ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor Philosophiæ

BEYOND THE STANDARD MODEL: TOPICS ON CHIRAL GAUGE ANOMALIES AND ON NEUTRINO OSCILLATIONS.

CANDIDATE:

Maurizio Piai

SUPERVISORS:

Marco Fabbrichesi

Antonio Masiero

Serguey T. Petcov

ACADEMIC YEAR 2001-2002

SISSA - SCUOLA INTERNAZIONALE SUPERIORE STUDI AVANZATI

TRIESTE

Via Beirut 2-4

TRIESTE

Thesis submitted in partial fulfillment of the requirements for the degree of **Doctor Philosophiæ**

BEYOND THE STANDARD MODEL: TOPICS ON CHIRAL GAUGE ANOMALIES AND ON NEUTRINO OSCILLATIONS.

CANDIDATE:

Maurizio Piai

SUPERVISORS:

Marco Fabbrichesi

Antonio Masiero

Serguey T. Petcov

ACADEMIC YEAR 2001-2002

Preface

Within the multi-folded research context of beyond the Standard Model physics, the author of this Ph. D. Thesis works on several different research lines: the study of low-energy Supersymmetry effects and of its underlying flavor and CP-violating structures at the B factories, the exploration of extra-dimension scenarios through gravitational effects on ordinary matter and bulk fields, the development of anomaly cancellation mechanisms for gauge theories in D dimensions, with the study of their phenomenological implications, the discussion about new experimental facilities for testing the structure of the neutrino mass matrix.

Not all of the material produced during the Ph. D. activity could be collected here, but the results in two of these fields have been selected: topics related to the cancellation of gauge anomalies and to the phenomenology of neutrino oscillations from artificial sources.

As a result, this Thesis has a quite unusual structure: it is divided into two separate parts, which share the effort of Particle Physics to go beyond the Standard Model, but differ by the use of drastically different research techniques and strategies.

Each of the two parts contains a brief introduction, which anticipates the main lines of the research program developed and the results obtained, is divided in Chapters and has a specific Bibliography.

Contents

Part I

CANCELLATION OF CHIRAL ANOMALIES

Introduction

The classical description of the fundamental forces of nature is based on symmetry assumptions. According to Noether theorem, these symmetries imply the existence of corresponding conservation laws. It can happen that the process of quantization of a theory is not compatible with some of these symmetries, so that the corresponding conservation laws are violated in the quantized theory: this is called an *anomaly*.

In this First Part we discuss a particular, but very important, class of anomalies: the *chiral* (gauge) anomalies, which arise in presence of chiral fermion representations of some symmetry group. We call *local* those anomalies the arise from infinitesimal variations and that can be treated perturbatively.

Even though anomalies can be expressed as local operators, functionals of the gauge bosons and linear in the gauge parameters, they come from the variation of non-local operators and cannot be removed by the addition of any local counter-term functional of the gauge bosons only. In this sense, they are long-distance effects, and they cannot be affected by the behavior of the underlying short-distance theory.

Local anomalies can be compute in various manners. From the impossibility, in some 1-loop amplitude in the perturbative expansion, to find a regularization prescription which is compatible at the same time with the conservation laws of axial and vectorial currents. In the path integral formulation, they originate from the impossibility of defining an invariant fermionic measure. The most efficient way to compute anomalies makes use of differential geometry techniques, and relates their existence to the index theorems describing the topological structure of the fiber bundle of a gauge symmetry based on the anomalous symmetry group.

While global symmetries with anomalous conservation laws are known to describe phenomenologically relevant processes, such as the decay of the π^0 meson into photons, an anomalous local symmetry implies the inconsistency of the theory: the presence of anomalies in gauge symmetries spoils the unitarity of the theory, in the sense that there is no way of defining consistently the space of physical states.

In gauge theories, there is also a non-perturbative class of anomalies: the *global* anomalies that occur when the gauge group has a non trivial homotopy group and that may lead to an inconsistent path integral formulation of the theory.

The main motivation for the study of chiral anomalies and of anomaly cancellation mechanisms is that we know that some of the fundamental forces of nature are chiral gauge theories. Model-building requires then severe constraints in order to cancel these anomalies.

Anomalies can be canceled either by appropriate choices of the chiral representations for fermions, as it happens in the Standard Model (SM), or by the addition to the action of higher order couplings between gauge bosons and other fields. Even though this second possibility is not compatible with renormalizability in $D = 4$ space-time dimensions, it is interesting to study non-renormalizable theories, such as the low-energy effective actions of electro-weak and strong interactions. Moreover, in $D > 4$ gauge theories are themselves non-renormalizable.

Two significant examples of non-trivial anomaly cancellation mechanisms, based on the presence of higher order (local) operators in theories with anomalous chiral spectrum, are the Green-Schwarz (GS) mechanism and the addition of a Wess-Zumino (WZ) coupling.

The GS mechanism cancel all the possible reducible (local) gauge anomalies thanks to the presence in the spectrum of even rank tensors with non-trivial properties under gauge transformations. Phenomenology is affected because some light components of these fields remain in the spectrum as pseudo-scalar fields with axion-like couplings.

The second example is based on effective field theory techniques. An anomalous theory with gauge group G can be made gauge invariant by adding the WZ term to the action and coupling G-valued scalar fields that transform non-linearly under the corresponding gauge symmetry.

Usually the WZ term is used in effective field theories in order to reproduce the anomalies of an underlying theory. It results to be very useful also in the computation of global anomalies. This interpretation can be reversed and instead of *reproducing* the anomalies of an underlying theory, we can use the WZ term to *cancel* the anomalies of a theory with chiral fermions. What is most interesting is that some, or even all, of the G-valued fields entering the WZ term can be identified with the would-be Goldstone bosons of the spontaneous breaking of the gauge symmetry, so that no new degrees of freedom are needed in order to cancel chiral anomalies of spontaneously broken gauge symmetries.

While the GS mechanism has some intrinsic limitations, since it can be applied only to (local) reducible anomalies, the WZ term can always be added and any kind of gauge anomaly canceled.

A detailed discussion of the case when both local and global anomalies are present in D-dimensional effective theories of a gauge group G spontaneously broken to a subgroup H requires to generalize the techniques for the construction of the WZ terms in order to avoid algebraic and topological obstructions. This leads to establish a set of conditions under which this cancellation is possible.

In order to illustrate the phenomenological consequences of the requirement of anomaly cancellation in chiral gauge theories, we discuss a $D = 6$ dimensional extension of the SM. The main assumption is that the gauge and matter fields propagate in $D = 6$ dimensions: matter fields are assigned to chiral representations Q, U, D, L and E of $G = SU(3)_c \times SU(2)_L \times U(1)_Y$, with the same quantum numbers of one family of SM fermions. A chiral field in $D = 6$ reduces to a Dirac spinor in $D = 4$: the 4-dimensional chirality is recovered after compactification on non-trivial background manifolds (an orbifold, for instance). The chirality assignents (in $D = 6$) are free parameters, to be choosen according to the requirement of anomaly cancellation

While the $D = 4$ effective theory is free from triangular and global anomalies in $D = 6$ quadrangular diagrams give rise to anomalies. Global anomalies are present as well. We therefore must cancel global and local gauge and gravitational anomalies. Since we cannot do this just by an appropriate choice of the chiral fields, we are forced to use one of the mechanisms descibed above.

If the GS mechanism is applied, we need first to ensure that global and irreducible anomalies are absent. Cancellation of irreducible anomalies is achieved within each family by choosing strong and gravitational couplings to be vector-like: we need to add a fermion singlet N to the spectrum, and assign opposite chirality to doublets and singlets with the same $SU(3)_c$ quantum numbers. Reducible anomalies can be compensated by the addition of two GS fields (2-forms). Cancellation of global anomalies is obtained by replicating the field content of each family: this is possible with a number of generations $N_g = 0 \bmod 3$ with identical quantum numbers and chirality assignements. On the other hand, axions remain in the $D = 4$ effective theory, besides the SM fields, due to the GS mechanism. These could provide a solution to the strong-CP problem, but impose a severe bound on the fundamental scale of the theory, which must be in the range of usual GUT models. This model cannot be a TeV-scale extra-dimension model.

Were the symmetry G unbroken, this would be the end of the story: cancellation of anomalies requires the introduction of new degrees of freedom, which modigy in a significant manner the low-energy phenomenology. However, the group G is broken by the Higgs mechanism to the subgroup $H = SU(3)_c \times U(1)_{\text{e.m.}}$, and the theory contains a set of pseudo-scalars transforming according to non linear realizations of the broken symmetry. These are the would-be Goldstone bosons, which in the unitary gauge constitute the longitudinal components of massive gauge bosons. They can be coupled to gauge bosons, through the WZ gauge non-invariant interactions, and cancel the anomalous variations, according to the second mechanism discussed.

The cancellation of all the local and global anomalies is possible as long as the vanishing of gravitational anomalies and gauge anomalies depending only on the generators of the unbroken group H are absent. This requirement forces the fermionic spectrum to be the same as before: $SU(3)_c$ and gravity must have vector-like couplings to matter. But now, no GS fields are present, and there are no axions, so that the scale of the theory can be lowered down to the TeV scale. The WZ couplings can be defined in such a way as to cancel also the global anomalies, so that there is no restriction on the number of families.

Chapter 1

Local and global gauge anomalies

In this Chapter we will discuss the chiral anomalies, which arise in presence of chiral fermion representations of some symmetry group. Since their discovery [1], chiral anomalies have played and continue to play an important role in high energy physics, for a two-fold reason: anomalies of global symmetries can explain the existence of symmetry-violating physical processes, such as the π_0 decay, while anomalies in local (gauge) symmetries lead to violation of unitarity and to the inconsistency of a given model.

In this chapter we review some of the classical techniques to compute perturbative anomalies of chiral symmetries, which will be called *local (gauge) anomalies* in the following. We will repeat the computation of chiral anomalies in three different (but equivalent) ways: with Feynman diagrams, with the Feynman path integral and using differential geometry.

Historically, the diagrammatic approach [2] first lead to the discovery of the chiral anomaly [1], and in this formulation it can be seen from the impossibility of defining a symmetry-preserving regulator. Fujikawa then discovered a procedure [3] which gave an elegant mathematical interpretation of the anomaly in terms of the non-invariance of the measure in the Feynman path integral. The origin of the chiral anomalies from the topology and geometry of the underlying theory can be seen introducing differential geometry techniques: the Stora-Zumino descent equations [4], based on the Wess-Zumino consistency conditions [5]. This is also the easiest procedure to compute anomalies in any gauge theories.

We explicitly discuss the main consequence of the presence of anomalies in gauge symmetries: the lack of unitarity of the theory, due to the fact that it is not possible to define a conserved nilpotent Becchi-Rouet-Stora-Tyutin (BRST) charge and define consistently the space of physical states.

