spatial average $L^{-3}\int \mathrm{d}^3x\,\Phi(\phi)$ and we integrated by parts in order to exhibit the chemical potential for Chern-Simons number

$$\mu = \frac{\mathrm{d}}{\mathrm{d}t}\Phi(t) \tag{4.61}$$

Therefore, the time derivative of Φ can be interpreted as a time-dependent chemical potential for Chern-Simons number and $\mathcal{L}_{eff}$ takes the form

$$\mathcal{L}_{eff} = \mu N_{CS} \tag{4.62}$$

The produced B asymmetry is given by

$$n_B = N_f\int \mathrm{d}t\,\frac{\Gamma\mu}{T} \sim N_F\frac{\Gamma(T_{eff})}{T_{eff}}\Delta\Phi \tag{4.63}$$

The transition rate for fluctuations between neighbouring minima depends entirely whether electroweak symmetry is broken. In the unbroken phase, sphaleron transitions are unsuppressed, while in the broken phase there arises the usual exponential suppression. The transition rates in the symmetric and broken phases are, respectively

$$\begin{aligned}\Gamma &\simeq 30\,\alpha_w^5 T^4\\ \Gamma &\simeq 30\,(\alpha_w T)^{-3} m_W^7\exp\left(-\frac{E_{sph}}{T}\right)\end{aligned} \tag{4.64}$$

where α_w is the weak fine structure constant, E_{sph} is the sphaleron potential energy and m_W is the value of the W boson mass which is varying through the wall [9]. Clearly the contribution from the symmetric phase dominates. In the spinodal decomposition transition, we may assume that these values interpolate smoothly. Treating Γ a as a step function and T as essentially constant during the phase transition, we can estimate the integral to find a total baryon asymmetry. Using the sphaleron rate in the EW symmetric phase $N_f\Gamma = 30\,\alpha_w^5 T^4 \sim \alpha_w^4 T^4$, this leads to

$$\frac{n_B}{s} = N_f \alpha_w^4 \left(\frac{T_{eff}}{T_{reh}}\right)^3 \Delta\Phi\, \frac{45}{2\pi^2 g_*(T_{reh})} \sim 10^{-7}\left(\frac{T_{eff}}{T_{reh}}\right)^3 \Delta\Phi \tag{4.65}$$

where T_{reh} is the reheat temperature after the EWPT. It may be significantly higher than the temperature of the EWPT, T_{EWPT}, if the EWPT was delayed and completed after a super-cooling stage (see section 4.5.1). For standard EW Baryogenesis, $T_{eff} = T_{EWPT} = T_{reh}$. In contrast, the key-point for cold Baryogenesis is that $T_{eff} \neq T_{EWPT}$. It is significantly higher than the temperature of the EWPT. It is a way to express the very efficient rate of B violation in terms of the equilibrium expression $\Gamma \sim \alpha_w^4 T^4$ although the system is very much out-of equilibrium. This estimate should be compared to the observed value, $n_B/s \simeq 10^{-10}$. Although this expression has been obtained for a spinodal decomposition transition with a number of simplifying assumptions, it is parametrically similar to the analogous expression for bubble nucleation. In the case of bubble nucleation, the parametric dependence on $\Delta\Phi$ is unaltered [10].

In the next section we will describe a procedure to compute the size of CP violation $\Delta\Phi$ given a particular functional Φ proposed in [10, 26].

4.8 Axion-induced CP violation

We will investigate whether the large values of the effective vacuum angle in Eq. (1.3) at early times can have any implications for EW Baryogenesis. We have

$$\bar\theta = \frac{a}{f_a} \sim \mathcal{O}(1) \quad \text{for} \quad T \gtrsim 1\,\text{GeV} \tag{4.66}$$

and then $\bar\theta$ quickly drops as the axion gets a mass $m_a^2 \propto \Lambda_{QCD}^3$ at the QCD phase transition [10], where Λ_{QCD} is the strong sector confinement scale, and starts oscillating around the minimum of its potential[17]. The axion Lagrangian reads

$$\mathcal{L}_a = \mathcal{L}(\partial_\mu a) - \frac{1}{2}\partial^\mu a \partial_\mu a + \frac{a}{f_a}\frac{\alpha_s}{8\pi} G\tilde G \tag{4.67}$$

where the last term describes axion coupling to the QCD field-strength G (see Section C). So that (compare with Section 3)

$$\frac{\partial V_{eff}}{\partial a} = -\frac{1}{f_a}\frac{\alpha_s}{8\pi} G\tilde G \tag{4.68}$$

Gluon condensation from $SU(3)$ instantons leads to a VEV for $G\tilde G$ and a potential for the axion that can be written as

$$V = f_\pi^2 m_\pi^2 \left(1 - \cos\frac{a}{f_a}\right) \approx f_a^2 m_a^2 \left(1 - \cos\frac{a}{f_a}\right) \tag{4.69}$$

As a result

$$\frac{\alpha_s}{8\pi}\left\langle G\tilde G\right\rangle = f_a^2 m_a^2 \sin\bar\theta \tag{4.70}$$

To make a connection between the axion and EW Baryogenesis, we have to construct an effective operator gathering gluons and EW gauge bosons. The main point of the previous section can be summarised, according to equations (4.59) (4.62) (4.61), as

$$\mathcal{L}_{eff} = \frac{\alpha_W}{8\pi}\Phi(T)\,\mathrm{Tr}\,F\tilde F \quad \leftrightarrow \quad \mathcal{L}_{eff} = \mu N_{CS} \quad \text{where} \quad \mu = \frac{\mathrm{d}}{\mathrm{d}t}\Phi(T) \tag{4.71}$$

An operator of the type (4.59) can arise, where Φ is controlled by the axion mass squared. In particular, the η' meson, which is a singlet under the approximate $SU(3)$ flavor symmetry of strong

[17] Classical oscillations of the axion background field around the minimum of the potential start at T_i when its mass is of the order of the Hubble scale $m_a(T_i) \sim 3H(T_i) \sim \Lambda_{QCD}^2/M_{Plank}$. The energy stored in these axion oscillations redshifts as non relativistic matter, and eventually behave exactly as cold dark matter, according to $\mathrm{d}^2\langle a\rangle/\mathrm{d}t^2 + 3H(T_i)\,\mathrm{d}\langle a\rangle/\mathrm{d}t + m_a^2(t)\langle a\rangle = 0$

interactions, can couple to both $G\tilde{G}$ and $F\tilde{F}$. At temperatures below the η' mass, $m_{\eta'} \approx 958$ MeV, we can use the effective operator

$$\mathcal{L}_{eff} = \frac{1}{M^4}\frac{\alpha_s}{8\pi}G\tilde{G}\frac{\alpha_w}{8\pi}F\tilde{F} \tag{4.72}$$

where $1/M^4 = 10/\big(F_\pi^2 m_{\eta'}^2\big)$. We end up with

$$\mathcal{L}_{eff} = \frac{1}{M^4}f_a^2 m_a^2(T)\sin\bar{\theta}\frac{\alpha_w}{8\pi}F\tilde{F} \tag{4.73}$$

hence

$$\Phi(T) = \frac{1}{M^4}f_a^2 m_a^2(T)\sin\bar{\theta} \quad \rightarrow \quad \mu = \frac{\mathrm{d}\Phi}{\mathrm{d}t} = \frac{f_a^2}{M^4}\frac{\mathrm{d}}{\mathrm{d}t}\big[m_a^2(T)\sin\bar{\theta}\big] \tag{4.74}$$

So that we can see, the time variation of the axion field and/or mass is a source for Baryogenesis:

$$n_B \propto \int \mathrm{d}t\,\frac{\Gamma(T)}{T}\frac{\mathrm{d}}{\mathrm{d}t}\big[m_a^2(T)\sin\bar{\theta}\big] \tag{4.75}$$

In general, the axion mass is very suppressed at temperatures above the QCD scale. A large B asymmetry is therefore produced only if the EWPT occurs not much earlier than the QCD phase transition. Also, if the baryon asymmetry of the Universe (BAU) could really be created due to the existence of a non-vanishing axion field at high temperatures, then one can expect that there is some relation between the observed magnitude of the BAU and the amount of the axionic dark matter. However, the calculated value of baryon asymmetry turns out to be the right order of magnitude only in the case of a maximal initial misalignment angle and when there is no reheating after the phase transition. The last condition requires a very small value for the Higgs-boson mass, $M_H \sim 100\,KeV$, but the model of this kind seems to be (almost) excluded by experiments at CERN, because experimentally $M_H \sim 125, GeV$ [1]. Which for some authors is sufficient to conclude that the electroweak Baryogenesis seems impossible to realise in the simplest versions of invisible axion models [19]. Others are more optimistic, and include the possibility that in the context of cold Baryogenesis the effective temperature characterising baryon number violation may be significantly higher than the actual temperature of the Universe [26]. In this case an EW scale dilaton provides a new route for departure from thermal equilibrium and can drive a parametric amplification of baryon number violation at low temperatures. This opens the possibility that the strong CP violation via the QCD axion could be responsible for the baryon asymmetry of the universe during a delayed EWPT.

5 The Quest for Axions

A priori the mass of the axion (or equivalently, the PQ symmetry breaking scale) is arbitrary. All values solve the strong CP problem equally well: the axion mass only determines the curvature of the potential that holds $\bar{\theta}$ at zero, and the strength of the axions interactions, as we will see below. If we arbitrarily set the symmetry breaking scale f_a be somewhere between $100\,GeV$ and $10^{19}\,GeV$, the axion mass then lies between about $1\,MeV$ and $10^{-12}\,eV$ a span of eighteen orders of magnitude to look for the particle [33].

As a consequence, an axion model has only one free parameter (compare with Section 3.2): the axion mass, or equivalently the PQ symmetry breaking scale. They are related by

$$m_a = \frac{\sqrt{z}}{1+z}\frac{f_\pi m_\pi}{f_a/\xi} \simeq \frac{0.62\,eV \times 10^7\,GeV}{f_a/\xi} \tag{5.1}$$

where $z = m_u/m_d \simeq 0.56$, $m_\pi = 135\,MeV$ and $f_\pi = 93\,MeV$ are the pion mass and decay constant, ξ is the color anomaly of the PQ symmetry. The axion field a is related to $\bar{\theta}$ by $a = (f_a/\xi)\bar{\theta}$. The Feynman diagrams associated with these interactions are shown in Fig. 11.

The most significant feature of all the axion couplings g_{aii} is that they are proportional to $1(f_a/\xi)$ or equivalently m_a: the smaller the axion mass, or the larger the PQ SSB scale, the more weakly the axion couples. The coupling of the axion to the photon arises through the electromagnetic anomaly of the PQ symmetry, and allows the axion to decay to two photons.

5.1 Laboratory searches

The original axion proposed by Peccei and Quinn was based upon a PQ symmetry breaking scale equal to that of the weak scale ($f_{PQ} \sim 250\,GeV$), leading to an axion mass of about $200\,keV$. For that reason, its interactions were roughly semi-weak, making it very accessible to laboratory searches. This kind of axion was quickly ruled out. Shortly after that, the axion was made "invisible" by raising the symmetry breaking scale f_{PQ} to a value much greater than $250\,GeV$ and by that reducing the coupling strength of the axion. Once the weak scale was excluded from research, all values of f_{PQ} became more or less equally attractive. Among the laboratory searches, the most sensitive were [33]

- The kaon decay process $K^+ \to \pi^+ + a$, where the axion is not found. The present experimental upper limit to the branching ratio for $K^+ \to \pi^+ + \text{nothing}$ is 3.8×10^{-8}. In an axion model this process arises either through axion/pion mixing (the decay $K^+ \to \pi^+ + \pi^0$ is observed) or an off-diagonal coupling of the axion to s and d quarks ($s \leftrightarrow d + a$).
- The decays of quarkonium ($Q\bar{Q}$) states, $J/\psi \to a + \gamma$, $Y \to a + \gamma$. The upper limits of the branching ratios for these two processes are 1.4×10^{-5} and 3×10^{-4}, respectively.