We also discuss a non-perturbative class of anomalies, called *global gauge anomalies* (non to be confused with the anomalies of a global symmetry), that occur when the gauge group has a non trivial homotopy group. In this case, even if local anomalies are absent, when performing gauge transformations that cannot be deformed to the identity one can still cause a phase shift of the effective action, leading to an inconsistent path integral formulation of the theory. We review the basic ideas of the proof of the existence of such anomalies, and of their computation [7].

This is by no means an exhaustive review of all the mathematical and physical aspects of the study of anomalies. We use the differential geometry language only in order to show a technical way of computing non Abelian anomalies, but we do not enter in details about the relation of anomalies with topological aspects (as the index theorems), and about the close correlation between Abelian, parity and non-Abelian anomalies in different dimensions. Moreover, we deal here only with anomalies of internal symmetries, and not discuss gravitational or Supersymmetry anomalies. Rather, we want to emphasize that each of the different techniques discussed allows to clarify some peculiar aspects of the anomaly, which must be kept in mind in defining consistent procedures for anomaly cancellation. This cancellation mechanisms will be the subject of the subsequent chapters.

Most of the content of this Chapter, besides on the original papers, can be found in several textbooks. We refer the interested reader to [8, 9] for a large overview on the subject, to [10] for the more mathematical aspects, such as the index theorems and the connection between Abelian, parity and non-Abelian anomalies in the descent equation procedure, and to [11] for a discussion on BRST formulation.

1.1 Adler-Bell-Jackiw anomaly

Let us consider first the computation that lead to the discovery of the Adler-Bell-Jackiw (ABJ) anomaly. We write the theory of one spinor field ψ with charge Q coupled to the a $U(1)$ gauge field A_μ with coupling g:

$$\mathcal{L} = \bar{\psi}\,(i\not{D} - m)\,\psi - \frac{1}{4}\,F_{\mu\nu}F^{\mu\nu}\,, \tag{1.1}$$

where the Dirac operator, containing the covariant derivative, id defined as:

$$\not{D} = \gamma^\mu\,(\partial_\mu + igQA_\mu) \tag{1.2}$$

We define the vector, axial and pseudo-scalar currents:

$$J_\mu = \bar{\psi}\,\gamma_\mu\,\psi\,, \tag{1.3}$$

$$J^5_\mu = \bar{\psi}\,\gamma_\mu\,\gamma^5\,\psi\,, \tag{1.4}$$

$$J^5 = \bar{\psi}\,\gamma^5\,\psi\,, \tag{1.5}$$

They satisfy the following conservation laws:

$$\partial^\mu J_\mu = 0\,, \tag{1.6}$$

$$\partial^\mu J^5_\mu = 2im\,J^5\,. \tag{1.7}$$

In the $m \to 0$ limit the spinor Lagrangian possesses two $U(1)$ global symmetries:

$$U(1)_V \;:\; \psi \to e^{igQ\alpha}\,\psi\,, \tag{1.8}$$

$$U(1)_A \;:\; \psi \to e^{igQ\beta\gamma^5}\,\psi\,, \tag{1.9}$$

so that J_μ and J^5_μ become the classical Noether currents of $U(1)_V$ and $U(1)_A$, respectively.

On the other hand, these two conservation laws cannot be both preserved after quantum corrections has been included. The simplest way to see this is the computation of the amplitude of the Feynman diagrams in Fig. 1.1.

The explicit computation, performed using a gauge invariant regulator (with a Pauli-Villars regulator or with the t'Hooft-Veltman dimensional regularization)

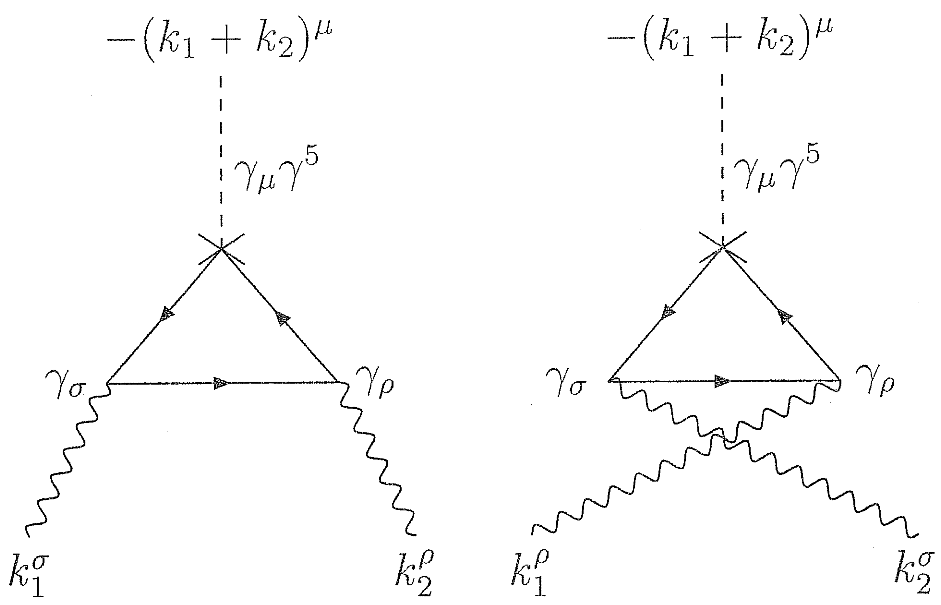

Figure 1.1: Diagrams contributing to the axial anomaly

gives a violation of the axial current conservation. Explicitly, one finds that the depicted amputated graphs sum up to give [1]: [1]

$$-(k_1+k_2)^\mu\, R_{\sigma\rho\mu} \;=\; -\mathcal{N}\,(A_1 - A_2)\, k_1^\tau k_2^\mu\, \epsilon_{\tau\sigma\rho\mu}\,,\,, \tag{1.10}$$

where the following definitions are used:

$$\begin{aligned}
A_1 &= (k_1\, k_2)\, A_3 + k_2^2\, A_4\,,\\
A_2 &= (k_1\, k_2)\, A_6 + k_1^2\, A_5\,,\\
A_3(k_1,k_2) &= -A_6(k_2,k_1) = -16\pi^2\, I_{11}(k_1,k_2)\,,\\
A_4(k_1,k_2) &= -A_5(k_2,k_1) = -16\pi^2\, \left(I_{20}(k_1,k_2) - I_{10}(k_1,k_2)\right)\,,\\
I_{st}(k_1,k_2) &= \int_0^1 \mathrm{d}x \int_0^{1-x} \mathrm{d}y\, \frac{x^s\, y^t}{y(1-y)k_1^2 + x(1-x)k_2^2 + 2xy(k_1\, k_2)}\,,\\
\mathcal{N} &= \frac{-ig^2Q^2}{(2\,\pi)^4}\,,
\end{aligned}$$

[1]The amplitude $R_{\sigma\rho\mu}$ of the amputated graphs contains actually other contributions proportional to antisymmetrized products of k_1 and k_2, which vanish when multiplied by the external momenta.

When one considers on-shell massless photons, the complicated non-analytic expression coming from the integral in eq. (1.10) reduces to a local function:

$$-(k_1+k_2)^\mu R_{\sigma\rho\mu} = \mathcal{N}\, 8\,\pi^2\, k_1^\tau k_2^\mu\, \epsilon_{\tau\sigma\rho\mu}\,. \tag{1.11}$$

We can rewrite eq. (1.11) as the non-vanishing of the divergence of the axial current:

$$\partial^\mu J^5_\mu = \mathcal{G}\,[A] \tag{1.12}$$

$$= \frac{g^2 Q^2}{16\,\pi^2}\,\epsilon_{\tau\sigma\rho\mu}\, F^{\tau\sigma} F^{\rho\mu}\,. \tag{1.13}$$

If gauge non-preserving regulators are introduced, as an explicit cut-off in the momentum integration, then part of the anomaly appears in the divergence of the vector-like current. It is even possible to define the regulator in such a way to preserve the axial current conservation law, at the price of loosing the conservation of the vector-like current. The anomaly consists indeed in the absence of any symmetry preserving regulator: in general there is some regularization dependence of the form of the anomaly itself, but it cannot be removed from the theory, the final result being physical (regularization independent).

The fact that the anomaly is a physical effect, and not just an artifact of the renormalization, is clear if one uses the effective vertex $R_{\sigma\rho\mu}$ to define an effective operator $\Gamma[A,B]$, where B_μ is an external field coupling to the axial current. This operator is a complicated non-local functional. The anomaly is the (local) variation of this non-invariant operator, but it is not possible to write any counter-term, in the form of a local functional of the fields A_μ and B_μ, in order to cancel this non-local operator and restore gauge invariance.

Although anomalies are identified through the ultraviolet behavior of the theory (the absence of any symmetry-preserving regulator in the renormalization procedure), they are intrinsically an infrared property, since it is not possible to remove them with the addition of any local counter-term, and therefore they cannot depend on the unknown modifications of the short distance physics. This crucial observation is the starting point of Chapter 3.

When a non-Abelian symmetry group G is present in the Lagrangian instead of

$U(1)$, the expressions eq. (1.12) and eq. (1.13) can be generalized to:

$$\begin{aligned} \partial^\mu J^5_\mu &= \mathcal{G}[A] && (1.14)\\ &= \frac{g^2}{16\,\pi^2}\,\epsilon_{\tau\sigma\rho\mu}\,\mathrm{Tr}\,[F^{\tau\sigma}F^{\rho\mu}]\,, && (1.15)\end{aligned}$$

where $F^{\tau\sigma}$ are the field strength tensors of G.

This is the so called *Abelian anomaly*. In the literature one often finds also the expression *singlet anomaly*. Notice that *Abelian* refers to the fact that the anomalous current does not carry an internal symmetry index, while in the r. h. s. of eq. (1.15) the field strength fields of any gauge group G may appear.

It must be noticed that also quadrangular and pentagonal diagrams contribute to this expression. On the other hand, all of these contributions are strongly related, and computing the triangle is enough to fix the complete one-loop form of the anomaly. The requirement of covariance is in fact enough to determine the other contributions. As we will see later in discussing the non-Abelian anomaly (the computation of which is postponed to Section 1.3) other conditions can be imposed, obtaining different results. All of those different expressions are indeed equivalent up to the addition of polynomial functionals to the definition of the currents.

An even stronger statement is true: the Adler-Bardeen theorem [12] states that the full structure of the chiral anomaly is given by the triangle loop. This means that higher order loop corrections can only renormalize the fields and the charges involved.

The ABJ anomaly solved an important problem related to the decay of the π^0 meson onto two photons. According to the Partial Conservation of Axial Current (PCAC) hypothesis, the conservation laws of the axial $SU(2)$ isospin currents have to be modified accordingly as:

$$\partial^\mu J^{5\,(3)}_\mu = f_\pi m_\pi^2 \phi^{(3)} a_\pi + \frac{\alpha}{8\,\pi}\,\epsilon^{\mu\nu\rho\sigma}\,F_{\mu\nu}\,F_{\rho\sigma}\,, \qquad (1.16)$$

in order for the pion to decay with the expected rate. The first term on the r. h. s. takes into account that the chiral symmetry is not exact, its violations being proportional to the mass of the Goldstone bosons $\phi^{(5)}$, while the second term is produced by the anomalous coupling of neutral component of the axial current $\phi^{(3)} \sim \pi^0$ to the electromagnetic fields. The estimate of the π^0 life-time performed with an effective Lagrangian taking into account only the first term would be several orders of magnitude longer than the experimental result. The fact that the anomaly couples axial currents to the electromagnetic fields, allows for the existence of a new decay channel with a stronger coupling, and the lifetime can be estimated to be in the correct experimental range.

The diagrammatic computation of anomalies can be easily extended to higher (even) space-time dimensions. In odd dimensions chirality is not defined, so that this kind of anomalies is absent.

In $D = 2n$ dimensions it is possible to define the matrix Γ_{D+1}, to play the role of the γ^5 in $D = 4$. The equivalent of the triangle diagram is a one-loop diagram with fermions on the internal lines and M vertexes, one of which contains an axial-like coupling $\Gamma_{D+1}\,\Gamma_N$, while in the others one has vector-like couplings trough the Γ_N matrices to the vector bosons. The amplitude is proportional to the trace of the product of Γ_{D+1} with all the Γ_N matrices coming from the vertexes and the fermion propagators.

The lowest order non-vanishing trace gives a D dimensional Levi-Civita tensor ϵ with all the D indexes saturated on independent external polarization and momentum vectors: only $M - 1$ out of the momenta are independent, so that:

$$\begin{aligned} D &\leq (M-1) + M = 2M - 1\,, \\ M &\geq 1 + \frac{D}{2}, = 1 + n\,. \end{aligned} \tag{1.17}$$

In D dimensions the first contribution to the anomaly comes from a loop with $1 + D/2$ vertexes (see Fig. 1.2).

It is of interest also the case when chiral currents carry some internal symmetry index:

$$J^{5\,a}_{\mu} \equiv \bar{\psi}\,\gamma_\mu\,\gamma^5\,T^a\,\psi\,, \tag{1.18}$$

where T^a are the (hermitian) generators of some Lie group G. The explicit expression of the anomaly in this case is slightly different from the Abelian case, and, although

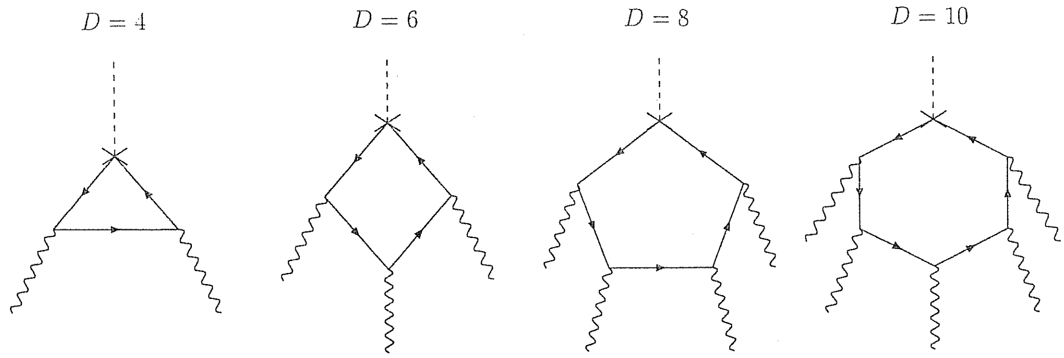

Figure 1.2: Diagrams contributing to the anomaly in $D = 2n$ even dimensions.

it can be computed from 1-loop amplitudes too, we will write its explicit form only in section 1.3. As far as the group structure of the anomaly is concerned, if one defines:

$$J^5_\mu \equiv J^{5\,a}_\mu T^a , \tag{1.19}$$

one understands that the anomaly must be written as a G-invariant quantity built out of the product of $1 + D/2$ generators T^a of G. In $D = 4$, this leads to the definition of the D^{abc} symbols as traces on symmetrized products:

$$D^{abc} \equiv \text{Tr}\left[\{T^a, T^b\}T^c\right] . \tag{1.20}$$

The non-vanishing of some of these symbols signals the presence of (non-Abelian) anomalies. This definition generalizes to any even dimension, where the anomaly is controlled by similar objects with $1 + D/2$ indexes.

For a simple compact Lie group G, the singlets which can be built as products of the generators are the Casimir invariants of the group and all the possible products of these. The number of such Casimir operators coincides with the rank of the group, and their order (number of generators the combination of which composes the Casimir invariant) has been classified according to Table 1.1 [13]

Group	Order
$SU(n+1)$	$2, \cdots, n+1$
$SO(2n+1)$	$2, 4, \cdots, 2n$
$Sp(2n)$	$2, 4, \cdots, 2n$
$SO(2n)$	$2, 4, \cdots, 2n-2, n$
G_2	$2, 6$
F_4	$2, 6, 8, 12$
E_6	$2, 5, 6, 8, 9, 12$
E_7	$2, 6, 8, 10, 12, 14, 18$
E_8	$2, 8, 12, 14, 18, 20, 24, 30$

Table 1.1: Order of the Casimir invariants of the classical compact simple Lie groups.

We call, in the context of non-Abelian anomalies, *non-factorizable anomalies* those which are proportional to one of the Casimir invariants of the group. In D dimensions, this requires the existence of a Casimir invariant of order $1 + D/2$. All the other (local) anomalies, *i.e.* the Abelian anomalies and the anomalies written as the product of n Casimir invariants of order m_i $(i = 1, \cdots, n)$, with $\sum_i m_i = 1+D/2$, are called *factorizable anomalies.*

1.2 Anomalies in the Feynman path integral formulation

In spite of the simplicity of the diagrammatic computation of the (Abelian) anomaly performed in the previous section, such a procedure is not very enlightening about the origin of the anomaly itself: it may seem quite obscure that, after deriving Feynman rules from a Lagrangian enjoying a certain symmetry, the application of these rules to a one-loop computation yields an amplitude which is not compatible with the symmetry itself.

The correct interpretation of this result has been given by Fujikawa [3], and can be seen explicitly in the Feynman path integral formulation In this formulation it is apparent that equations of motion and transition probabilities are not derived directly from the classical action $\mathcal{S}$, but rather the Feynman path integral, which not necessarily enjoys all the symmetries of $\mathcal{S}$. Consider the effective action $W[A]$ defined by the functional integration over the fermionic degrees of freedom:

$$e^{iW[A]} = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \; e^{i\mathcal{S}}, \tag{1.21}$$

where

$$\mathcal{S} = \int \mathrm{d}^4x \, \mathcal{L}. \tag{1.22}$$

Because the fermionic measure $\mathcal{D}\bar{\psi}\mathcal{D}\psi$ is not invariant under chiral transformations, after quantization we find a violation of the symmetries of the classical theory.

In the Euclidean space formulation of the functional integral (analytic continuation $ix^0 = x^4$):

$$e^{iW[A]} = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \; e^{-\mathcal{S}_e}, \tag{1.23}$$

where

$$\mathcal{S}_e = \int \mathrm{d}^4x_e \mathcal{L}_e, \tag{1.24}$$

consider the massless Lagrangian:

$$\mathcal{L} = \bar{\psi} \, i\not{D} \, \psi. \tag{1.25}$$

In this simple case the path integral reduces to the computation of the fermionic determinant of the hermitian Dirac operator $\not{D}$:

$$e^{iW[A]} = \det(i\not{D}). \tag{1.26}$$

In order to compute this determinant one needs to regularize properly the fermionic measure. This is done by decomposing the spinors into eigenfunctions φ_n

$$\psi = \sum_n a_n \varphi_n, \tag{1.27}$$

$$\bar{\psi} = \sum_m \varphi_m^\dagger \bar{b}_m, \tag{1.28}$$

relative to the real eigenvalues λ_n of the Dirac operator

$$\not{D}\,\varphi_n \;=\; \lambda_n\,\varphi_n\,, \tag{1.29}$$

and then regularizing the infinite product with a smooth functional, for instance a Gaussian cut-off [3]. The normalizations are chooses according to:

$$\delta_{nm} \;=\; \int \mathrm{d}^4x\,\varphi_n(x)^\dagger\varphi_m(x)\,, \tag{1.30}$$

$$\delta(x-y) \;=\; \sum_n \varphi_n(x)\varphi_n^\dagger(y)\,, \tag{1.31}$$

The functional integration measure is defined in terms of Grassmann variables $\bar{b}_n$ and a_n as:

$$\mathcal{D}\bar{\psi}\mathcal{D}\psi \;\equiv\; \lim_{N\to\infty}\prod_{n=1}^{N}\mathrm{d}\bar{b}_n\,\mathrm{d}a_n\,, \tag{1.32}$$

so that

$$e^{iW[A]} \;=\; \det\,(i\not{D}) \;=\; \prod_n i\lambda_n\,. \tag{1.33}$$

Performing now a change of integration variables corresponding to an infinitesimal chiral transformation in the form of eq. (1.9), but with (local) parameter $\beta(x)$:

$$\delta\,\psi \;=\; igQ\beta\,\gamma_5\,\psi\,, \tag{1.34}$$

$$\delta\,\bar{\psi} \;=\; \bar{\psi}\,igQ\beta\,\gamma_5\,, \tag{1.35}$$

the measure transforms into:

$$\mathcal{D}\bar{\psi}\mathcal{D}\psi \;\to\; \mathcal{D}\bar{\psi}\mathcal{D}\psi\,J[\beta,A_\mu]\,, \tag{1.36}$$

where

$$J[\beta,A_\mu] \;=\; \mathrm{Exp}\left[-2igQ\sum_n\int \mathrm{d}^4x_e\,\beta(x)\;\varphi_n^\dagger(x)\,\gamma_5\,\varphi_n(x)\right]. \tag{1.37}$$

The computation of the Jacobian can be performed introducing a Gaussian regulator: [2]

$$
\begin{aligned}
&\sum_n \varphi_n^\dagger(x)\,\gamma_5\,\varphi_n(x) = \\
&= \lim_{M\to\infty} \sum_n \varphi_n^\dagger(x)\,\gamma_5\,\mathrm{Exp}\left[\frac{-\lambda_n^2}{M^2}\right]\varphi_n(x) \qquad (1.38)\\
&= \lim_{M\to\infty} \sum_n \varphi_n^\dagger(x)\,\gamma_5\,\mathrm{Exp}\left[\frac{-\not{D}^2}{M^2}\right]\varphi_n(x)\\
&= \lim_{M\to\infty} \int \frac{\mathrm{d}^4k_e}{(2\,\pi)^4}\,\mathrm{Tr}\left\{e^{-ikx}\,\gamma_5\,\mathrm{Exp}\left[\frac{-\not{D}^2}{M^2}\right]e^{-ikx}\right\}\\
&= \lim_{M\to\infty} \int \frac{\mathrm{d}^4k_e}{(2\,\pi)^4}\,M^4e^{-k^2}\,\mathrm{Tr}\left\{\gamma_5\,\mathrm{Exp}\left[-\frac{2ik_\mu D^\mu}{M} - \frac{D_\mu D^\mu}{M^2} - \frac{gQ\gamma^\mu\gamma^\nu F_{\mu\nu}}{2M^2}\right]\right\}\\
&= \lim_{M\to\infty} \frac{1}{2!}\,\frac{M^4}{4M^4}\,(gQ)^2 \int \frac{\mathrm{d}^4k_e}{(2\,\pi)^4}\,e^{-k^2}\,\mathrm{Tr}\left\{\gamma_5\,\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\right\}F_{\mu\nu}F_{\rho\sigma}\\
&= \frac{-g^2Q^2}{32\,\pi^2}\,\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}
\end{aligned}
$$