Based upon the three processes just mentioned, one can safely conclude that, if an axion exists, it must be characterised by

$$f_{PQ} \geq 10^3\,GeV \quad \text{or} \quad m_a \lesssim 6\,keV \tag{5.2}$$

While laboratory-based experiments rule out axion masses in the range of about $10\,keV$ to $1\,MeV$ and most certainly exclude the original axion, they leave open an enormous window, $10^{-12}\,eV$ to $10\,keV$, one which has only been explored by exploiting the astrophysical and cosmological effects of the axion.

5.2 Stars

Given the intrinsically short timescale associated with nuclear interactions, it is surprising that stars live as long as they do. A star like our sun can live burning hydrogen for 10^{10} years. The reason for stellar longevity is that the rate at which a star can liberate its nuclear free energy is controlled not by its nuclear reaction rates, but rather by the rate at which the nuclear energy released inside can be transported through the star and radiated outside, into the space. Under the conditions that exist in a typical star, like our Sun, the mean free path of a photon is only about $1\,cm$, and the time required for a photon emitted at the center of the sun to make its way to the surface is $\sim 10^7$ years. Ordinary matter is very opaque to photons because of the strength

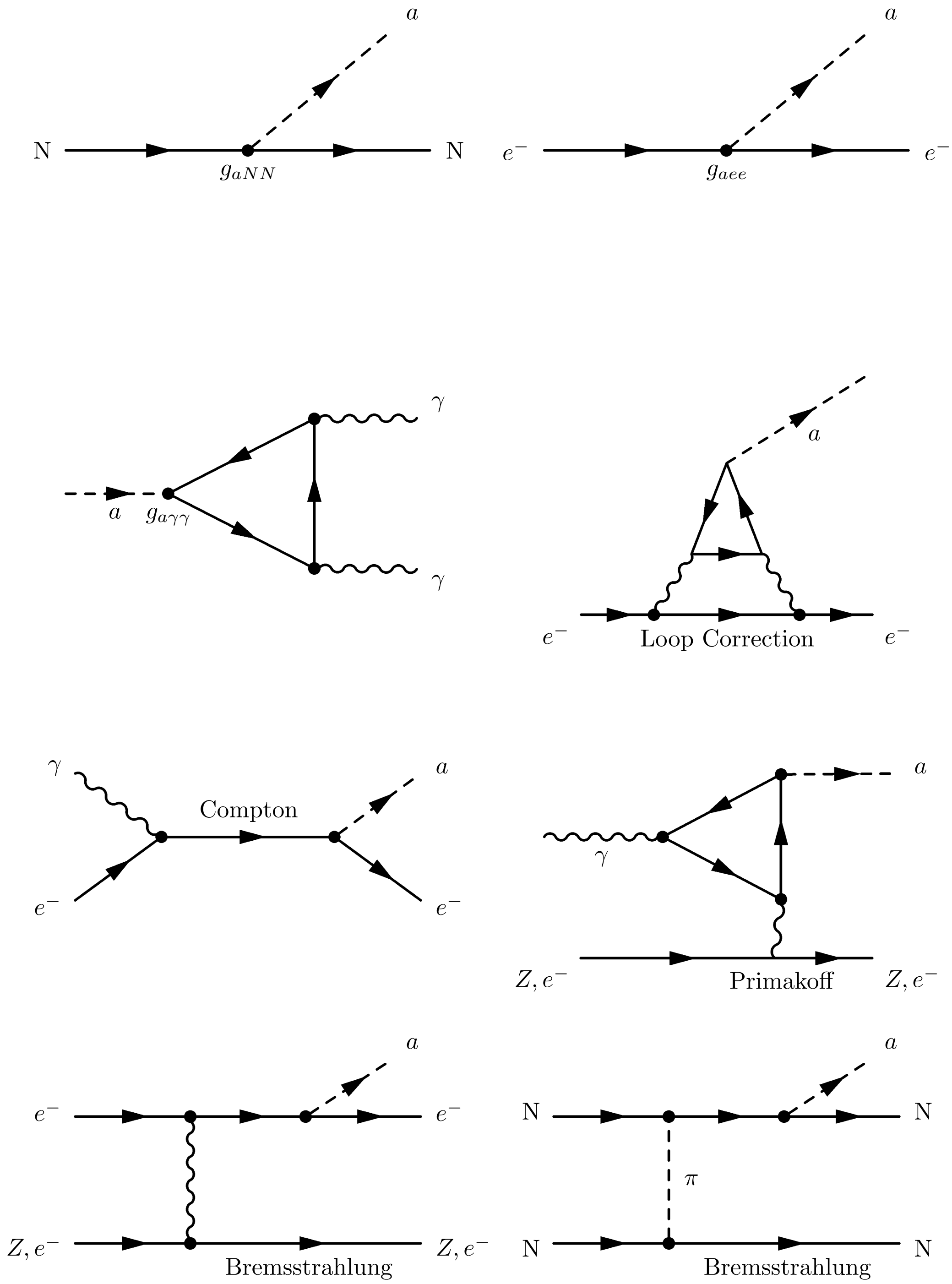

Figure 11: Axion couplings to ordinary matter (electrons, nucleons, and photons) and the dominant axion emission processes in stars. The dominant emission processes in main sequence stars, red giant stars and white dwarfs are the Compton-like and bremsstrahlung processes, the rates for both of which are proportional to the axion-electron coupling squared. For the hadronic axion the dominant emission process in these objects is the Primakoff process, owing to the fact that the tree-level axion-electron coupling is highly suppressed. In neutron stars the dominant emission process for both types of axions is nucleon-nucleon axion bremsstrahlung. In the early Universe the dominant axion production process is axion-pion conversion $\pi + N \to a + N$, which corresponds to the Feynman diagram for $N + N \to N + N$ with the lower nucleon line removed.

of the electromagnetic interactions. The Thomson cross section is $\sigma \simeq 0.67 \times 10^{-24}\, cm^2$. Stellar longevity is explained by the large interaction cross section of the photon with ordinary matter. The existence of a light, weakly interacting particle has the potential to greatly accelerate the evolutionary process of stars of all types by more efficiently carrying energy away, and by that shortening their lifetimes. To effectively remove the free energy liberated in the nuclear reactions in a star, the axion must interact weakly enough so that it streams right out without interacting, but strongly enough so that it is produced in sufficient numbers to move away large amounts of energy. The "optimal" interaction strength is such that the interaction length is comparable to the size of the astrophysical object.

Neutrinos are examples of efficient "star coolants": neutrino emission is the primary cooling mechanism for a newly born neutron star, as they remove the greatest part of the gravitational binding energy that is released in the formation of the star. Based upon the observed duration of the neutrino bursts detected in the SN 1987A supernova, we can take the conclusion that the number of light neutrino flavors must be less than about nine. For more than nine neutrino flavors the cooling of the neutron star would have proceeded too rapidly, shortening the neutrino burst by about the same factor, in contradiction to the observations of the Kamiokande II and Irvine-Michigan-Brookhaven water Cherenkov detectors [33].

The potential effect of axion emission on stars is the acceleration of their evolution and shortening of their lifetimes. In main sequence and red giant stars the primary axion emission processes are

- The Compton-like process $\gamma + e^- \rightarrow a + e^-$.
- The Axion bremsstrahlung $e^- + Z \rightarrow a + e^- + Z$
- The Primakoff process $\gamma + Z\,(\text{or}\, e^-) \rightarrow a + Z\,(\text{or}\, e^-)$

The first two are proportional to $g_{aee}^2 \propto m_a^2$. Of lesser importance, unless g_{aee} vanishes at tree level, as it does for a hadronic axion, is the third one. In very low-mass stars ($M \lesssim 0.2\, M_\odot$) emission through the axio-electric effect (the analogue of the photo-electric effect) is also very important. Roughly speaking, if one "turns on" axion emission in a star, the star contracts to raise its temperature and nuclear energy liberation rate to balance its axion energy losses. In the process the star also raises its photon luminosity, and as a result of both axion emission and enhanced photon emission its lifetime is shortened. It is found that the axion mass limits that follow from the Sun is $m_a \lesssim 1\, eV$.

5.3 SN 1987A

SN 1987A was triggered by the gravitational collapse of the blue super giant Sanduleak $-69^\circ\, 202$. During the catastrophic collapse of the iron core about $3 \times 10^{53}\, erg$ of energy was radiated in the form of thermal neutrinos of all three types. If an axion exists, it can play an important role in the cooling of this nascent neutron star. Axions carry away energy from the core and thus accelerate the cooling of the core. The observable effect of axion cooling is to shorten the neutrino burst. Under the conditions that existed in the post-collapse core the dominant axion emission process is nucleon-nucleon axion bremsstrahlung. The axion coupling relevant for this process is that to nucleons, of order $m_N(f_a/\xi) \simeq 10^{-7}(m_a/eV)$. Based upon the observed length of the neutrino burst associated with SN 1987A, an axion of mass $10^{-3}\, eV$ to $2\, eV$ can be excluded.

The SN l987A axion mass constraint is the most stringent astrophysical constraint to the axion mass. The known uncertainties in the axion emission rate amount to no more than a factor of 10. Since that rate itself scales as the axion mass squared, the uncertainty in the axion mass bounds amounts to no more than a factor of 3. Moreover, because the physics of the cooling of the nascent neutron star is so simple and the observable (the neutrino burst) is so clean and direct, the SN 1987A constraint is probably the most reliable of the astrophysical bounds [33].

6 Conclusions

In this paper, a mechanism was exposed to explain the asymmetry between matter and antimatter in the Universe. In particular, we have shown that the QCD axion could play a key role in providing a source of CP violation in Baryogenesis.

We have seen that the solution to the $U(1)$ problem came from the realisation that Quantum Chromodynamics has a rich structure, characterised by a topological integer or winding number N_{CS}. Instanton solutions provide a mechanism of tunnelling between topological inequivalent n-vacua. So that as a consequence the theory possesses a hidden parameter, the vacuum angle θ, which implies violation of CP invariance in strong interactions. Since it is experimentally verified that the strong interaction is CP-invariant, the solution to the $U(1)$ problem causes another problem, dubbed the “strong CP problem” i.e. why in Nature θ has a very small value. Peccei-Quinn proposed a solution that introduces a new additional chiral symmetry which allows naturally to set $\theta = 0$. However, the symmetry is accompanied by a new Nambu-Goldstone boson, the ”axion” that has so far eluded detection.

Several theories try to incorporate the axion idea to solve the BAU riddle. In this report, we focused on Electroweak Baryogenesis, a theory that has all the elements for generating the baryon asymmetry, and we analysed a model that makes use of θ to leverage the effect of the Electroweak theory to generate an optimal violation of CP invariance.

In my opinion, the Peccei-Quinn solution is a good candidate to solve the strong CP problem. It is an elegant and ingenious theory, that introduce a seemingly simple explanation to one of the unsolved problems in physics. Between the various solutions proposed for the strong CP problem, is the only one that provides a “strong” solution, since it imposes a symmetry on the Lagrangian from the outset so that $\theta = 0$ initially and remains stable after the symmetry breaking. All other theories, apart from those based on Supersymmetry, have serious problems or were ruled out due to inconsistencies with observed quantities.

However, the axion has never been detected despite extensive search. Laboratory searches and cosmological considerations have helped narrowing down the possible values its mass may have. New experiments will be performed soon at the LHC that could provide the definitive answer. The next few years will be crucial to confirm or reject the theory.

Acknowledgements

I would like to express my sincere gratitude to Professor Kellogg Stelle for agreeing to supervise this dissertation and for his advice during its preparation and writing. I am really grateful for this opportunity to learn more about a fascinating subject.