In the computation we used the following relations: [3]

$$
\begin{aligned}
\varphi_n(x) &= \int \frac{\mathrm{d}^4k}{(2\,\pi)^2}\,e^{ikx}\tilde{\varphi}_n(k)\,, &(1.39)\\
\sum_n \tilde{\varphi}_n^\dagger(l)\Gamma\tilde{\varphi}_n(k) &= \mathrm{Tr}\,\Gamma\,\delta(l-k)\,, &(1.40)\\
\not{D}^2 = D_\mu D^\mu &+ \frac{igQ}{4}\,[\gamma^\mu,\gamma^\nu]\,F_{\mu\nu}\,, &(1.41)\\
\int_{-\infty}^{+\infty} \mathrm{d}^{2n}k\,e^{-k^2} &= \pi^n\,, &(1.42)\\
\mathrm{Tr}\ \gamma_5\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma &= -4\epsilon^{\mu\nu\rho\sigma}\,. &(1.43)
\end{aligned}
$$

And finally the Jacobian becomes:

$$
\begin{aligned}
J[\beta, A_\mu] &= \mathrm{Exp}\left[-igQ\int \mathrm{d}^4x_e\,\beta\,\frac{-g^2Q^2}{16\,\pi^2}\,\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right]\,, \qquad (1.44)\\
&= \mathrm{Exp}\left[-gQ\int \mathrm{d}^4x\,\beta\,\frac{g^2Q^2}{16\,\pi^2}\,\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right]\,,
\end{aligned}
$$

[2] One could also see that only the zero modes of the Dirac operator do contribute to the Jacobian: the anomaly is related to the Atiyah-Singer index theorem (see for instance [10] and references therein) which counts the difference between the number of Left and Right Handed chiral zero-modes of the Dirac operator. This index is a topological invariant of the theory: this is another way of seen that the anomaly is intrinsically a long-distance effect.

[3] In Euclidean space, with $\gamma^4 = i\gamma^0$.

$$= \mathrm{Exp}\left[-gQ\int \mathrm{d}^4x\,\beta\,\mathcal{G}[A]\right]. \tag{1.45}$$

The meaning of this non-trivial Jacobian can be understood going back to the partition function eq. (1.21), with the Lagrangian in eq. (1.25), and computing the variation under a chiral transformation. To do this we get back to the change of variables we started from:

$$\mathcal{L} \to \mathcal{L}' = \mathcal{L} - gQ\,\partial^\mu\beta\,J^5_\mu\,, \tag{1.46}$$

$$e^{iW[A]} \to e^{iW'[A]} = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, e^{i\int \mathrm{d}^4x\mathcal{L}}\;\mathrm{Exp}\, i\int \mathrm{d}^4x\,\left[-gQ\;\partial^\mu\beta\,J^5_\mu - gQ\,\beta\,\mathcal{G}[A]\right] \tag{1.47}$$

$$= \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, e^{i\int \mathrm{d}^4x\mathcal{L}}\left[1 + i\int \mathrm{d}^4x\, gQ\,\beta\,(\partial^\mu J^5_\mu - \mathcal{G}[A]) + \cdots\right].$$

The gauge field being invariant under the action of these chiral transformations, it must be $W[A] = W'[A]$. Then for consistency one finds:

$$\langle \partial^\mu J^5_\mu\rangle = \mathcal{G}[A]\,, \tag{1.48}$$

where we define, for a generic operator Ω:

$$\langle\Omega\rangle \equiv \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, e^{-\int \mathrm{d}^4_e x\mathcal{L}}\,\Omega\,, \tag{1.49}$$

This result agrees with the previous section. We have thereby again showed that after quantization the axial current is not conserved.

Let us stress that we used here the invariance of the effective action because the symmetry we are dealing with is a global symmetry. On the other hand, it can be interesting to see what happens if one takes the $U(1)_A$ symmetry to be local, by introducing a gauge field B_μ, and modifying the covariant derivative in such a way as to have a gauge invariant classical Lagrangian.

$$D_\mu = \partial_\mu + igQ\,A_\mu + igQ\,B_\mu\gamma^5\,, \tag{1.50}$$

$$U(1)_V \;:\; \begin{cases} A_\mu \to A_\mu - \partial_\mu\alpha \\ B_\mu \to B_\mu \end{cases}\,, \tag{1.51}$$

$$U(1)_A \;:\; \begin{cases} A_\mu \to A_\mu \\ B_\mu \to B_\mu - \partial_\mu\beta \end{cases}\,. \tag{1.52}$$

In this case it is convenient to work in terms of Left- and Right-Handed fields:

$$P^{L,R} \equiv \frac{1 \pm \gamma^5}{2}, \tag{1.53}$$

$$\psi_{L,R} \equiv P^{L,R}\psi, \tag{1.54}$$

$$A_\mu^{L,R} \equiv A_\mu \pm B_\mu, \tag{1.55}$$

$$\alpha^{L,R} \equiv \alpha \pm \beta, \tag{1.56}$$

Then, if one defines the partition function as:

$$e^{iW[A^L,A^R]} = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, e^{i\mathcal{S}}, \tag{1.57}$$

$$\mathcal{S} = \int \mathrm{d}^4x\, \bar{\psi}\, i(\partial_\mu + igQ\, A_\mu^{L,R}\, P^{L,R})\,\psi, \tag{1.58}$$

the final result is that, in spite of $\mathcal{S}$ being invariant under local transformations of both $U(1)_{L,R}$, the effective action is not, its variation being:

$$\delta W[A^L, A^R] = -gQ \int \mathrm{d}^4x\, \alpha^{L,R}(x)\, \mathcal{G}[A^{L,R}], \tag{1.59}$$

$$\mathcal{G}[A] = \frac{g^2Q^2}{32\,\pi^2}\,\epsilon_{\tau\sigma\rho\mu}\, F^{\tau\sigma}F^{\rho\mu}. \tag{1.60}$$

Notice the factor of 2 in $\mathcal{G}[A]$, due to the projector operators.

Finally, a comment is in order: the computation of chiral anomalies by means of functional integration techniques is quite elegant, and allows one to understand their origin, but some subtleties are not apparent in this approach, and require some caution. In particular, we have seen in the previous section that the diagrams giving rise to the anomaly has a very complicated non-analytic structure in the off-shell part of the amplitude: the anomaly is a local functional resulting from the variation of a non-local functional. This structure is very difficult to identify in the path integral formulation. When we will discuss the renormalizability of a chiral theory, in the presence of anomalies canceled by effective operators, the existence of these non-analytic amplitudes will prove of crucial importance.

1.3 Anomalies and differential geometry.

Even though chiral anomalies have been discovered computing one loop diagrams, and their origin from the fermionic functional measure results quite clear in the path integral formulation, still some aspects are not very easy to deal with in neither of the two languages. The explicit computations involved are quite complicated and not very easy to extend to higher dimensions, non-Abelian groups and to the case where other fields give contribution to the anomalies besides the chiral fermions. Further, some ambiguities in the correct definition of the anomaly, and its relation to the underlying geometry and topology of the theory, such as the relation to the Atiyah-Singer index theorem, or to the homotopy classification of the gauge symmetry groups, are very difficult to extract.

A very powerful tool has been designed [4] in order to discuss all of these aspects: a differential geometry formulation of the gauge and space-time structure of the theory, which is written in terms of elements of $\Omega^p(G)$ and $\Omega^r(M)$, which are, respectively, the space of smooth p-forms on G and of r-forms on the base manifold $M = \mathcal{M}^D$. It allows also to identify the relation to the non-perturbative aspects of the theory itself, as like the existence of solitonic solutions. In this language the anomaly takes a very compact and simple form, and is rather easy to compute. Furthermore, this formulation makes quite clear the interplay between gauge anomalies and unitarity in the BRST quantization procedure.

Let us consider a generic compact semi-simple (non-Abelian) group G, with hermitian generators T^a such that

$$\begin{cases} \mathrm{Tr}\, T^a T^b & = & \frac{1}{2}\,\delta^{ab}\,, \\ \left[T^a, T^b\right] & = & i\, f^{abc}\, T^c\,. \end{cases} \tag{1.61}$$

Let us couple the gauge fields of G to a chiral fermion ψ_L, with the Lagrangian

$$\mathcal{L} \;=\; \bar{\psi}_L\, i\not{D}\, \psi_L\,, \tag{1.62}$$

where the covariant derivative is

$$D \;=\; \partial + iT^a A^a\,. \tag{1.63}$$

The action of gauge transformations with parameter $\alpha = \alpha^a T^a$

$$U \;\equiv\; e^{i\alpha^a T^a}\,, \tag{1.64}$$

is given by

$$\psi_L \to \psi'_L = U\,\psi_L\,, \tag{1.65}$$

$$A \to A' = U\,A\,U^{-1} + i\,\partial\,U\,U^{-1}\,, \tag{1.66}$$

or, in infinitesimal form,

$$\delta\psi_L = i\alpha^a T^a\,\psi_L\,, \tag{1.67}$$

$$\delta A^a_\mu = -\partial_\mu\alpha^a - f^{abc}\,\alpha^b\,A^c_\mu\,. \tag{1.68}$$

These transformations lead to the classical conservation law of the chiral current:

$$D^\mu\,J^L_\mu = 0\,, \tag{1.69}$$

where:

$$J^L_\mu = (J^L_\mu)^a\,T^a\,, \tag{1.70}$$

$$(J^L_\mu)^a = \bar{\psi_L}\,i\gamma_\mu\,T^a\,\psi_L\,. \tag{1.71}$$

$$\tag{1.72}$$

These formulae hold also in the case of anomalies related to a global symmetry group G, if one formulates the theory in terms of external fields A^a_μ. We work in generic $D = 2n$ even dimensional Minkoski space-time. Notice that we re-scale out, in this section, the gauge coupling, by setting $g = 1$, in order to simplify the expressions.