A special thanks go to my wife Maria, for her support and dedication throughout this course of studies and my life in general.

Appendices

A The Chiral Anomaly

In quantum physics an *anomaly* is the failure of a symmetry of a theory's classical action to be a symmetry of any regularization of the full quantum theory. In particular, a chiral anomaly is the anomalous non-conservation of a chiral current. In some theories of fermions with chiral symmetry, the quantisation may lead to the breaking of this (global) chiral symmetry. In that case, the charge associated with the chiral symmetry is not conserved.

A.1 Current conservation and Ward Takahashi identities

We work in the usual spinor electrodynamics described by the Lagrangian density [3]

$$\mathcal{L}_{QED} = \bar{\psi} i\gamma^\mu(\partial_\mu - ieA_\mu)\psi - m\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + B\partial^\mu A_\mu + i\partial^\mu \bar{c}\partial_\mu c \tag{A.1}$$

where $F_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu$. The last two terms in $\mathcal{L}_{QED}$ stand for the gauge fixing term and the *Faddeev Popov ghost* term.

We can form two currents from the vector and pseudo-vector Dirac field bilinears [24]

$$j^\mu(x) = \bar{\psi}(x)\gamma^\mu\psi(x) \quad j^{\mu 5}(x) = \bar{\psi}(x)\gamma^\mu\gamma^5\psi(x) \tag{A.2}$$

Computing the divergences of these currents, using the equations of motion which follow from Eq. (A.1) gives the two equations

$$\partial_\mu j^\mu = 0 \tag{A.3}$$
$$\partial_\mu j^{\mu 5} = 2im\bar{\psi}\gamma^5\psi \tag{A.4}$$

Thus j^μ is always conserved if $\psi(x)$ satisfies the Dirac equation. When we couple the Dirac field to the electromagnetic field, j^μ will become the electric current density. If $m = 0$, the current $j^{\mu 5}$, called the *axial vector current* is also conserved.

In gauge theories such as quantum electrodynamics, it is of fundamental importance to perform calculations by preserving gauge invariance in any finite order of perturbation theory. The important relations which express the gauge invariance of the starting Lagrangian in terms of Green's functions are called Ward-Takahashi (WT) identities [14].

We write the generating functional of correlation functions as

$$\begin{aligned} Z(J^\mu, \eta, \bar{\eta}) &= \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}A_\mu\mathcal{D}B\mathcal{D}c\mathcal{D}\bar{c}\exp\left[i\int \mathrm{d}^4x\,(\mathcal{L}_{QED} + \mathcal{L}_J)\right] \\ &\equiv \int \mathrm{d}\mu\exp\left[i\int \mathrm{d}^4x\,(\mathcal{L}_{QED} + \mathcal{L}_J)\right] \end{aligned} \tag{A.5}$$

where the source terms are defined by

$$\mathcal{L}_J \equiv -A_\mu(x)J^\mu(x) + \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x) \tag{A.6}$$

The WT identity resulting from the gauge invariance of the starting action is derived by considering the following change of integration variables (gauge transformation)

$$\begin{aligned} \psi'(x) &= e^{i\alpha(x)}\psi(x) \\ \bar{\psi}'(x) &= \psi(x)e^{-i\alpha(x)} \end{aligned} \tag{A.7}$$

and the following identity in the path integral formulation holds

$$\begin{aligned} &\int \mathrm{d}\mu\exp\left\{i\int \mathrm{d}^4x\left[\mathcal{L}_{QED}(\bar{\psi},\psi,A_\mu,B,c,\bar{c}) + \mathcal{L}_J(\bar{\psi},\psi,A_\mu)\right]\right\} \\ &= \int \mathrm{d}\mu'\exp\left\{i\int \mathrm{d}^4x\left[\mathcal{L}_{QED}(\bar{\psi}',\psi',A_\mu,B,c,\bar{c}) + \mathcal{L}_J(\bar{\psi}',\psi',A_\mu)\right]\right\} \end{aligned} \tag{A.8}$$

This derivation is based on the fact that the value of a definite integral does not depend on the naming of integration variables. This relation when combined with the invariance of the path integral measure under the above change of variables

$$\mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}A_\mu\mathcal{D}B\mathcal{D}c\mathcal{D}\bar{c} = \mathcal{D}\bar{\psi}'\mathcal{D}\psi'\mathcal{D}A_\mu\mathcal{D}B\mathcal{D}c\mathcal{D}\bar{c} \tag{A.9}$$

and the change of the action in the order linear in $\alpha(x)$

$$\begin{aligned}\mathcal{L}_{QED}\left(\bar{\psi}', \psi', A_\mu, B, c, \bar{c}\right) &= \mathcal{L}_{QED}\left(\bar{\psi}, \psi, A_\mu, B, c, \bar{c}\right) - \partial_\mu\alpha(x)\bar{\psi}\gamma^\mu\psi \\ \mathcal{L}_J\left(\bar{\psi}', \psi', A_\mu\right) &= \mathcal{L}_J\left(\bar{\psi}, \psi, A_\mu\right) - i\alpha(x)\bar{\psi}(x)\eta(x) + i\alpha(x)\bar{\eta}(x)\psi(x)\end{aligned} \tag{A.10}$$

finally gives the identity

$$\int \mathrm{d}^4x\, \alpha(x) \left\langle \partial_\mu\left[\bar{\psi}(x)\gamma^\mu\psi(x)\right] - i\bar{\psi}(x)\eta(x) + i\bar{\eta}(x)\psi(x)\right\rangle_J = 0 \tag{A.11}$$

by keeping the terms linear in $\alpha(x)$. We also performed a partial integration by using the fact that $\alpha(x)$ is local, namely, it vanishes at space-time infinity. We also defined

$$\begin{aligned}\langle O(x)\rangle_J &\equiv \langle 0, +\infty|\hat{O}(x)|0, -\infty\rangle_J \\ &= \int \mathrm{d}\mu\, O(x) \exp\left[i\int \mathrm{d}^4x\, (\mathcal{L}_{QED} + \mathcal{L}_J)\right]\end{aligned} \tag{A.12}$$

for a general operator $\hat{O}(x)$.

If one chooses $\alpha(x)$ in the identity (A.11) to be a function with a δ-functional peak at x, one obtains

$$\partial_\mu \left\langle \bar{\psi}(x)\gamma^\mu\psi(x)\right\rangle_J = i\left\langle \bar{\psi}(x)\eta(x)\right\rangle_J + i\left\langle \bar{\eta}(x)\psi(x)\right\rangle_J \tag{A.13}$$

which is the basic relation known as a WT identity.

If one functionally differentiates this identity with respect to $J^\nu(y)$ once and then sets all the sources to be 0, one obtains

$$\partial^x_\mu \left\langle T\bar{\psi}(x)\gamma^\mu\psi A_\nu(y)\right\rangle = 0 \tag{A.14}$$

The Fourier transformation of this last, relation means that the. probability amplitude for electron-positron pair creation from the current $j^\mu(x) = \bar{\psi}(x)\gamma^\mu\psi(x)$ and then pair annihilation into a photon vanishes when multiplied by the momentum. In common language, this corresponds to the current conservation of Eq. (A.3).

A.2 Quantum breaking of chiral symmetry

On the basis of the simplest gauge theory, quantum electrodynamics, we perform a calculation of the quantum breaking of chiral symmetry (or chiral anomaly) associated with an axial-vector current. The Euclidean[18] generating functional of correlation functions for quantum electrodynamics is given by

$$Z(J^\mu = 0, \eta = 0, \bar{\eta} = 0) = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi[\mathcal{D}A_\mu] \exp\left(\int \mathrm{d}^4x\, \mathcal{L}_{QED}\right) \tag{A.15}$$

where we set the source term at $\mathcal{L}_J = 0$. We also defined

$$[\mathcal{D}A_\mu] = \mathcal{D}A_\mu\mathcal{D}B\mathcal{D}c\mathcal{D}\bar{c} \tag{A.16}$$

in the path integral measure. Namely, the measure $[\mathcal{D}A_\mu]$ includes all the terms related to gauge fixing. We consider the following change of path integral variables

[18] Euclidean space-time is obtained upon analytic continuation of physical amplitudes for imaginary time: we replace t by ix_4 with now x_4 a real coordinate. This is exactly what one needs to do if one wishes to compute a tunnelling amplitude

$$\begin{aligned} \psi'(x) &= e^{i\alpha(x)\gamma^5}\psi(x) \\ \bar{\psi}'(x) &= \bar{\psi}(x)e^{i\alpha(x)\gamma^5} \end{aligned} \tag{A.17}$$

which is called the *chiral transformation.* Here $\psi'(x)^\dagger = \psi(x)^\dagger e^{-i\alpha(x)\gamma^5}$ but because of the anti-commuting property of γ^5 and γ^0, we have the expression above. We have an identity under this change of variables

$$\begin{aligned} &\int \mathcal{D}\bar{\psi}\mathcal{D}\psi[\mathcal{D}A_\mu]\exp\left[\int \mathrm{d}^4x\,\mathcal{L}_{QED}\left(\bar{\psi},\psi A_\mu\right)\right] \\ &= \int \mathcal{D}\bar{\psi}'\mathcal{D}\psi'[\mathcal{D}A_\mu]\exp\left[\int \mathrm{d}^4x\,\mathcal{L}_{QED}\left(\bar{\psi}',\psi',A_\mu\right)\right] \end{aligned} \tag{A.18}$$

namely, we have the statement that the definite integral does not depend on the naming of path integral variables, which is a generalisation of the ordinary integral $\int \mathrm{d}x\,f(x) = \int \mathrm{d}y\,f(y)$. We look at parts which depend on fermionic variables in this identity. Firstly, the variation of the action in the exponential factor is given for an infinitesimal $\alpha(x)$ by

$$\begin{aligned} &\int \mathrm{d}^4x\left[\bar{\psi}(x)e^{i\alpha(x)\gamma^5}i\gamma^\mu(\partial_\mu - ieA_\mu)e^{i\alpha(x)\gamma^5}\psi(x) - m\bar{\psi}(x)e^{i\alpha(x)\gamma^5}e^{i\alpha(x)\gamma^5}\psi(x)\right] \\ &= \int \mathrm{d}^4x\left[\bar{\psi}(x)i\gamma^\mu(\partial_\mu - ieA_\mu)\psi(x) - m\bar{\psi}(x)\psi(x)\right] \\ &\quad + \int \mathrm{d}^4x\left[-\partial_\mu\alpha(x)\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x) - 2mi\alpha(x)\bar{\psi}(x)\gamma^5\psi(x)\right] \end{aligned} \tag{A.19}$$

where we used $\gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0$ and Eq. (A.17). If one further assumes that the path integral measure does not change under the above change of variables, one has

$$\mathcal{D}\bar{\psi}\mathcal{D}\psi[\mathcal{D}A_\mu] = \mathcal{D}\bar{\psi}'\mathcal{D}\psi'[\mathcal{D}A_\mu] \tag{A.20}$$

The above identity is written after those calculations as

$$\begin{aligned} &\int \mathcal{D}\bar{\psi}\mathcal{D}\psi[\mathcal{D}A_\mu]\left\{\int \mathrm{d}^4x\left[-\partial_\mu\alpha(x)\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x) - 2mi\alpha(x)\bar{\psi}(x)\gamma^5\psi(x)\right]\right\} \\ &\qquad \times \exp\left(\int \mathrm{d}^4x\,\mathcal{L}_{QED}\right) \\ &\equiv \int \mathrm{d}^4x\left\langle\left[-\partial_\mu\alpha(x)\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x) - 2mi\alpha(x)\bar{\psi}(x)\gamma^5\psi(x)\right]\right\rangle = 0 \end{aligned} \tag{A.21}$$

when one retains only the terms linear in $\alpha(x)$ by expanding the action in the exponential factor in powers of $\alpha(x)$. If one considers $\alpha(x)$ which has a δ-functional peak in the neighborhood of x, one obtains after partial integration the“naive” chiral identity

$$\partial_\mu\left\langle\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x)\right\rangle = 2im\left\langle\bar{\psi}(x)\gamma^5\psi(x)\right\rangle \tag{A.22}$$

As seen in Equations (A.2), (A.3) and (A.4), the current $\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x)$ which contains γ^5 is called the axial-vector current. In the case of the vanishing fermion mass $m = 0$, the axial-vector current is conserved and such a theory is called *chiral invariant.*

However, within the framework of perturbation theory, that the axial-vector vertex has anomalous properties which disagree with those found by the formal manipulation of field equations. In particular, because of the presence of closed-loop “triangle diagrams” (Fig. 12), the divergence of the axial-vector current is not the usual expression calculated from the field equations, and the axial-vector current does not satisfy the usual Ward identity: because the integral defining the triangle graph is linearly divergent, the value of the triangle graph is ambiguous and depends on the labelling convention and the method of evaluation of the integral. A consequence is that, in massless electrodynamics, despite the fact that the theory is invariant under γ^5 transformations, the axial-vector current is not conserved [3].

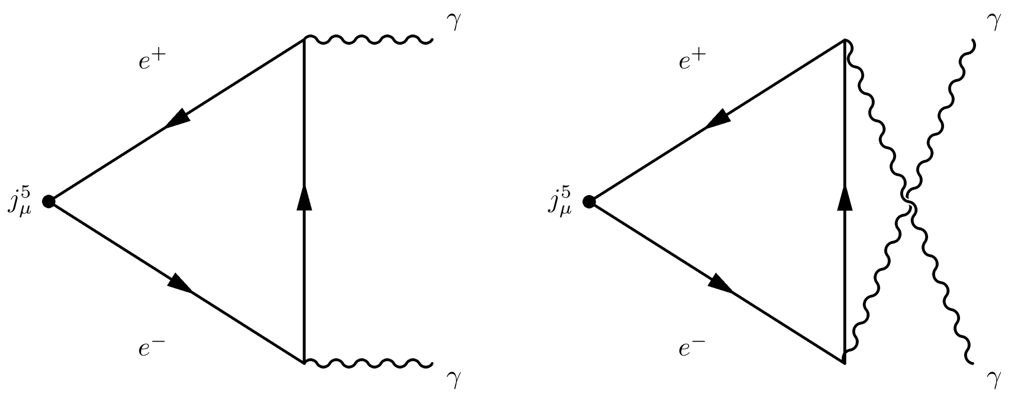

Figure 12: Feynman diagrams which give rise to the triangle anomaly

The actual perturbative calculation is very complicated and one has to deal with subtle momentum integrals which include a linear divergence. However, if one uses the covariant regularization [14], the calculation becomes simpler, and one can show the deviation from the naive chiral identity without relying on perturbative calculations.

The current is rewritten as

$$\begin{aligned}\left\langle\bar{\psi}(x)\gamma^{\mu}\gamma^{5}\psi(x)\right\rangle &= \lim_{y\to x}\left\langle T\bar{\psi}(y)\gamma^{\mu}\gamma^{5}\psi(x)\right\rangle \\ &= -\lim_{y\to x}\left\langle T\gamma^{\mu}_{\alpha\beta}\gamma^{5}\psi_{\beta}(x)\bar{\psi}_{\alpha}(y)\right\rangle \\ &= \lim_{y\to x}\operatorname{tr}\left[\gamma^{\mu}\gamma^{5}\frac{1}{i\not{D}-m}\delta(x-y)\right] \end{aligned} \tag{A.23}$$

We expand the denominator, where $\not{D}=\gamma^{\mu}(\partial_{\mu}-ieA_{\mu})$ in powers of eA_{μ}. In the present case, the term linear in eA_{μ} vanishes due to charge conjugation properties and the quadratic term in eA_{μ} gives the lowest-order term. In terms of the language of Feynman diagrams, one evaluates the triangular diagrams in Fig. 12 where the current appearing on the left-hand side of Eq. (A.22) is denoted by j^5_{μ}. For this reason, the chiral anomaly is also called the “triangle anomaly”. The degree of divergence d of this triangle diagram is $d=1$ and the diagram diverges linearly.

We apply the covariant regularization to this calculation

$$\left\langle\bar{\psi}(x)\gamma^{\mu}\gamma^{5}\psi(x)\right\rangle_{cov} \equiv \lim_{y\to x}\operatorname{tr}\left[\gamma^{\mu}\gamma^{5}\frac{1}{i\not{D}-m}f\left(\frac{\not{D}^{2}}{\Lambda^{2}}\right)\delta(x-y)\right] \tag{A.24}$$

with a smooth regulator function $f(x)$ ($f(0)=1$, $f(\infty)=0$, $\lim_{x\to 0}xf'(x)=0$ and $\lim_{x\to\infty}xf'(x)=0$) and the trace is taken over the freedom of Dirac matrices. A direct evaluation of this quantity is possible, but what we want to know is the deviation from Eq. (A.22) and thus it is sufficient to evaluate the derivative of this regularised current

$$\partial_{\mu}\left\langle\bar{\psi}(x)\gamma^{\mu}\gamma^{5}\psi(x)\right\rangle_{cov} \tag{A.25}$$

This calculation is performed in the following manner. We first define a complete set of eigenfunctions of the hermitian operator $\not{D}$

$$\begin{aligned}\not{D}\phi_n(x) &= \lambda_n\phi_n(x) \\ \int \mathrm{d}^4x\,\phi^{\dagger}_m(x)\phi_n(x) &= \delta_{m,n} \\ \sum_n \phi_n(x)\phi^{\dagger}_n(y) &= \delta(x-y) \end{aligned} \tag{A.26}$$

and rewrite Eq. (A.24) as

$$\begin{aligned}
\left\langle\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x)\right\rangle_{cov} &= \lim_{y\to x} \operatorname{tr}\left[\gamma^\mu\gamma^5\frac{1}{i\not{D}-m}f\left(\frac{\not{D}^2}{\Lambda^2}\right)\sum_n \phi_n(x)\phi_n^\dagger(y)\right] \\
&= \lim_{y\to x}\sum_n \phi_n^\dagger(y)\gamma^\mu\gamma^5\frac{1}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) \\
&= \sum_n \phi_n^\dagger(x)\gamma^\mu\gamma^5\frac{1}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) \qquad \text{(A.27)}
\end{aligned}$$

The summation here converges if the function $f\left(\lambda_n^2/\Lambda^2\right)$ rapidly approaches 0 for large λ_n^2/Λ^2. We now take the derivative of this expression

$$\begin{aligned}
\partial_\mu\left\langle\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x)\right\rangle_{cov} &= \partial_\mu\sum_n \phi_n^\dagger(x)\gamma^\mu\gamma^5\frac{1}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) \\
&= \sum_n\Bigg[\partial_\mu\phi_n^\dagger(x)\gamma^\mu\gamma^5\frac{1}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) \\
&\qquad + \phi_n^\dagger(x)\gamma^\mu\gamma^5\frac{1}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\partial_\mu\phi_n(x)\Bigg] \\
&= -\sum_n\left(\not{D}\phi\right)_n^\dagger(x)\gamma^5\frac{1}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) \\
&\qquad -\sum_n \phi_n^\dagger(x)\gamma^5\frac{1}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\not{D}\phi_n(x) \\
&= -\sum_n \phi_n^\dagger(x)\gamma^5\frac{2\lambda_n}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) \qquad \text{(A.28)}
\end{aligned}$$

where we used the properties of the γ matrices. This formula is further rewritten as

$$\begin{aligned}
&\partial_\mu\left\langle\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x)\right\rangle_{cov} \\
&= 2im\sum_n \phi_n^\dagger(x)\gamma^5\frac{1}{i\lambda_n-m}f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) + 2i\sum_n \phi_n^\dagger(x)\gamma^5 f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) \\
&= 2im\left\langle\bar{\psi}(x)\gamma^5\psi(x)\right\rangle_{cov} + 2i\sum_n \phi_n^\dagger(x)\gamma^5 f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x)
\end{aligned} \qquad \text{(A.29)}$$

This final expression shows that the gauge covariant regularization gave rise to an extra last term which can break chiral symmetry, in contrast to the naive chiral identity (A.22). The chiral identity is modified from a naive form by quantum effects if the above extra term does not vanish.

This last term is evaluated as follows: We first assume that the operator $f\left(\not{D}^2/\Lambda^2\right)$ is sufficiently convergent, and rewrite the sum over the four-component functions $\phi_n(x)$ to an integral over a complete set of plane waves

$$\begin{aligned}
\sum_n \phi_n^\dagger(x)\gamma^5 f\left(\frac{\lambda_n^2}{\Lambda^2}\right)\phi_n(x) &= \sum_n \phi_n^\dagger(x)\gamma^5 f\left(\frac{\not{D}^2}{\Lambda^2}\right)\phi_n(x) \\
&= \operatorname{tr}\int\frac{\mathrm{d}^4k}{(2\pi)^4}e^{-ikx}\gamma^5 f\left(\frac{\not{D}^2}{\Lambda^2}\right)e^{ikx} \qquad \text{(A.30)}
\end{aligned}$$

where the trace is over the Dirac indices which take four values. This calculation is performed as

$$
\begin{aligned}
&\mathrm{tr}\int \frac{\mathrm{d}^4k}{(2\pi)^4} e^{-ikx}\gamma^5 f\left(\frac{\not{D}^2}{\Lambda^2}\right)e^{ikx} \\
&= \mathrm{tr}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\gamma^5 f\left(\frac{(ik_\mu + D_\mu)(ik^\mu + D^\mu) - \frac{ie}{4}[\gamma^\mu,\gamma^\nu]F_{\mu\nu}}{\Lambda^2}\right) \qquad (A.31)\\
&= \Lambda^4\, \mathrm{tr}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\gamma^5 f\left(-k_\mu k^\mu + \frac{2ik^\mu D_\mu}{\Lambda} + \frac{D^\mu D_\mu}{\Lambda^2} - \frac{ie}{4\Lambda^2}[\gamma^\mu,\gamma^\nu]F_{\mu\nu}\right)
\end{aligned}
$$

where we used the following relation

$$
\begin{aligned}
\not{D}^2 &= \frac{1}{2}\{\gamma^\mu,\gamma^\nu\} D_\mu D_\nu + \frac{1}{2}[\gamma^\mu,\gamma^\nu]D_\mu D_\nu \\
&= D_\mu D^\mu - \frac{ie}{4}[\gamma^\mu,\gamma^\nu]F_{\mu\nu} \qquad (A.32)
\end{aligned}
$$

where we then moved the factor e^{ikx} through the operator $f\left(\not{D}^2/\Lambda^2\right)$ and performed a scaling of the variable $k_\mu \to \Lambda k_\mu$. We next expand the contents of the function $f(x)$ around the value $x = -k_\mu k^\mu = \left|k^2\right|$ (note that $x \geq 0$ in our Euclidean metric convention)

$$
\begin{aligned}
&f\left(-k_\mu k^\mu + \frac{2ik^\mu D_\mu}{\Lambda} + \frac{D^\mu D_\mu}{\Lambda^2} - \frac{ie}{4\Lambda^2}[\gamma^\mu,\gamma^\nu]F_{\mu\nu}\right) \\
&= f(-k_\mu k^\mu) + f'(-k_\mu k^\mu)\left(\frac{2ik^\mu D_\mu}{\Lambda} + \frac{D^\mu D_\mu}{\Lambda^2} - \frac{ie}{4\Lambda^2}[\gamma^\mu,\gamma^\nu]F_{\mu\nu}\right) \qquad (A.33)\\
&+ \frac{1}{2!}f''(-k_\mu k^\mu)\left(\frac{2ik^\mu D_\mu}{\Lambda} + \frac{D^\mu D_\mu}{\Lambda^2} - \frac{ie}{4\Lambda^2}[\gamma^\mu,\gamma^\nu]F_{\mu\nu}\right)^2 + \ldots
\end{aligned}
$$

and use the fact that only the terms of order $1/\Lambda^4$ or larger survive in the limit $\Lambda \to \infty$ in Eq. (A.31). We also note that the trace with the factor γ^5 is non-vanishing only when the trace contains four or more γ-matrices. The term which contains $([\gamma^\mu,\gamma^\nu]F_{\mu\nu})^2$ in Eq. (A.33) is the only one satisfying these non-vanishing conditions.