The action of gauge transformations on a generic functional $F[A]$ of the fields A, in particular on the effective action $W[A]$ defined in eq. (1.21) is represented by the functional differential operator:

$$X^a \equiv -\partial_\mu\frac{\delta}{\delta A^a_\mu} + f^{abc}\,A^b_\mu\frac{\delta}{\delta A^c_\mu}\,, \tag{1.73}$$

so that

$$\delta\,F[A] = -\int \mathrm{d}^D x\,\alpha^a\,X^a\,F[A]\,, \tag{1.74}$$

$$\delta\,W[A] = -\int \mathrm{d}^D x\,\alpha^a\,X^a\,W[A]\,. \tag{1.75}$$

With these definitions, generalizing eq. (1.59), the anomaly is given by:

$$\mathcal{G}[A]^a = X^a W[A]. \tag{1.76}$$

One can verify that

$$\left[-iX^a, -iX^b\right] = i f^{abc} \left(-iX^c\right) , \tag{1.77}$$

from which one deduces the so called Wess-Zumino consistency conditions [5]:

$$X^a \mathcal{G}^b - X^b \mathcal{G}^a = -f^{abc} \mathcal{G}^c \tag{1.78}$$

The anomaly, defined as the variation of the effective action under a gauge transformation, can be written in terms of the chiral current defined by

$$\langle J^a_\mu \rangle \equiv -\frac{\delta}{\delta A^a_\mu} W[A] , \tag{1.79}$$

as

$$D_\mu \langle J^a_\mu \rangle = \mathcal{G}[A]^a , \tag{1.80}$$

and is a solution of eq. (1.78).

in order to solve eq. (1.78), let us first rewrite the gauge theory in terms of anti-commuting objects, introducing the BRST formalism and rewriting eq. (1.78) in this language in terms of BRST transformations.

We introduce scalar anti-commuting fields (Faddeev-Popov ghosts) along the group generators v^a, and define:

$$A \equiv -iA^a_\mu T^a \, \mathrm{d}x^\mu , \tag{1.81}$$

$$v \equiv -iv^a T^a , \tag{1.82}$$

where also $\mathrm{d}x^\mu$ are anti-commuting. A is a 1-form on M, and a 0-form on G, while v is a 0-form on M and a 1-form on G.

Define the quantity:

$$G[v, A] = -\int v^a \mathcal{G}[A] \, \mathrm{d}^D x . \tag{1.83}$$

The exterior derivative is the operator:

$$\mathrm{d} \equiv \mathrm{d}x^{\mu} \frac{\partial}{\partial x^{\mu}} . \tag{1.84}$$

The anti-commuting BRST operator s is defined by:

$$sA \equiv \mathrm{d}\,v - \{A, v\} , \tag{1.85}$$

$$sv \equiv -v^2 , \tag{1.86}$$

which in components reduces to

$$sA^a_\mu = -\partial_\mu v^a - f^{abc} v^b A^c_\mu , \tag{1.87}$$

$$sv^a = -\frac{1}{2} f^{abc} v^b v^c . \tag{1.88}$$

In this way the gauge transformation of A_μ can be written as:

$$\delta_\alpha A_\mu = (\theta\, s\, A)_\mu \tag{1.89}$$

$$\alpha = \theta\, v , \tag{1.90}$$

where we introduced the anti-commuting parameter θ.

By definition, s satisfies

$$sd + ds = 0 . \tag{1.91}$$

The action of s on v is such that the consistency conditions on $\mathcal{G}$ are equivalent to:

$$sG[v, A] = 0 . \tag{1.92}$$

In the language of differential forms, the anomaly is a solution of the consistency conditions. Trivial solutions can be constructed starting from any local functional $F[A]$:

$$G[v, A] = s\, F[A] . \tag{1.93}$$

On the other hand, these solutions are not the anomalies we are looking for: one could add these local functionals to the original action and cancel the anomaly. The anomaly, being a physical property of the action, is the variation of a non-local operator, which cannot be removed from the action, as we already discussed. This observation makes clear that, once a non-trivial solution $G[v, A]$ of eq. (1.92) has been found, one can add to the action a generic local functional $F[A]$, and modify accordingly the anomaly to $G[v, A] + s\,F[A]$.

A non-trivial solution of eq. (1.92) can be computed using the Stora-Zumino descent equations. We start adding two dimensions to the $D = 2n$ dimensional space-time and defining the field strength 2-form and a the Abelian anomaly in $D + 2$ dimensions as the (2n+2)-form (the normalization is suppressed here, we will put it back only at the end of the computation):

$$F = \mathrm{d}A + A^2\,, \tag{1.94}$$

$$\mathrm{Tr}\,F^{n+1}\,, \tag{1.95}$$

and observing that $\mathrm{Tr}\,F^{n+1}$ is closed:

$$\mathrm{d}\,F = [A, F]\,, \tag{1.96}$$

$$\mathrm{d}\,\mathrm{Tr}\,F^{n+1} = 0\,. \tag{1.97}$$

So, assuming that the space-time is simply-connected, there exists a (2n+1)-form ω^0_{2n+1} such that:

$$\mathrm{Tr}\,F^{n+1} = \mathrm{d}\,\omega^0_{2n+1} \tag{1.98}$$

The Chern-Simons forms denoted by ω^p_r are p-forms on G (ghost number p) and r-forms on the base space M. The BRST operator we defined acts as a derivative in $\Omega(G)$, the space of all the forms on G.

Being $\mathrm{Tr}\,F^{n+1}$ a gauge invariant functional of the gauge fields, it is also BRST invariant, so that, using eq. (1.91):

$$0 = s\,\mathrm{Tr}\,F^{n+1} \tag{1.99}$$

$$= s\,\mathrm{d}\,\omega^0_{2n+1} \tag{1.100}$$

$$= -\mathrm{d}\,s\,\omega^0_{2n+1}\,, \tag{1.101}$$

$$s\,\omega^0_{2n+1} \equiv -\mathrm{d}\,\omega^1_{2n}\,. \tag{1.102}$$

Iterating this procedure one finds the descent equations:

$$\begin{aligned} s\,\omega^0_{2n+1} &= -\,\mathrm{d}\,\omega^1_{2n}\,, \\ s\,\omega^1_{2n} &= -\,\mathrm{d}\,\omega^2_{2n-1}\,, \\ &\dots \\ s\,\omega^{2n+1}_0 &= 0\,. \end{aligned} \tag{1.103}$$

One sees that, ω^0_{2n+1} depending only on gauge fields, ω^1_{2n} is linear in v, and satisfies eq. (1.92), so that it can be identified with the anomaly. A solution is given by:

$$\omega^1_{2n} = n(n+1)\int_0^1 \mathrm{d}t\,(1-t)\,\mathrm{Tr}\,\left[v\,\mathrm{d}(AF^n_t - 1)\right]\,, \tag{1.104}$$

$$F_t = tF + (t-t^2)\,A^2\,, \tag{1.105}$$

$$G[v,A] = -\int \omega^1_{2n}\,. \tag{1.106}$$

In particular, in $D=4$ this reduces (up to the normalization) to:

$$\omega^1_4 = \mathrm{Tr}\,v\,\mathrm{d}\left(A\mathrm{d}A + \frac{1}{2}A^3\right)\,. \tag{1.107}$$

We call BRST transformation with anti-commuting (global) parameter θ:

$$\delta_B\,F \equiv \theta\,s\,F\,. \tag{1.108}$$

A gauge transformation can be written as a BRST transformation with the identifications:

$$\delta_\alpha\,A = \delta_B\,A\,, \tag{1.109}$$

$$\alpha = \theta\,v\,. \tag{1.110}$$

Accordingly, in $D=4$ one finds that the anomaly, properly normalized starting from the direct diagrammatic computations of the Abelian anomaly [4, 14] and making now explicit the dependence on the gauge parameter g, is:

$$\delta_\alpha\,W[A] = -g\int \mathrm{d}^4x\,\alpha^a\,\mathcal{G}^a[A]\,, \tag{1.111}$$

$$\mathcal{G}^a[A] = \frac{g^2}{24\,\pi^2}\,\epsilon^{\mu\nu\rho\sigma}\,\mathrm{Tr}\left[T^a\,\partial_\mu(A_\nu\partial_\rho A_\sigma + \frac{1}{2}A_\nu A_\rho A_\sigma)\right]\,. \tag{1.112}$$

Notice that this expression is slightly different from the one we found in the Abelian case in the previous sections: it is expressed in terms of non-covariant objects. This expression is known as the *consistent non-Abelian anomaly*. This form, by going back to its components, with the gauge parameter replacing the ghost field v, gives correctly the anomaly in terms of the variation of the functional $W[A]$, and of the *consistent current* defined as its variation in respect to the gauge bosons.

On the other hand, we have already discussed the fact that the anomaly is only defined up to the variation of local functionals of the gauge fields only. One can make use of this property to define the *covariant current* $\tilde{J}^a_\mu$ by the addition to the current of a local (polynomial) functional of the gauge fields and their field strength [16], in such a way that this new current can be written in terms of covariant objects only. In the $D = 4$ case, one gets:

$$\left(D^\mu \tilde{J}_\mu\right)^a = \tilde{\mathcal{G}}[A]^a \,, \tag{1.113}$$

$$\tilde{\mathcal{G}}[A]^a = \frac{1}{32\,\pi^2}\epsilon^{\mu\nu\rho\sigma}\,\mathrm{Tr}\,\left[T^a\,F_{\mu\nu}\,F_{\rho\sigma}\right]\,. \tag{1.114}$$

This expression is known as the *covariant anomaly*. This latter expression is closer to the Abelian anomaly in eq. (1.59)

1.4 Local anomalies in gauge theories

We have seen that anomalies of a global symmetry do not cause any problem for the formulation of the quantum theory, because the modification of the related conservation laws does not affect the invariance of the action. They have useful phenomenological consequences, explaining for instance the decay rate of the π^0 meson.

But when a gauge theory is based on an anomalous symmetry, it is not possible to give a consistent probabilistic interpretation of the corresponding quantum theory, which becomes meaningless. In order to understand how this happens, we briefly review the basic features of the BRST quantization procedure. A complete treatment of this subject can be found in several textbooks, besides the original papers, and goes beyond the aims of this thesis. For simplicity, we work here in $D = 4$ dimensions.

If no anomalies are present, *i.e.* $s\,W = 0$, the BRST transformation defined previously is a (global) symmetry not only of the gauge invariant classical action, but (with appropriate definitions of the action of s on the other fields involved) also of the whole action after the introduction of the gauge-fixing terms and of the Fadeev-Popov determinant in the path integral formulation; that is, for

$$\mathcal{S}_{\text{tot}} = \mathcal{S}_{\text{gauge inv.}} + \mathcal{S}_{\text{gauge fixing}} + \mathcal{S}_{\text{Fadeev-Popov}}, \tag{1.115}$$

we have that

$$\delta_B\,\mathcal{S}_{\text{tot}} = 0 \tag{1.116}$$

This invariance allows for the definition of a Noether conserved current J^μ_B, and in particular of a conserved charge:

$$Q_B(t) = \int \mathrm{d}^3x\, J^\mu_B(x,t)\,, \tag{1.117}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_B(t) = -i\,[Q_B(t), H]\,, \tag{1.118}$$

where H is the Hamiltonian, that determines the evolution of states according to

$$|\psi(t)\rangle = e^{-i\,H\,t}\,|\psi(0)\rangle\,. \tag{1.119}$$

From the nil-potency of s one deduces that $Q_B^2 = 0$. The physical states of the theory are then defined starting from the Fock space, and selecting those state which are annihilated by Q_B:

$$Q_B\,|\psi_{\text{phys}}\rangle = 0\,, \tag{1.120}$$

Then we identify all the states which differ by the action of Q_B on a generic state:

$$|\psi_{\text{phys}}\rangle \leftrightarrow |\psi_{\text{phys}}\rangle + Q_B\,|\psi\rangle\,. \tag{1.121}$$

With eq. (1.121) we define a physical state as an equivalence class of states satisfying eq. (1.120): this is the so-called BRST-cohomology.