So, we have

$$
\begin{aligned}
&\mathrm{tr}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\gamma^5 \frac{1}{2!} f''(-k_\mu k^\mu)\left(-\frac{ie}{4}[\gamma^\mu,\gamma^\nu]F_{\mu\nu}\right)^2 \\
&= \mathrm{tr}\,\gamma^5 \frac{1}{2}\left(-\frac{ie}{4}[\gamma^\mu,\gamma^\nu]F_{\mu\nu}\right)^2 \frac{1}{16\pi^2}\int_0^\infty \mathrm{d}x\, x f''(x) \qquad (A.34)\\
&= \left(\frac{e^2}{32\pi^2}\right)\epsilon^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta}
\end{aligned}
$$

where we used

$$
\mathrm{tr}\,\gamma^5 \frac{1}{16}[\gamma^\mu,\gamma^\nu]\left[\gamma^\alpha,\gamma^\beta\right] = -\epsilon^{\mu\nu\alpha\beta} \qquad (A.35)
$$

and also $\mathrm{d}^4k = \pi^2\,\mathrm{d}\left|k^2\right|\left|k^2\right| = \pi^2\,\mathrm{d}x\, x$. Furthermore, we performed the integral

$$
\begin{aligned}
\int_0^\infty \mathrm{d}x\, x f''(x) &= x f'(x)\Big|_0^\infty - \int_0^\infty \mathrm{d}x\, f'(x) \\
&= -\, f(x)\Big|_0^\infty \\
&= f(0) = 1 \qquad (A.36)
\end{aligned}
$$

having in mind the properties of the regulator function (See comments after Eq. (A.24)).

We thus derived the chiral identity with the anomalous term (the last term)

$$\partial_\mu \left\langle \bar{\psi}(x)\gamma^\mu\gamma^5\psi(x)\right\rangle_{cov} = 2im\left\langle \bar{\psi}(x)\gamma^5\psi(x)\right\rangle_{cov} + 2i\left(\frac{e^2}{32\pi^2}\right)\epsilon^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta} \tag{A.37}$$

Equation (A.37), which is the principal result of this section, states the surprising fact that the axial-vector current divergence, as calculated in perturbation theory, contains a well defined extra term which is not obtained when the axial-vector divergence is calculated by formal use of the equations of motion. It can be shown [3] that the axial-vector divergence with the extra term included, is not multiplicatively renormalizable. It can also be shown that in perturbation theory is impossible to maintain gauge invariance and conservation of axial current at the same time. One may argue that there is no a priori reason to demand gauge invariance with respect to the photon indices as opposed to a normal axial-vector Ward identity, or, for that matter, to demand either. However, there are strong physical restrictions on the matrix element for an axial-vector meson to decay into two photons, and on the behaviour of the triangle loop diagrams at low energy, that force us to choose a value for the regulator function that preserves gauge invariance. So that it is not possible to get rid of the anomalies here discussed.

A.3 Connection with the PCAC puzzle

Despite the effective coupling constant for $\pi^0 \to \gamma\gamma$ should vanish for zero pion mass in theories with *Partially Conserved Axial Current* (PCAC) and gauge invariance, it does not so vanish in an explicit perturbation calculation in the Gell-Mann & Lévy σ-model [5].

The invariant amplitude for $\pi^0 \to \gamma\gamma$ is obtained by contracting the polarisation vectors of the photons with a tensor

$$T^{\mu\nu}(p,q) = \epsilon^{\mu\nu\alpha\gamma}p_\alpha q_\beta T\big(k^2\big) \tag{A.38}$$

where p and q are the photon momenta and k is the pion momentum. The above general form of $T^{\mu\nu}$ is dictated by Lorentz invariance, parity conservation, gauge invariance and Bose symmetry. Now $T\big(k^2\big)$, when calculated in perturbation theory from diagrams similar to those in Fig. 12, with j^5_μ replaced by the pion-nucleon coupling $ig_0\gamma^5$, gives

$$T(0) = g4\pi^2m^{-1} \tag{A.39}$$

where m is the nucleon mass.

On the other hand, when the neutral pion field is the divergence of the PCAC, as calculated by formal use of the equations of motion, then

$$T(0) = 0 \tag{A.40}$$

Reconciling equations (A.39) and (A.40) is the PCAC puzzle.

In [5] the authors attempt to circumvent this contradiction by introducing a regulator nucleon field ψ_1 which is quantised with commutators rather than anti-commutators. The coupling of the regulator field to the mesons is such that as the regulator mass approaches infinity, the regulator coupling to the mesons becomes infinite as well. As a consequence, even in the limit of infinite regulator mass the regulator field triangle diagram makes a contribution to the amplitude for $\pi^0 \to \gamma\gamma$ so that the total amplitude does vanish at $(p+q)^2$ in accord with the PCAC prediction.

However, it has been noted in [3] that the regulator procedure described above leads to grave difficulties when we turn to purely strong interaction phenomena. In fact, it introduces unrenormalizable infinities into the strong interactions in the σ- model, and therefore is not satisfactory.

A.4 Chiral anomaly in QCD-type theory

To derive the chiral anomaly in this case we will use a different method which, as opposed to the non-vanishing contribution of a triangle diagram, reinterprets it as a change in the partition function measure under a chiral transformation. The quantum anomaly is identified as the Jacobian arising from the symmetry transformation of path integral variables [14].

We start with an explanation of a straightforward generalisation of quantum electrodynamics, which is called Abelian gauge theory, to a non-Abelian gauge theory which has the same structure as QCD (quantum chromodynamics). The Euclidean path integral for this theory is given by

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi[\mathcal{D}A_\mu] \exp\left[\int \mathrm{d}^4x\, \bar{\psi}(i\not{D} - m)\psi + S_{YM}\right] \tag{A.41}$$

Dirac's γ-matrix convention is the same as for QED and γ^μ is anti-hermitian[19]. The covariant derivative D_μ is defined by

$$\not{D} \equiv \gamma^\mu D_\mu = \gamma^\mu \left(\partial_\mu - \sum_a igA^a_\mu T^a\right) \equiv \gamma^\mu(\partial_\mu - igA_\mu) \tag{A.42}$$

by using the generator T^a of a non-Abelian gauge group. The non-Abelian gauge field carries the same number of components as the generators of the group. We often use the notation $A_\mu = A^a_\mu T^a$ as in the last expression above. The field strength tensor $F^a_{\mu\nu}$ of the gauge field, which is a generalisation of the electric and magnetic fields $\vec{\boldsymbol{E}}$ and $\vec{\boldsymbol{B}}$ is defined by

$$[D_\mu, D_\nu] = -ig\big(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu\big)T^a \equiv -igF_{\mu\nu} \tag{A.43}$$

S_{YM} is an action for the non-Abelian gauge field, which is called the *Yang-Mills* field,

$$\begin{aligned} S_{YM} &= -\frac{1}{4}\sum_{\mu\nu\alpha}\int \mathrm{d}^4x\, F^a_{\mu\nu}F^{a\mu\nu} \\ &= -\frac{1}{4}\int \mathrm{d}^4x\,\big(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu\big)^2 \end{aligned} \tag{A.44}$$

and it gives a generalisation of Maxwell's action of the electromagnetic field. The path integral measure for the gauge field $[\mathcal{D}A_\mu]$ in Eq. (A.41) contains a suitable gauge fixing term also, and the details of the gauge fixing are not important for the analysis in this chapter[20].

The actual QCD, which describes the strong interaction, is based on the gauge group $SU(3)$, and thus we deal with a three-component field consisting of three conventional Dirac fields. In the case of the gauge group $SU(2)$, for example, an arbitrary function which belongs to $SU(2)$ is written as

$$g(x) = \exp\left[i\sum_{a=1}^{3}\omega^a(x)T^a\right] \in SU(2) \tag{A.45}$$

by using three real functions $\omega^a(x)$ and the generators T^a of $SU(2)$, which are given in terms of the Pauli matrices. The non-Abelian (local) gauge transformation is then defined by the replacement of variables

$$\begin{aligned} \psi(x) &\;\to\; \psi'(x) = g(x)\psi(x) \\ \bar{\psi}(x) &\;\to\; \bar{\psi}'(x) = \bar{\psi}(x)g^\dagger(x) \\ A_\mu(x) \equiv \sum_a A^a_\mu(x)T^a &\;\to\; A'_\mu(x) = g(x)A_\mu(x)g^\dagger(x) + \frac{1}{ig}[\partial_\mu g(x)]g^\dagger(x) \end{aligned} \tag{A.46}$$

The gauge transformation for matter fields thus corresponds to a rotation in internal space specified by a local $g(x)$ at each space-time point; for the gauge field A_μ, the transformation is given by a combination of rotation (the first term) and translation (the second term). The covariant derivative D_μ is then transformed under the gauge transformation as

$$D'_\mu = \partial_\mu - igA'_\mu(x) = g(x)[\partial_\mu - igA_\mu(x)]g^\dagger(x) = g(x)D_\mu g^\dagger(x) \tag{A.47}$$

and consequently

$$D'_\mu\psi'(x) = \big[\partial_\mu - igA'_\mu(x)\big]\psi'(x) = g(x)[\partial_\mu - igA_\mu(x)]\psi(x) = g(x)D_\mu\psi(x) \tag{A.48}$$

is transformed in the same manner as the field variable $\psi(x)$ itself. For this reason, D_μ is called the covariant derivative. The field strength tensor of the gauge field, which is expressed in terms of the covariant derivative, is transformed as

[19] See [24] for the properties of γ^μ and γ^5

[20] For an explanation of non-Abelian gauge groups see [27]

$$\left[D'_\mu, D'_\nu\right] = -igF'_{\mu\nu} = g(x)[D_\mu, D_\nu]g^\dagger(x) = -igg(x)F_{\mu\nu}g^\dagger(x) \tag{A.49}$$

If one recalls $g(x)g^\dagger(x) = 1$, with the convention $\mathrm{tr}\, T^aT^b = (1/2)\delta^{ab}$, the action appearing in Eq. (A.41) remains invariant under the gauge transformation, namely, the action is *gauge invariant*

$$\begin{aligned}\int \mathrm{d}^4x \left[\bar\psi(i\not{D} - m)\psi - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu}\right] &= \int \mathrm{d}^4x \left[\bar\psi(i\not{D} - m)\psi - \frac{1}{2}\,\mathrm{tr}\, F_{\mu\nu}F^{\mu\nu}\right] \\ &= \int \mathrm{d}^4x \left[\bar\psi'(i\not{D} - m)\psi' - \frac{1}{2}\,\mathrm{tr}\, F'_{\mu\nu}F'^{\mu\nu}\right] \end{aligned} \tag{A.50}$$

We examine the localised (space-time dependent) infinitesimal chiral transformations

$$\begin{aligned}\psi'(x) &= e^{i\alpha(x)\gamma^5}\psi(x) = \psi(x) + i\alpha(x)\gamma^5\psi(x) \\ \bar\psi'(x) &= \bar\psi(x)e^{i\alpha(x)\gamma^5} = \bar\psi(x) + \bar\psi(x)i\alpha(x)\gamma^5\end{aligned} \tag{A.51}$$

for the non-Abelian gauge theory. To analyse the Jacobian for the chiral transformations, we expand the fermionic variables

$$\begin{aligned}\psi(x) &= \sum_n a_n\varphi_n(x) \\ \bar\psi(x) &= \sum_n \bar b_n\varphi^\dagger_n(x)\end{aligned} \tag{A.52}$$

in terms of the eigenfunctions of the hermitian operator $\not{D}$ (see Eq. (A.26))

$$\begin{aligned}\not{D}\varphi_n(x) &= \lambda_n\varphi_n(x) \\ \int \mathrm{d}^4x\, \varphi^\dagger_m(x)\varphi_n(x) &= \delta_{m,n}\end{aligned} \tag{A.53}$$

The action for the fermion is formally diagonalized by this expansion

$$\int \mathrm{d}^4x\, \bar\psi(i\not{D} - m)\psi = \lim_{N\to\infty}\sum_{n=1}^{N}(i\lambda_n - m)\bar b_n a_n \tag{A.54}$$

The path integral measure for this case is written as

$$\mathcal{D}\bar\psi\mathcal{D}\psi = \lim_{N\to\infty}\prod_{n=1}^{N}\bar b_n a_n \tag{A.55}$$

and it leads to the following Jacobian for an infinitesimal chiral transformation [21]

$$J = \exp\left[-2i\lim_{N\to\infty}\sum_{n=1}^{N}\int \mathrm{d}^4x\, \alpha(x)\varphi^\dagger_n(x)\gamma^5\varphi_n(x)\right] \tag{A.56}$$

The actual evaluation of this Jacobian proceeds by replacing the mode cut-off by the eigenvalue cut-off as

$$\begin{aligned}&\lim_{M\to\infty}\lim_{N\to\infty}\sum_{n=1}^{N}\int \mathrm{d}^4x\, \alpha(x)\varphi^\dagger_n(x)\gamma^5 f\left(\frac{\lambda^2}{M^2}\right)\varphi_n(x) \\ &\lim_{M\to\infty}\sum_{n=1}^{\infty}\int \mathrm{d}^4x\, \alpha(x)\varphi^\dagger_n(x)\gamma^5 f\left(\frac{\not{D}^2}{M^2}\right)\varphi_n(x) \\ &\equiv \lim_{M\to\infty}\mathrm{tr}\,\alpha(x)\gamma^5 f\left(\frac{\not{D}^2}{M^2}\right)\end{aligned} \tag{A.57}$$

[21] see [14] for details

where $f(x)$ is an arbitrary function of x which approaches 0 rapidly as $x \to \infty$ with a normalisation $f(0) = 1$. After some manipulation, which is the generalisation to non-abelian case of that in section A.2 we finally get for the Jacobian of the chiral transformation

$$J = \exp\left[-2i \int \mathrm{d}^4x\, \alpha(x) \frac{g^2}{32\pi^2} \operatorname{tr} \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta}\right] \tag{A.58}$$

where the remaining trace is over the indices of matrix generators of the Yang-Mills field.

A.5 The Instanton

The Euclidean non-Abelian gauge theory, unlike abelian theory, accommodates a classical solution which is called an *instanton*. This solutions is the long range field $A_\mu(x)$ which minimise locally the Yang-Mills actions S_E and for which $S_E < \infty$. The instanton is fundamental since it describes tunnelling from one vacuum in field theory to another. In fact, in Euclidean theory we use an imaginary time to describe tunnelling in field theory just as in quantum mechanics. In particular, the classical solution of Euclidean theory gives the stationary point of the action and thus describes the path where the tunnelling probability becomes a maximum. The instanton solution precisely gives such a stationary point of the action.

All fields we are interested in satisfy the condition $F^a_{\mu\nu}(x) \to 0$ at space-time infinity $x \to \infty$. Consider a very large sphere, written as S^3, in our Euclidean 4-dimensional space $\left(\hat{x}^4, \vec{\mathbf{x}}\right)$. The sphere itself is of course 3-dimensional. Since the action S_E is invariant under the gauge transformation, the gauge field $A_\mu(x)$ itself should approach the configuration which is gauge equivalent to the vacuum $A_\mu(x) = 0$. This is equivalent to [4]

$$A_\mu(x)\Big|_{S^3} \approx ig(x)\partial_\mu g^\dagger(x)\Big|_{S^3} \tag{A.59}$$

where $g(x)$ is defined in Eq. (A.45). It can be shown that one can regard the element $g(x) \in SU(2)$ as describing a unit hypersurface (also written as S^3), so that the a point on the sphere in four-dimensional Euclidean space-time described above and a point on the hypersurface described by $g(x)$ are in $1:1$ correspondence. Namely, when the coordinates $\left(\hat{x}^4, \vec{\mathbf{x}}\right)$ of space-time cover S^3 once, the element $g(x)$ covers the hypersurface S^3 in the gauge space once. Hence every field $A_\mu(x)$ produces a certain mapping of the sphere S^3 onto the gauge group $SU(2)$. It is clear that if two such mappings belong to different homotopy classes then the corresponding fields $A^{(1)}_\mu$ and $A^{(2)}_\mu$ cannot be continuously deformed one into another. It is known that there exists an infinite number of different classes of mappings of $S^3 \to G$ if G is a non-abelian simple Lie group. Hence, the phase space of the Yang-Mills fields are divided into an infinite number of components, each of which is characterised by some value of N_{CS}, where N_{CS} is a certain integer, called the winding number. This is written by using the instanton solution as

$$N_{CS} = \frac{1}{32\pi^2} \operatorname{tr} \int \mathrm{d}^4x\, \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F^{\mu\nu} \tag{A.60}$$

To prove this, we use the Jacobi identity

$$\operatorname{tr} \epsilon^{\mu\nu\alpha\beta} [A_\mu, A_\nu][A_\alpha, A_\beta] = \operatorname{tr} \epsilon^{\mu\nu\alpha\beta} A_\mu [A_\nu, [A_\alpha, A_\beta]] = 0 \tag{A.61}$$

and the Gauss law together with the behavior of the instanton solution at infinity to write

$$\begin{aligned}
N_{CS} &= \frac{1}{8\pi^2} \operatorname{tr} \int \mathrm{d}^4x\, \partial_\mu \left[\epsilon^{\mu\nu\alpha\beta}\left(A_\nu \partial_\alpha A_\beta - \frac{2i}{3} A_\nu A_\alpha A_\beta\right)\right] \\
&= \frac{1}{24\pi^2} \operatorname{tr} \int \mathrm{d}S_\mu\, \epsilon^{\mu\nu\alpha\beta} A_\nu A_\alpha A_\beta \\
&= \frac{1}{24\pi^2} \operatorname{tr} \int \mathrm{d}S_\mu\, \epsilon^{\mu\nu\alpha\beta} \left(g\partial_\nu g^{-1}\right)\left(g\partial_\alpha g^{-1}\right)\left(g\partial_\beta g^{-1}\right)
\end{aligned} \tag{A.62}$$

Indeed, the integral in the last expression stands for a surface integral over the hypersurface S^3 located at infinity of four-dimensional space-time, and $\mathrm{d}S_\mu$ stands for the surface element which is orthogonal to the μ-axis. So, we see that the integrand in Eq. (A.62) is precisely the Jacobian of the mapping of S^3 on $SU(2)$. Hence N_{CS} is the number of times the $SU(2)$ is covered under this mapping. It is just the definition of the mapping degree[22].

Now, if we define $\tilde{F}_{\mu\nu} = (1/2)\epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}$ and use the Schwartz inequality

$$\begin{aligned} \int d^4x\, F^a_{\mu\nu} F^a_{\mu\nu} &= \sqrt{\int d^4x\, F^a_{\mu\nu} F^a_{\mu\nu} \int d^4x\, \tilde{F}^a_{\mu\nu} \tilde{F}^a_{\mu\nu}} \\ &\geq \left| \int d^4x\, F^a_{\mu\nu} \tilde{F}^a_{\mu\nu} \right| \end{aligned} \tag{A.63}$$

Using Eq. (A.60) we have

$$\frac{1}{4g^2} \int d^4x\, F^a_{\mu\nu} F^a_{\mu\nu} \geq \frac{8\pi^2}{g^2} |N_{CS}| \tag{A.64}$$

and the equality holds only for the case

$$F^a_{\mu\nu} = \pm \tilde{F}^a_{\mu\nu} \tag{A.65}$$

where $\pm$ corresponds to the signature of the winding number N_{CS}. If one uses the $+$ signature in this relation and defines $A_\mu(x) = i f(r^2) g(x) \partial_\mu g^\dagger(x)$, we have $f(r^2)^2 - f(r^2) = -r^2 f'(r^2)$ which can be shown to be satisfied by

$$f(r^2) = \frac{r^2}{r^2 + \rho^2} \tag{A.66}$$

where the parameter ρ is a real constant. This shows that the instanton gives the winding number $N_{CS} = 1$ and at the same time it gives the minimum of the action for $N_{CS} = 1$. The field configuration which gives the stationary point of the action is a solution of the field equation derived from the action, and thus the instanton is in fact the solution of the Euclidean Yang-Mills field equation.

We finally write down an expression for the instanton solution

$$A_\mu(x) = i \frac{r^2}{r^2 + \rho^2} g(x) \partial_\mu g^\dagger(x) \tag{A.67}$$

that approaches the configuration $i g(x) \partial_\mu g^\dagger(x)$ at infinity, which is gauge equivalent to the vacuum. Physically, this gauge function $g(x)$ is interpreted as describing tunnelling starting from one vacuum at $x^4 = -\infty$ to another vacuum at $x^4 = \infty$ following the Euclidean imaginary time. (For this reason, this solution is called an "instant-on", indicating that it appears and then disappears instantly in time unlike the ordinary soliton). The value of the action for this solution $S_E = -8\pi^2/g^2$ describes the height of the tunnelling barrier, and in the path integral the factor with this action in the exponential

$$e^{-\frac{8\pi^2}{g^2}} \tag{A.68}$$

is understood as giving the leading term of the tunnelling probability.

[22]The quantity appearing in Eq. (A.62), $K^\mu = \frac{1}{8\pi^2} \operatorname{tr} \int d^4x\, \epsilon^{\mu\nu\alpha\beta} \left(A_\nu \partial_\alpha A_\beta - \frac{2i}{3} A_\nu A_\alpha A_\beta \right)$, is the *Chern-Simons form*

B CP Violation

Any relativistic field theory must be invariant under continuous Lorentz transformations. In addition to continuous Lorentz transformations, there are two other space-time operations that are potential symmetries of a theory's Lagrangian: parity and time reversal. Parity, denoted by P, sends $(t, x) \to (t, -x)$, reversing the handedness of space. Time reversal, denoted by T, sends $(t, x) \to (-t, x)$, interchanging the forward and backward light-cones. It will be convenient to introduce a third (non-space-time) discrete operation: charge conjugation, denoted by C. Under this operation, particles and antiparticles are interchanged [24].

Before 1956 it was firmly believed that physical laws are symmetric with respect to P, T and C. Though none of these symmetries followed from any fundamental principle the belief was quite strong and it was a great shock when it was found that space parity is not conserved [20] so that a mirror reflected process could be physically impossible. It was assumed immediately that simultaneous mirror reflection and charge conjugation, CP, restore the symmetry so that for each process with particles the mirror reflected process with antiparticles, and otherwise the same, is possible. However in 1964 it was found that CP is also broken [8] and after this discovery a theory of non conservation of baryons became possible.

The only discrete symmetry which survived to the present day is the combination of all three transformations, CPT. The so called CPT-theorem can be proven which states that any Lorentz-invariant theory with positive definite energy and the normal relation between spin and statistics is invariant with respect to CPT-transformation. As a result of this symmetry masses of particles and antiparticles and their total decay widths must be exactly equal. However the probabilities of specific channels should be different for charged conjugated processes if both C and CP are broken.