The theory possesses also an additive symmetry, with a conserved charge Q_c (ghost number). One imposes also

$$Q_c \, |\psi_{\text{phys}}\rangle \;=\; 0\,. \qquad (1.122)$$

This construction is enough to ensure that all the unphysical states of the theory (ghosts and unphysical polarizations) are absent in the (Hilbert) space of asymptotic states. The fact that both Q_B and Q_c are conserved ensures that this unphysical states cannot be produced in the causal evolution of the system.

In presence of anomalies, eq. (1.118) does not hold, so that even if the initial Hilbert space has been constructed in such a way as not to contain unphysical states, transitions from physical to unphysical states are allowed. As a consequence, one produces, for instance, negative norm states. The probabilistic interpretation of the theory is lost, and the scattering matrix S is no more unitary, violating the postulates of quantum field theory. The theory is then meaningless.

There are a number of possible ways to write down consistent gauge theories which are free from local anomalies:

- G can be an anomaly safe group, for which it is not possible to write an invariant out of $1 + D/2$ generators (this is the case of a chiral $SU(2)$ in $D = 4$),
- all of the representations of G are vector-like, so that the contribution to the anomaly due to Left Handed fields cancels with that of their Right Handed partners (as for the $SU(3)_c$ theory of strong interactions),
- the matter field content is chiral, but chosen in such a way as to set to zero the anomaly after summing over all the chiral fields (this is the case of the SM in $D = 4$, where the anomaly due to the leptons cancels exactly with the anomaly due to the quarks),
- one can introduce non-invariant higher order local operators coupling the gauge bosons to new fields, transforming non-linearly under the action of the gauge

transformation, in such a way as to compensate the anomalous variation of the fermionic determinant.

In this thesis we will focus on the phenomenological consequences of the application of some mechanisms belonging to this last class.

An example of such a mechanism was already discussed in [6], in order to explain the connection between anomalies of gauge symmetries, causality and renormalizability.

Consider an Higgs model with symmetry group $G = U(1)$, a chiral field ψ_L and a scalar field φ transforming under a gauge transformation according to

$$\varphi \;\rightarrow\; e^{ig\alpha}\,\varphi\,, \tag{1.123}$$

$$\psi \;\rightarrow\; e^{ig\alpha}\,\psi\,. \tag{1.124}$$

Suppose that the scalar potential has a minimum for

$$\langle\varphi\rangle \;=\; v \neq 0\,, \tag{1.125}$$

and, accordingly, rewrite the Higgs field as

$$\varphi = \rho\, e^{-i\theta}\,. \tag{1.126}$$

The transformation properties of θ are inherited by φ:

$$\theta \;\rightarrow\; \theta - g\alpha\,. \tag{1.127}$$

It is well known that, after gauge fixing in the unitary gauge, θ provides the longitudinal components of the (massive) gauge bosons A_μ.

On the other hand, the theory is anomalous, according to eq. (1.59):

$$\delta\, W[A] \;=\; -g \int \mathrm{d}^4x\, \alpha(x)\, \mathcal{G}[A]\,, \tag{1.128}$$

$$\mathcal{G}[A] \;=\; \frac{g^2}{32\,\pi^2}\,\epsilon_{\tau\sigma\rho\mu}\, F^{\tau\sigma} F^{\rho\mu}\,. \tag{1.129}$$

This anomaly can be canceled adding to the classical action the following (one-loop) term:

$$\mathcal{S}_1 \;\equiv\; -\frac{g^2}{32\,\pi^2} \int \mathrm{d}^4x\, \theta\, \epsilon_{\tau\sigma\rho\mu}\, F^{\tau\sigma} F^{\rho\mu}\,, \tag{1.130}$$

in such a way that

$$\delta W[A] \;=\; -\delta \mathcal{S}_1 \,, \tag{1.131}$$

thus ensuring that the effective action generated by $\mathcal{S}+\mathcal{S}_1$ in gauge invariant, and unitarity is restore.

On the other hand, the operator defined in eq. (1.130) has dimension 5, and this theory can be seen to be non-renormalizable. We will discuss more in details the non-renormalizability in chapter 3, where we will generalize recalling the results of [47], this idea to the case of non-Abelian symmetries. Here, it is enough to recall that the local counter-term we added removes only the (local) variation of the amplitude we called $R_{\mu\nu\rho}$ in section 1.1. This amplitude has a non-local part which cannot be removed without the addition of a non-local operator.

In effective field theories, as well as in gauge theories in $D > 4$ space-time dimensions, renormalizability is not important, and then these class of mechanism can be of interest, as we will see.

1.5 Global gauge anomalies

When a D dimensional theory with gauge group G has a chiral matter field content, the local (chiral) gauge anomaly discussed in the previous sections is not the only possible anomaly causing the inconsistency of the theory. In general, we work on a D-dimensional Minkoski space-time $\mathcal{M}^D$, but one can always reformulate the theory in Euclidean space-time via a Wick rotation, in such a way that the base manifold of the theory is $\mathbf{R}^D$:

$$\mathcal{M}^D \to \mathbf{R}^D \,. \tag{1.132}$$

The boundary conditions are usually chosen is such a way that one can add consistently the point $\{\infty\}$, compactifying the space-time on a sphere (of infinite radius, but clearly this is topologically equivalent to a finite sphere of large radius), so that:

$$\mathbf{R}^D \cup \{\infty\} \sim \mathcal{S}^D \,. \tag{1.133}$$

In other words, the topological properties of the base manifold, after including boundary conditions, are those of the sphere. In particular, the maps (gauge transformations) $U(x)$

$$U : \mathcal{M}^D \to G\,, \tag{1.134}$$

are classified by the D-dimensional homotopy group:

$$\pi_D\,(G)\,. \tag{1.135}$$

In Table 1.2 we collect some of the relevant homotopy groups of the simple classical Lie groups.

Group	π_1	π_2	π_3	π_4	π_5	π_6
$SO(3)$	$\mathbf{Z}_2$	0	$\mathbf{Z}$	$\mathbf{Z}_2$	$\mathbf{Z}_2$	$\mathbf{Z}_{12}$
$SO(4)$	$\mathbf{Z}_2$	0	$\mathbf{Z}+\mathbf{Z}$	$\mathbf{Z}_2+\mathbf{Z}_2$	$\mathbf{Z}_2+\mathbf{Z}_2$	$\mathbf{Z}_{12}+\mathbf{Z}_{12}$
$SO(5)$	$\mathbf{Z}_2$	0	$\mathbf{Z}$	$\mathbf{Z}_2$	$\mathbf{Z}_2$	0
$SO(6)$	$\mathbf{Z}_2$	0	$\mathbf{Z}$	0	$\mathbf{Z}$	0
$SO(n)\,,\ n>6$	$\mathbf{Z}_2$	0	$\mathbf{Z}$	0	0	0
$U(1)$	$\mathbf{Z}$	0	0	0	0	0
$SU(2)$	0	0	$\mathbf{Z}$	$\mathbf{Z}_2$	$\mathbf{Z}_2$	$\mathbf{Z}_{12}$
$SU(3)$	0	0	$\mathbf{Z}$	0	$\mathbf{Z}$	$\mathbf{Z}_6$
$SU(n)\,,\ n>3$	0	0	$\mathbf{Z}$	0	$\mathbf{Z}$	0
G_2	0	0	$\mathbf{Z}$	0	0	$\mathbf{Z}_3$
F_4	0	0	$\mathbf{Z}$	0	0	0
E_6	0	0	$\mathbf{Z}$	0	0	0
E_7	0	0	$\mathbf{Z}$	0	0	0
E_8	0	0	$\mathbf{Z}$	0	0	0

Table 1.2: Homotopy groups of some classical compact Lie groups up to $D = 6$ [10].

When discussing the local anomalies in the previous sections, we always computed the gauge variation under infinitesimal transformations, *i.e.* transformations which differ from the identity transformation by an infinitesimal amount. By combining these transformations one can reach only the possible gauge configurations U that can be continuously deformed to the identity. This means that we know what are the variations of the effective action and what are the anomalous conservation laws restricted just to the trivial homotopy class of G. If $\pi_D(G) \neq 0$, the absence, or the cancellation, of these anomalies is not enough to ensure that the theory is invariant under gauge transformations $\hat{U}$ which are in non-trivial homotopy classes.

It was first discovered by Witten [7] studying the $D = 4$ chiral realization of $SU(2)$, that the effective action may receive a shift by the action of $\hat{U}$, even if no local anomaly of $SU(2)$ is present [4].

Before showing how to compute this anomaly, let us discuss why its presence is dangerous for the formulation of the theory. Just for concreteness, we restrict ourselves to the case discussed by Witten in the original paper.

Starting from the definition of $W[A]$ in eq. (1.21), where now ψ is a chiral (left-handed) field, and from the observation that

$$\pi_4\left(SU(2)\right) = \mathbf{Z}_2\,, \tag{1.138}$$

one can identify a gauge transformation $\hat{U}$ in the non-trivial homotopy class as the generator of the homotopy group. Applying $\hat{U}$, in general the variation of the effective action is:

$$W[A] \to W'[A'] = \Delta(\hat{U})\, W[A]\,. \tag{1.139}$$

Due to the absence of local anomalies, $\Delta(\hat{U})$ depends only on the homotopy class of $\hat{U}$: this means we are effectively studying in which representations of the discrete

[4] No $SU(2)$ anomaly is present in four dimensions, the triangle diagram being proportional to the (vanishing) $D_{a_1a_2a_3}$:

$$D_{a_1a_2a_3} \equiv \mathrm{Tr}\left[\{T_{a_1}, T_{a_2}\}T_{a_3}\right], \tag{1.136}$$
$$T^a = \frac{\tau^a}{2}, \tag{1.137}$$

where τ^a are the hermitian 2×2 Pauli matrices. This is due to the fact that $SU(2)$ admits Casimir operators only of order 2 (see Table 1.1), so that no order 3 singlets can be built.

group $\pi_4(SU(2))$ the fields of the theory are. From the observation that the double iteration of $\hat{U}$ must belong to the trivial homotopy class, we have:

$$\Delta(\hat{U})^2 = 1, \tag{1.140}$$

$$\Delta(\hat{U}) = e^{i\phi}, \tag{1.141}$$

with $\phi = 0$ or $\phi = \pi$. In the first case, no anomaly is present, and the theory is consistent. In the second case, for gauge invariance to be preserved, one needs:

$$W[A] = W'[A'] = -W[A] = 0. \tag{1.142}$$

This means that the partition function, the generator functional of all the connected correlation functions of the theory, must vanish, with all the correlation functions themselves, and the theory becomes trivial.

The source of the problem is the fact that, even after the introduction of the ghost fields for the gauge fixing procedure, there is a multiple counting in the path integral when integrating over the gauge fields A_μ, because for every configuration A_μ there are equivalent configurations A_μ^n obtained by applying the representative elements of each homotopy class. The multiple counting cannot be avoided, because one can deform continuously these A_μ^n configurations to reach A_μ.

In the canonical quantization, this can be understood from the fact that the representation of the Lie algebra of G does not provide a representation of the whole Lie group G, but there is a non-trivial center. Under the action of the non-trivial operators of the center the fields are not invariant, so that the space of physical states of the theory is empty.

We briefly get back to the proof given by Witten in [7] of the fact that for $SU(2)$ in $D = 4$ with one chiral spinor on the fundamental representation one has $\phi = \pi$. The proof is based on the use of a particular version of the Atiyah-Singer index theorem.

For a Dirac spinor representation of $SU(2)$ one defines the determinant of the Dirac operator according to:

$$\det i\not{D} = \int (\mathcal{D}\bar{\psi}\mathcal{D}\psi)_{\text{Dirac}}\, e^{\bar{\psi} i \not{D} \psi}. \tag{1.143}$$

Then, for Weyl fermions, one gets exactly the square root of this result, up to an ambiguity in the definition of the sign:

$$\pm \left[\det i\not{D}\right]^{1/2} = \int (\mathcal{D}\bar{\psi}\mathcal{D}\psi)_{\text{Weyl}} \; e^{\bar{\psi}\, i\not{D}\,\psi} . \tag{1.144}$$

This is consistent with the fact that the eigenvalues of the Dirac operator, defined now on a large (finite radius) sphere $\mathcal{S}^D$, come always in (real) pairs $(\lambda_i, -\lambda_i)$:

$$\begin{cases} i\not{D}\,\psi_i &= \lambda_i\,\psi_i \\ i\not{D}\,\gamma_5\,\psi_i &= -\lambda_i\,\psi_i \end{cases} . \tag{1.145}$$

If we choose to fix the field A_μ, and define the Weyl determinant as the product of all the positive eigenvalues of $i\not{D}(A_\mu)$, after applying a gauge transformation $\hat{U}$ homotopically non trivial, one finds that

$$\left[\det i\not{D}(A_\mu)\right]^{1/2} = -\left[\det i\not{D}(A_\mu^{\hat{U}})\right]^{1/2} . \tag{1.146}$$

This can be seen building a 1-parameter family of gauge configurations A_μ^t, with $t \in [0,1]$, continuously transforming $A_\mu^0 = A_\mu$ onto $A_\mu^1 = A_\mu^{\hat{U}}$, and applying the mod two Atiyah-Singer theorem to the five dimensional cylinder, which implies that along the path from A_μ to $A_\mu^{\hat{U}}$ an odd number of eigenvalues $(\lambda_i, -\lambda_i)$ are interchanged (see Fig. 1.3), thus changing the sign of the square root [17]. This means there is no consistent way to define this sign, *i.e.* $\phi = \pi$.

In chapter 3 we will discuss a simpler and more general technique for the computation of these anomalies, suggested by an observation of Witten in [17] in discussing the WZ action, and elaborated by Elitzur and Nair in [18]. This technique is based on the observation that global anomalies can be computed starting from local anomalies of a (larger) homotopically trivial group in which the group G is embedded.

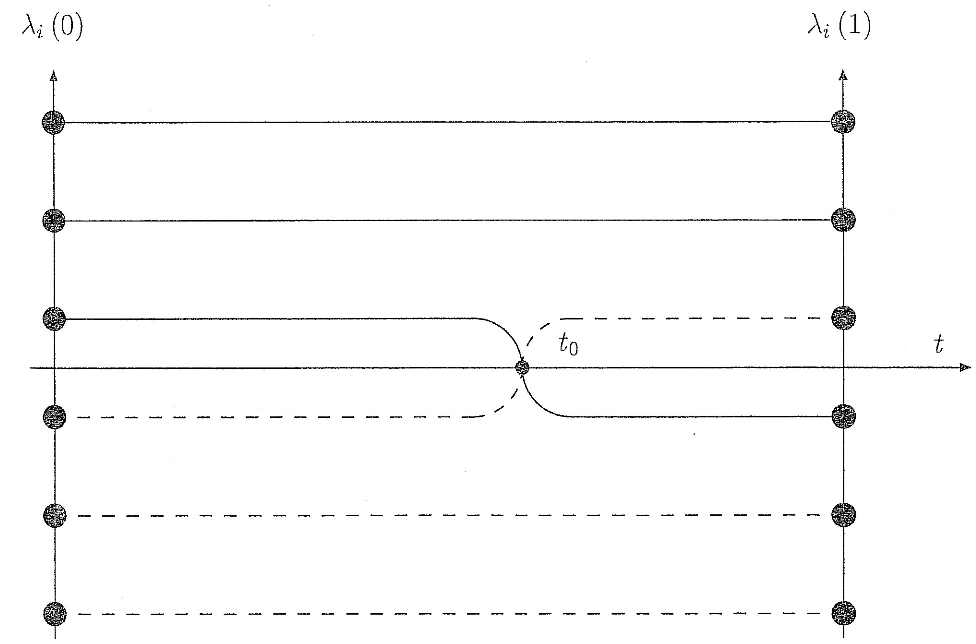

Figure 1.3: Diagrammatic representation of the flow of the eigenvalues of the Dirac operator as a function of A^t. The square root of the determinant cannot have a definite sign: after defining it as the product of eigenvalues on solid lines, the application of a topologically non-trivial gauge transformation $\hat{U}$ which transforms A^0 to A^1 is equivalent to following a continuous deformation from A^0 to A^1, which passes through t_0, where the square root vanishes and changes sign.

Chapter 2

The Green-Schwarz mechanism

The first mechanism for anomaly cancellation we wish to discuss is the GS mechanism. It has been discovered in the context of Superstring theories [19], showing that the Type I (open) Superstring based on the gauge group $SO(32)$ has an anomaly free low-energy effective action in $D = 10$ dimensions.

We show how this mechanism works in general in canceling all the possible reducible (local) gauge anomalies in the context of low energy effective field theories [20]. On the other hand, the mechanism requires the presence in the spectrum of bosonic fields with non-trivial properties under gauge (and gravitational) transformations. Some light components of these fields remain, in general, in the spectrum: these (pseudo-scalar) fields have axion-like couplings that can be relevant for phenomenology. For this reason we briefly review the physics of the axion, and recall the main features of the strong-CP problem, the Peccei-Quinn (PQ) solution [21] and the experimental status of the searches for the PQ axion.

We then discuss the phenomenology of a concrete proposal [22] for the (non-Supersymmetric) extension to $D = 6$ dimensions of the SM, in which the requirement of the cancellation of global and irreducible local gauge (and gravitational) anomalies is enough to fix the matter field content, giving a prediction for the existence of $N_g = 3$ matter families. The application of the GS mechanism in order to cancel reducible local gauge anomalies produces a set of axion-like fields in the $D = 4$ effective theory. We identify one of the combinations of this fields with the PQ axion (the other components are shown to decouple from gauge fields), so that in this model we find a natural solution to the strong-CP problem. Further, we find an upper bound on the volume of the two compact extra-dimensions, deduced from the experimental bounds on the mass and coupling of the axion [23].

2.1 The Green-Schwarz mechanism

Non-Abelian anomalies in $D = 2n$ dimensions (we consider here a simple Lie group G and neglect for the moment gravity) can always be computed, using the descent equations, from the $(2n+2)$-form representing the Abelian anomaly in $D+2$ dimensions. If the gauge bosons of the theory transform accordingly to the adjoint representation of the group G, in general one finds that the starting point of the computation is given by the terms in Table 2.1

n	Non-factorizable	Factorizable
1	$a_2 \,\mathrm{Tr}\, F^2$	–
2	$a_3 \,\mathrm{Tr}\, F^3$	–
3	$a_4 \,\mathrm{Tr}\, F^4$	$b_{22} \,\mathrm{Tr}\, F^2 \,\mathrm{Tr}\, F^2$
4	$a_5 \,\mathrm{Tr}\, F^5$	$b_{23} \,\mathrm{Tr}\, F^2 \,\mathrm{Tr}\, F^2$
5	$a_6 \,\mathrm{Tr}\, F^6$	$\left\{ \begin{array}{c} b_{24} \,\mathrm{Tr}\, F^2 \,\mathrm{Tr}\, F^4 \\ b_{222} \,\mathrm{Tr}\, F^2 \,\mathrm{Tr}\, F^2 \,\mathrm{Tr}\, F^2 \end{array} \right.$

Table 2.1: $(2n+2)$-forms, starting point for the computation of non-abelian anomalies in $D = 2n$ dimensions, for a simple Lie group G. The coefficients a_i and b_j depend on the group and on the fermionic representations, F are the field strength forms, and $\wedge$ products are understood.

For non-factorizable anomalies, as discussed at length in the previous chapter, one deduces the form of eq. (1.106):

$$G_{\mathrm{nf}}[v, A] \;=\; -\int a_{n+1}\omega^1_{2n} \,. \tag{2.1}$$

In the factorizable case, the descent equations have to be applied to each of the factors, so that the (consistent) non-abelian anomaly reads as the sum of all the possible terms obtained from the third column of Table 2.1 replacing a $\mathrm{Tr}\, F^{m+1}$ with the Chern-Simons form ω^1_{2m}:

$$G_{\mathrm{f}}[v, A] \;=\; -\int \sum_{n=m+l+\cdots+k} b_{2m\, l \cdot k}\, \omega^1_{2m}\, \mathrm{Tr}\, F^l \cdots \mathrm{Tr}\, F^k \,. \qquad (2.2)$$

We add to the spectrum of the theory, besides the gauge bosons and the fermions responsible for the anomaly, $(2m)$-forms $B^{(2m)}$ which do not carry indexes of a representation of G (and with ghost number 0), but are not invariant under the action of BRST transformations, transforming according to:

$$s\, B^{(2m)} \;=\; -\omega^1_{2m} \,. \qquad (2.3)$$

If one adds to the classical action the couplings:

$$S^{GS} \;=\; -\int \sum_{n=m+l+\cdots+k} b_{2m\, l \cdots k}\, B^{(2m)}\, \mathrm{Tr}\, F^l \cdots \mathrm{Tr}\, F^k \,, \qquad (2.4)$$

using the invariance property $s\, \mathrm{Tr}\, F^l \;=\; 0$, one sees that these coupling cancel exactly the factorizable anomalies of eq. (2.2). There is actually a restriction on this procedure, due to the massless representations of the Lorentz group allowed: a GS field exists only if

$$2m \;\leq\; n - 1 \,. \qquad (2.5)$$

The GS mechanism cannot be applied to the case of non-factorizable anomalies. In the case when G is not simple, but contains some abelian factor, the GS mechanism generalizes with the introduction of neutral scalars (0-forms) C transforming according to:

$$s\, C \;=\; -v \,. \qquad (2.6)$$

We call *reducible anomalies* all those anomalies which can be cancelled with the GS mechanism. Non-factorizable anomalies, and factorizable ones for which one of the factors corresponds to a $(2m)$-form with $2m > n$ (so that no GS field exists such as to cancel them) are called *irreducible anomalies.*

When also gravity is included, all of this generalizes easily: mixed (gauge-gravitational) anomalies can be compensated by generalizing eq. (2.3) and eq. (2.4) with the introduction of properly defined gravitational Chern-Simons forms (the technology of descent equations, with the definition of the Chern-Simons forms, can be generalized to gravitational anomalies, with some subtleties that are not crucial in our context, replacing in the first step the field strength 2-forms F with the form representing the Ricci tensor $\mathcal{R}$).