Experimentally the first Sakharovs condition is well justified, CP-violation is directly observed in the decays of K_0-mesons, but it is not yet known what mechanism is responsible for it. A knowledge of that is very important for baryogenesis because baryogenesis took place at a much larger energy scale where the data on the kaon decays cannot be simply applied. Anyway it is known in principle that antiparticles are not just mirror reflections of particles, they have essentially different interactions and are produced with different probabilities in charge conjugated processes.

B.1 P, T and C

The existence of the operations of parity and time reversal is related to the connectedness of the Lorentz group itself [6]. In fact not all coordinate transformations permitted in special relativity can be built infinitesimally as a continuous Lorentz transformation starting from the identity. In particular, the two transformations of coordinates seen above cannot: the parity transformation

$$x^\mu \to P^\mu_\nu x^\nu, \quad P^\mu_\nu = \begin{pmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{B.1}$$

which reflects each space coordinate, and the time reversal transformation,

$$x^\mu \to T^\mu_\nu x^\nu, \quad T^\mu_\nu = \begin{pmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{pmatrix} \tag{B.2}$$

which reverses the sign of time. The representations of P and T in the theorys Hilbert space are denoted by $\mathcal{P}$ and $\mathcal{T}$ respectively. $\mathcal{P}$ can always be chosen to be a unitary operator and although $\mathcal{T}$ cannot be made unitary, it may always be chosen to be anti-unitary (that is, an operator which flips the sign of i)[23]. Parity and time reversal are potential symmetries of a theory's Lagrangian.

The unitary operator that represents this interchange in the Hilbert space will be denoted by $\mathcal{C}$.

[23]The reason $\mathcal{T}$ is anti-unitary is that H must transform under the symmetry into an operator which still has a positive spectrum; this will be satisfied if $\mathcal{P}H\mathcal{P}^* = H$ and $\mathcal{T}H\mathcal{T}^* = H$. On the other hand, time evolution by a positive amount of time t, e^{-iHt}, should be changed under time reversal to time evolution by a negative amount of time t, $\mathcal{T}e^{-iHt}\mathcal{T}^* = e^{iHt}$. The only way that both of these can be true is if T is an anti-unitary operator, reversing the sign of i.

B.2 Invariance of the SM to CPT transformation rules

The action of $\mathcal{P}$, $\mathcal{T}$, and $\mathcal{C}$ on particle states and on fields is determined (up to a conventionally fixed freedom to redefine fields) by their transformation properties under Lorentz transformations. Their action on a state, $|\mathbf{p}, \sigma\rangle$, that describes a particle of three-momentum $\mathbf{p}$, total spin j, and third component of angular momentum σ, may be chosen to be

$$\begin{aligned} \mathcal{P}\,|\mathbf{p},\sigma\rangle &= \alpha_p\,|-\mathbf{p},\sigma\rangle \\ \mathcal{T}\,|\mathbf{p},\sigma\rangle &= \alpha_t(-)^{j-\sigma}\,|-\mathbf{p},-\sigma\rangle \\ \mathcal{C}\,|\mathbf{p},\sigma\rangle &= \alpha_c\overline{|-\mathbf{p},\sigma\rangle} \end{aligned} \tag{B.3}$$

In these expressions, α_p, α_t and α_c are phases that are characteristic of each particle type, and the state $\overline{|\ldots\rangle}$ denotes the antiparticle for the state $|\ldots\rangle$. The transformation properties of the corresponding creation and annihilation operators are determined by those of the particle states

$$\begin{aligned} \mathcal{P}a^*_{\mathbf{p},\sigma}\mathcal{P}^* &= \alpha_p a^*_{-\mathbf{p},\sigma} \\ \mathcal{T}a^*_{\mathbf{p},\sigma}\mathcal{T}^* &= \alpha_t(-)^{j-\sigma}a^*_{-\mathbf{p},-\sigma} \\ \mathcal{C}a^*_{\mathbf{p},\sigma}\mathcal{C}^* &= \alpha_c\bar{a}^*_{\mathbf{p},\sigma} \end{aligned} \tag{B.4}$$

The transformation rules for the fields are then determined by their expansions in terms of creation and annihilation operators. Since these have the generic form

$$\phi \sim \sum_{\mathbf{p},\sigma}\left[u(\mathbf{p},\sigma)a_{\mathbf{p},\sigma} + v(\mathbf{p},\sigma)\bar{a}^*_{\mathbf{p},\sigma}\right] \tag{B.5}$$

the transformation rules for fields representing spin-zero particles become

$$\begin{aligned} \mathcal{P}\phi(x)\mathcal{P}^* &= \alpha_p^*\phi(x_p) \\ \mathcal{C}\phi(x)\mathcal{C}^* &= \alpha_c^*\phi(x) \end{aligned} \tag{B.6}$$

in which $x_p = (-\mathbf{x}, t)$ denotes the image of $x = (\mathbf{x}, t)$ under parity. (Since invariance of the theory under the combination CPT is guaranteed on general grounds, T-invariance is equivalent to CP-invariance. For this reason it suffices to have explicit expressions for the transformation rules under C and P in order to determine its symmetry properties.)

For (Majorana) spinor fields we have instead,

$$\begin{aligned} \mathcal{P}\psi(x)\mathcal{P}^* &= \alpha_p^*\beta\psi(x_p) \\ \mathcal{C}\psi(x)\mathcal{C}^* &= \alpha_c^*C\bar{\psi}^T(x) \end{aligned} \tag{B.7}$$

in which

$$\beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} = i\gamma^0 \tag{B.8}$$

and C is the charge conjugation matrix [24]. The factor β exchanges left- and right-handed components and is necessary because parity flips handedness. Finally, for spin-one gauge potentials, V_a^μ, that correspond to the gauge generator, t_a, we have (up to gauge transformations)

$$\begin{aligned} \mathcal{P}[t_a V_a^\mu(x)]\mathcal{P}^* &= P^\mu_{\ \nu}[t_a V_a^\mu(x_p)] \\ \mathcal{C}[t_a V_a^\mu(x)]\mathcal{C}^* &= -[t_a V_a^\mu(x)]^* \end{aligned} \tag{B.9}$$

The phase in the transformation rule for the gauge potentials is fixed by the requirement that the covariant derivative, $D = \partial - iT_aV_a$, transforms properly. Using these transformation rules, we can test the standard model interactions of Appendix C for invariance under the three independent symmetries of C, P, and CP.

The typical interaction Lagrangian density is the sum of several local operators, $\mathcal{O}_n(x)$, with some constant coefficients, $c_n : \mathcal{L}_{int} = \sum_n c_n \mathcal{O}_n(x)$. The resulting transformation rule for the interaction Lagrangian is

$$\begin{aligned} \mathcal{P}\mathcal{L}_{int}\mathcal{P}^* &= \sum_n (\alpha_n)_p c_n \mathcal{O}_n(x_p) \\ \mathcal{C}\mathcal{L}_{int}\mathcal{C}^* &= \sum_n (\alpha_n)_c c_n \mathcal{O}_n^*(x) \\ (\mathcal{C}\mathcal{P})\mathcal{L}_{int}(\mathcal{C}\mathcal{P})^* &= \sum_n (\alpha_n)_p (\alpha_n)_c c_n \mathcal{O}_n^*(x_p) \end{aligned} \tag{B.10}$$

where the phases $(\alpha_n)_p$ and $(\alpha_n)_c$ are products of the phases associated with the transformation of each field. Since the action is given by the integral of $\mathcal{L}(x)$ over space-time, the condition $\mathcal{P}\mathcal{L}(x)\mathcal{P}^* = \mathcal{L}(x_p)$ suffices to ensure that the action is invariant. The condition for parity invariance is therefore that there exist a choice of phases for each of the fields for which

$$(\alpha_n)_p = 1 \quad \text{for all}\, n \tag{B.11}$$

The Lagrangian is also required by unitarity to be Hermitian, so the following relation among the operators is also true

$$\sum_n c_n^* \mathcal{O}_n^* = \sum_n c_n \mathcal{O}_n \tag{B.12}$$

The action is therefore charge-conjugation invariant provided that there exists a choice of charge-conjugation phases, for each of the fields for which the coefficient of $\mathcal{O}_n^*$ is unchanged:

$$(\alpha_n)_c c_n = c_n^* \quad \text{for all}\, n \tag{B.13}$$

CP-invariance is similarly ensured if phases can be chosen such that

$$(\alpha_n)_c (\alpha_n)_p c_n = c_n^* \quad \text{for all}\, n \tag{B.14}$$

If we apply this formalism to the standard model Lagrangian then we find[6] that 7he Higgs interactions, gluon interactions, and electromagnetic interactions all respect each of the three discrete symmetries, C, P, and CP. The neutral current couplings of the fermions to the neutral Z boson break both C and P but in such a way that the combination CP is unbroken. Finally, the charged-current coupling of the fermions to the W boson not only violates C and P, but can also violate CP, provided that there is not sufficient freedom to make the Kobayashi-Maskawa matrix real.

As an illustration we show the manipulations for the charged-current quark interactions, Eq. (C.9). In this case the transformation rules for the spin-one fields become $\mathcal{C}W_\mu^\pm \mathcal{C}^* = -W_\mu^\mp$ and $\mathcal{P}W_\mu^\pm \mathcal{P}^* = P_\mu^\nu W_\nu^\pm$. Then, under charge conjugation, we have

$$\begin{aligned} \mathcal{C}\mathcal{L}_{cc}\mathcal{C}^* = \frac{ig_2}{2\sqrt{2}} \Big\{ &(\alpha_{u_m})_c (\alpha_{d_n})_c^* V_{mn} W_\mu^- \left[\bar{d}_n \gamma^\mu (1-\gamma^5) u_m\right]^* \\ &+ (\alpha_{u_n})_c^* (\alpha_{d_m})_c (V^\dagger)_{mn} W_\mu^+ \left[\bar{u}_n \gamma^\mu (1-\gamma^5) d_m\right]^* \Big\} \end{aligned} \tag{B.15}$$

and under parity transformations we get

$$\begin{aligned} \mathcal{P}\mathcal{L}_{cc}\mathcal{P}^* = \frac{ig_2}{2\sqrt{2}} \Big[&(\alpha_{u_m})_p (\alpha_{d_n})_p^* V_{mn} W_\mu^+ \bar{u}_m \gamma^\mu (1-\gamma^5) d_n \\ &+ (\alpha_{d_m})_p^* (\alpha_{u_n})_p (V^\dagger)_{mn} W_\mu^- \bar{d}_m \gamma^\mu (1-\gamma^5) u_n \Big] \end{aligned} \tag{B.16}$$

It is clear that there is no choice of phases for which the Lagrangian is parity or charge-conjugation invariant, because any choice that would make the term involving γ^μ invariant would make the $\gamma^5\gamma^\mu$ term not invariant (and vice versa). The point is that each operation replaces the projector $P_L = \left(1+\gamma^5\right)/2$ with the projector $P_R = \left(1-\gamma^5\right)/2$. Combining both transformations, however, gives the following result:

$$\begin{aligned}\mathcal{CP}\mathcal{L}_{cc}\mathcal{CP}^* = & \frac{ig_2}{2\sqrt{2}} \\ & \times \Big\{(\alpha_{u_m})_c(\alpha_{d_n})_c^*(\alpha_{u_m})_p(\alpha_{d_n})_p^* V_{mn} W_\mu^- \left[\bar{d}_n\gamma^\mu\left(1+\gamma^5\right)u_m\right]^* \\ & +(\alpha_{u_n})_c^*(\alpha_{d_m})_c(\alpha_{u_n})_p^*(\alpha_{d_m})_p \left(V^\dagger\right)_{mn} W_\mu^+ \left[\bar{u}_n\gamma^\mu\left(1+\gamma^5\right)d_m\right]^*\Big\}\end{aligned} \tag{B.17}$$

If the phases can be chosen to satisfy $(\alpha_{u_m})_c(\alpha_{d_n})_c^*(\alpha_{u_m})_p(\alpha_{d_n})_p^* = 1$, and the CKM matrix, Eq. (C.10), can be simultaneously chosen to be real, then this last equation would be precisely the complex conjugate of the original Lagrangian. Inspection of the other terms in the Lagrangian confirms that the phase choice can be made provided that the CKM matrix may be chosen to be real. Therefore, the standard model fails to conserve CP invariance only in that the CKM matrix *cannot* be made purely real.

C The Standard Model

The Standard Model of particle physics is a theory concerning the electromagnetic, weak, and strong nuclear interactions, as well as classifying all the subatomic particles known. The strong, weak, and electromagnetic interactions are understood as arising due to the exchange of various spin-one bosons amongst the spin-half particles that make up matter [6]. The gauged symmetry group of the standard model is $SU_c(3) \times SU_L(2) \times U_Y(1)$. We have:

1. Eight spin-one particles called *gluons*, $G^{\alpha}_{\mu}(x)$, associated with the factor $SU_c(3)$. The associated subscript "c" is meant to denote "color". Any particle that transforms with respect to this factor of the gauge group, and so which couples to the gluons, is said to be colored or to carry color. This interaction is also called the "strong interaction", and any particle which couples to the gluons is said to be "strongly interacting".

2. Three spin-one particles, $W^{\alpha}_{\mu}(x)$, associated with the factor $SU_L(2)$. The subscript "L" is meant to indicate that only the left-handed fermions turn out to carry this quantum number.

3. One particle $B_{\mu}(x)$, associated with the factor $U_Y(1)$. The subscript "Y" is meant to distinguish the group associated with the quantum number of *weak hypercharge*, denoted Y, from the group associated with ordinary electric charge, denoted Q.

The four spin-one bosons associated with the factors $SU_L(2)\times U_Y(1)$ are related to the physical bosons that mediate the weak interactions, $W^{\pm}$ and Z^0, and the familiar photon from Quantum Electrodynamics (QED).

Apart from spin-one particles we are aware of a number of fundamental spin-half particles. Our knowledge to date about the character of the interactions of these fermions may be compactly summarized by giving their transformation properties with respect to the gauge group $SU_c(3) \times SU_L(2) \times U_Y(1)$.

C.1 The SM Lagrangian

The SM Lagrangian takes the form

$$\mathcal{L}_{SM} = \mathcal{L}_{fg} + \mathcal{L}_{Higgs} \tag{C.1}$$

$$\begin{aligned}\mathcal{L}_{fg} = & -\frac{1}{4}G^{\alpha}_{\mu\nu}G^{\alpha\mu\nu} - \frac{1}{4}W^{a}_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{g_3^2\theta}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}G^{\alpha\mu\nu}G^{\alpha\lambda\rho} \\ & -\frac{g_2^2\Theta_2}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}W^{a\mu\nu}W^{a\lambda\rho} - \frac{g_1^2\Theta_1}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}B^{\mu\nu}B^{\lambda\rho} - \frac{1}{2}\bar{L}_m \not{D} L_m \\ & -\frac{1}{2}\bar{E}_m \not{D} E_m - \frac{1}{2}\bar{Q}_m \not{D} Q_m - \frac{1}{2}\bar{U}_m \not{D} U_m - \frac{1}{2}\bar{D}_m \not{D} D_m\end{aligned} \tag{C.2}$$

$$\begin{aligned}\mathcal{L}_{Higgs} = & -(D_{\mu}\phi)^{\dagger}(D^{\mu}\phi) - V\left(\phi^{\dagger}\phi\right) \\ & -\left(f_{mn}\bar{L}_m P_R E_n \phi + h_{mn}\bar{Q}_m P_R D_n \phi + g_{mn}\bar{Q}_m P_R U_n \tilde{\phi} + \text{h.c.}\right)\end{aligned} \tag{C.3}$$

$$\begin{aligned}V\left(\phi^{\dagger}\phi\right) &= \lambda\left[\phi^{\dagger}\phi - \frac{\mu^2}{2\lambda}\right]^2 \\ &= \lambda\left(\phi^{\dagger}\phi\right)^2 - \mu^2\phi^{\dagger}\phi + \frac{\mu^4}{4\lambda}\end{aligned} \tag{C.4}$$

in which the gauge field-strengths are given by

$$G^{\alpha}_{\mu\nu} = \partial_{\mu}G^{\alpha}_{\nu} - \partial_{\nu}G^{\alpha}_{\mu} + g_3 f^{\alpha}_{\beta\gamma}G^{\beta}_{\mu}G^{\gamma}_{\nu} \tag{C.5}$$

$$W^{a}_{\mu\nu} = \partial_{\mu}W^{a}_{\nu} - \partial_{\nu}W^{a}_{\mu} + g_2\epsilon_{abc}G^{b}_{\mu}G^{c}_{\nu} \tag{C.6}$$

$$B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu} \tag{C.7}$$

C.2 Charged current fermion interactions

The electroweak interactions come from the couplings that involve the electroweak gauge bosons, those spin-one particles that correspond to the $U_L(2) \times U_Y(1)$ factor of the gauge group. These come in two basic types. There are self-couplings that arise due to the non-linear terms in the gauge potentials within the $U_L(2) \times U_Y(1)$ field strenghts, and there are couplings with other particles that arise due to the use of gauge covariant derivatives in the kinetic-energy terms of the Lagrangian. In particular, consider the couplings between the electroweak bosons and spin-half and spin-zero particles.

The W^a_μ and B_μ-fermion couplings arise from the following kinetic terms of the Lagrangian (C.2),

$$\mathcal{L} = -\frac{1}{2}\bar{L}_m \not{D} L_m - \frac{1}{2}\bar{E}_m \not{D} E_m - \frac{1}{2}\bar{Q}_m \not{D} Q_m - \frac{1}{2}\bar{U}_m \not{D} U_m - \frac{1}{2}\bar{D}_m \not{D} D_m \tag{C.8}$$

The couplings between fermions and the charged spin-one particle, W^+_μ, are called the *charged-current interactions*

$$\mathcal{L}_{cc} = \frac{ig_2}{2\sqrt{2}}\left[V_{mn} W^+_\mu \bar{u}_m \gamma^\mu \left(1+\gamma^5\right) d_n + \left(V^\dagger\right)_{mn} W^-_\mu \bar{d}_m \gamma^\mu \left(1+\gamma^5\right) u_n\right] \tag{C.9}$$

Where we have defined the 3×3 unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix.

$$V_{mn} = \left(U^{(u)^\dagger} U^{(d)}\right)_{mn} \tag{C.10}$$

which arises due to the necessity to perform different field redefinitions for up- and down-type quarks when diagonalizing masses.

References

[1] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.

[2] P.A.R. Ade et al. Planck 2013 results. XVI. Cosmological parameters. *Astron.Astrophys.*, 571:A16, 2014.

[3] Stephen L. Adler. Axial-vector vertex in spinor electrodynamics. *Phys. Rev.*, 177:2426–2438, Jan 1969.

[4] A. A. Belavin, Alexander M. Polyakov, A. S. Schwartz, and Yu. S. Tyupkin. Pseudoparticle Solutions of the Yang-Mills Equations. *Phys. Lett.*, B59:85–87, 1975.

[5] J. S. Bell and R. Jackiw. A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ-model. *Nuovo Cim.*, A60:47–61, 1969.

[6] C. P. Burgess and G. D. Moore. *The standard model: A primer.* 2007.

[7] Hai-Yang Cheng. The Strong CP Problem Revisited. *Phys. Rept.*, 158:1, 1988.

[8] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2 pi Decay of the $k\,(2)\,0$ Meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[9] Andrew G. Cohen, D. B. Kaplan, and A. E. Nelson. Progress in electroweak baryogenesis. *Ann. Rev. Nucl. Part. Sci.*, 43:27–70, 1993.

[10] Nathaniel Craig and John March-Russell. Axion-Assisted Electroweak Baryogenesis. 2010.

[11] Michael Dine, Olaf Lechtenfeld, Bunji Sakita, Willy Fischler, and Joseph Polchinski. Baryon Number Violation at High Temperature in the Standard Model. *Nucl. Phys.*, B342:381–408, 1990.

[12] A. D. Dolgov. Baryogenesis, 30 years after. In *Surveys in high-energy physics. Proceedings, 25th ITEP Winter School, Snegiri, Russia, February, 1997*, 1997.

[13] Hugo Fort and R. Gambini. The U(1) and strong CP problem from the loop formulation perspective. 1997.

[14] K. Fujikawa and H. Suzuki. *Path integrals and quantum anomalies.* 2004.

[15] Kerson Huang. *Statistical Mechanics.* John Wiley & Sons, 2 edition, 1987.

[16] R. Jackiw and C. Rebbi. Vacuum Periodicity in a Yang-Mills Quantum Theory. *Phys. Rev. Lett.*, 37:172–175, 1976.

[17] Edward W. Kolb and Michael S. Turner, editors. *THE EARLY UNIVERSE. REPRINTS.* 1988.

[18] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov. On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe. *Phys. Lett.*, B155:36, 1985.

[19] V. A. Kuzmin, M. E. Shaposhnikov, and I. I. Tkachev. Strong CP violation, electroweak baryogenesis, and axionic dark matter. *Phys. Rev.*, D45:466–475, 1992.

[20] T. D. Lee and Chen-Ning Yang. Question of Parity Conservation in Weak Interactions. *Phys. Rev.*, 104:254–258, 1956.

[21] R. D. Peccei. The Strong CP problem and axions. *Lect. Notes Phys.*, 741:3–17, 2008. [,3(2006)].

[22] R. D. Peccei and Helen R. Quinn. Constraints Imposed by CP Conservation in the Presence of Instantons. *Phys. Rev.*, D16:1791–1797, 1977.

[23] R. D. Peccei and Helen R. Quinn. CP Conservation in the Presence of Instantons. *Phys. Rev. Lett.*, 38:1440–1443, 1977.

[24] Michael Edward Peskin and Daniel V. Schroeder. *An introduction to quantum field theory.* Advanced book program. Westview Press Reading (Mass.), Boulder (Colo.), 1995. Autre tirage : 1997.

[25] A. D. Sakharov. Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967. [Usp. Fiz. Nauk161,61(1991)].

[26] Geraldine Servant. Baryogenesis from Strong CP Violation and the QCD Axion. *Phys.Rev.Lett.*, 113(17):171803, 2014.

[27] Kellogg Stelle. Unification and The Standard Model - Lecture Notes. 2014.

[28] Gerard 't Hooft. Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle. *Phys. Rev.*, D14:3432–3450, 1976. [Erratum: Phys. Rev.D18,2199(1978)].

[29] Gerard 't Hooft. Symmetry Breaking Through Bell-Jackiw Anomalies. *Phys. Rev. Lett.*, 37:8–11, 1976.

[30] Gerard 't Hooft. How Instantons Solve the U(1) Problem. *Phys. Rept.*, 142:357–387, 1986.

[31] Gerard 't Hooft. The Physics of instantons in the pseudoscalar and vector meson mixing. 1999.

[32] Mark Trodden. Electroweak baryogenesis. *Rev. Mod. Phys.*, 71:1463–1500, 1999.

[33] Michael S. Turner. Windows on the Axion. *Phys. Rept.*, 197:67–97, 1990.

[34] Steven Weinberg. The U(1) Problem. *Phys. Rev.*, D11:3583–3593, 1975.

[35] Steven Weinberg. A New Light Boson? *Phys. Rev. Lett.*, 40:223–226, 1978.

[36] Frank Wilczek. Problem of Strong p and t Invariance in the Presence of Instantons. *Phys. Rev. Lett.*, 40:279–282, 1978.