The result of this procedure is that, given a chiral (gauge) theory in $D = 2n$ dimensions, absence of local (gauge) anomalies can be obtained by first choosing appropriately the group G and the chiral fields in such a way that no irreducible anomalies are present (the a_{n+1} symbols must vanish, for instance), and then by introducing a certain number of GS fields with the transformation properties of eq. (2.3), and couplings in the form of eq. (2.4), in such a way as to cancel all the remaining reducible anomalies.

Finally, one needs to write down appropriate gauge invariant kinetic terms for the GS filds $B^{(2m)}$. This can be achieved defining the k-forms ($k = 2m + 1$):

$$H \equiv \mathrm{d}\, B^{(2m)} - \omega^0_{2m+1}\,. \tag{2.7}$$

In such a way $s\,H = 0$ (recall that $s\,\omega^0_{2m+1} = \mathrm{d}\omega^1_{2m}$), and the kinetic term for these k-forms is (up to normalization factors):

$$\mathcal{S}^{kin} \simeq -\int H_{\mu_1\cdots\mu_k}\, H^{\mu_1\cdots\mu_k}\,. \tag{2.8}$$

The invariant term $\mathcal{S}^{kin}$ gives rise to a coupling of the $B^{(2m)}$ field to $2m$ gauge bosons, while in the gauge non-invariant action $\mathcal{S}^{GS}$, $B^{(2m)}$ couples to $(n+1-2m)$ gauge bosons, so that the combination of these couplings in a diagram with the exchange of $B^{(2m)}$ in the internal lines cancels exactly the anomalous diagrams, as illustrated in Fig. 2.1

In the case of Superstring theories, Green and Schwarz observed [20] that the $N = 1$ Supergravity theory in $D = 10$ with gauge group $SO(32)$ is free from gravitational

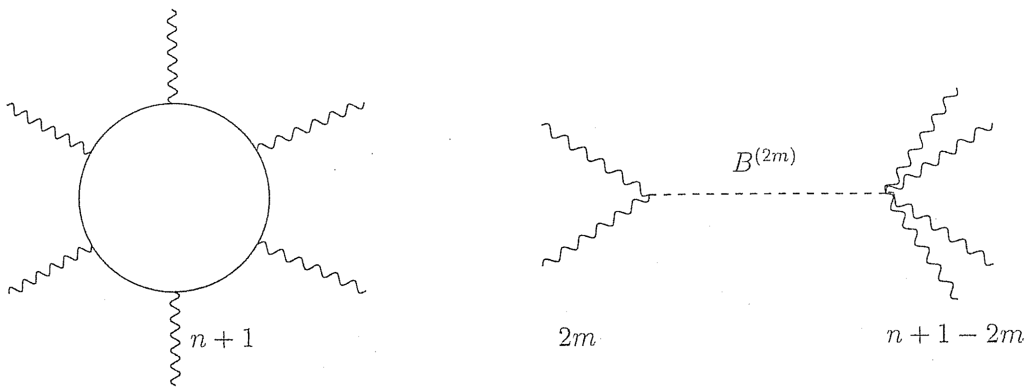

Figure 2.1: Cancellation of the diagram on the left giving rise to the anomaly in $D = 2n$ (even) dimensions ($n + 1$ external gauuge bosons) with the GS mechanism represented by the diagram on the right, containing the two couplings coming from the kinetic term of the GS field ($2m$ external gauge bosons) and from $\mathcal{S}^{GS}$ ($n+1-2m$ external gauge bosons).

as well as non-factorizable anomalies, and the only factorizable anomaly is that in the form

$$b_{24} \operatorname{Tr} F^2 \operatorname{Tr} F^4 . \tag{2.9}$$

Since

$$\operatorname{Tr} F^2 \operatorname{Tr} F^4 \;=\; \mathrm{d}(\omega_3^1 \operatorname{Tr} F^4) \;=\; \mathrm{d}(\omega_7^1 \operatorname{Tr} F^2) , \tag{2.10}$$

this anomaly is reducible, and can be cancelled by the order-2 antisymmetric tensor B_{MN} of the Supergravity multiplet, according to eq. (2.4) and eq. (2.3). This lead to the discovery of the Type I $SO(32)$ Superstring Theory.

All the GS fields in what follows are 2-forms B_{MN}, so that we from now on will refer to 2-forms as GS fields.

An important point, which will prove crucial in phenomenology, concerns compactification: what are the fields that appear in the low energy $D = 4$ effective action if in the original $D \geq 4$ theory the GS mechanism is at work, and what are their couplings? We are especially interested in the zero-modes of what will become the Kaluza-Klein decomposition of the fields. Considering the space-time to be $\mathcal{M}^4 \times \mathcal{K}$, with $\mathcal{K}$ some $D-4$ dimensional compact manifold, the number of zero modes of a p-form is controlled by the Betti numbers $b_p(\mathcal{K})$ [1]. In particular, for what massless modes are concerned, a 2-form in D dimensions decomposes always in a 2-form in $D=4$ (by restricting the indexes to $0 \cdots 3$) which is equivalent to a scalar. It can be seen [25] (see also [24]) that such a scalar inherits from eq. (2.7) an axionic coupling to gauge bosons, so that it is usually referred to as the *model-independent axion*. Out of the other $D=4$ remnants of B, the number of (massless) scalars is equal to $b_2(\mathcal{K})$: these are called *model-dependent axions*, their existence depending specifically on the topology of the compactification manifold $\mathcal{K}$.

Of particular interest is the application of the GS mechanism in the $D=4$ case [26]: the only GS fields allowed are 0-forms (pseudoscalar fields) θ, so that the mechanism applies only to abelian ($U(1)$) symmetries. Then eq. (2.3), eq. (2.4), eq. (2.7) and eq. (2.8) become respectively (we drop the explicit indication of the correct normalizations, in order to simplify the notation):

$$s\,\theta = -v\,, \tag{2.11}$$

$$\mathcal{S}^{GS} \propto -\int \theta \mathrm{Tr}\, F^2\,, \tag{2.12}$$

$$H \equiv \mathrm{d}\theta + A\,, \tag{2.13}$$

$$\mathcal{S}^{kin} \simeq -\int H_\mu H^\mu\,. \tag{2.14}$$

Under a gauge transformation of parameter α, θ transforms as

$$\delta_\alpha \theta = -\alpha\,, \tag{2.15}$$

so that it is always possible to choose the gauge fixing conditions in such a way as to set $\theta = 0$. In this case $H_\mu = A_\mu$, the gauge boson of the underlying symmetry, so that $\mathcal{S}^{kin}$ provides a mass term for the gauge boson, and the symmetry is non-linearly realized: the GS field provides the longitudinal component of the gauge boson of $U(1)$, which becomes heavy, so that the original gauge symmetry is not

[1] The pth Betti number of a manifold M is defined as the dimension of the pth cohomology group on M with real coefficients, see for instance [10].

manifest in the low energy spetrum. On the other hand the symmetry survives as an unbroken global symmetry of the theory.

In Superstring models, several $U(1)$ factors survive as gauge symmetries of the $D = 4$ effective theory, and most of them have anomalous chiral field content. In all these cases, some pseudo-scalar must be present for the GS mechanism to be at work, and the result is the decoupling of the (massive) gauge bosons of these abelian factors from the low-energy spectrum. On the other hand, some of the $U(1)$ symmetries may be identified with phenomenologically relevant symmetries (the barion number, for instance), and might explain the suppressed rate of some rare processes.

2.2 The strong-CP problem and the axion

Consider a non-Abelian gauge theory of group G with Yang-Mills $D = 4$ action:

$$\mathcal{L} = -\frac{1}{2\, g^2}\, \mathrm{Tr}\, F^{\mu\nu}\, F_{\mu\nu}\,. \tag{2.16}$$

As can be deduced from Table 1.2, or applying the Bott's theorem, for which all the maps from S^3 to G can be continously deformed onto maps to an $SU(2)$ subgroup of G, the homotopy group for non-Abelian $G \supseteq SU(2)$ is:

$$\pi_3\,(G) = \pi_3\,(SU(2)) = \mathbf{Z}\,. \tag{2.17}$$

Therefore, there exist istantonic (classical) solutions of the theory. Let us denote with $|n\rangle$, $(n \in Z)$, the istantonic solution with Pontryagin index $\mathcal{Q} = n$, where:

$$\mathcal{Q} = \frac{1}{16\,\pi^2} \int \mathrm{d}^4x\, \mathrm{Tr}\, F^{\mu\nu}\, \tilde{F}_{\mu\nu}\,, \tag{2.18}$$

and with G_n $(n \in \mathbf{Z})$ the (unitary) gauge transformations representative for the homotopy class n, such that:

$$G_n\,|m\rangle = |m+n\rangle\,. \tag{2.19}$$

The vacuum $|\theta\rangle$ can be defined as a superposition of instantons $|n\rangle$, $(n \in Z)$, with all possible values of the Pontryagin index $\mathcal{Q}$:

$$|\theta\rangle \equiv \sum_n e^{i\,n\,\theta}\,|n\rangle\,, \tag{2.20}$$

where the parameter θ is defined by:

$$G_1 \, |\theta\rangle \equiv e^{i\,\theta} \, |\theta\rangle \, . \tag{2.21}$$

It can be shown that the transition amplitude between all the possible vacua labelled by θ can be written as:

$$\langle \theta' \, \mathrm{out} | \theta \, \mathrm{in} \rangle = \delta(\theta - \theta') \int \mathcal{D}A_\mu \, \exp\left[- \int \mathrm{d}^4_e x (\mathcal{L} + \mathcal{L}_\theta) \right] , \tag{2.22}$$

where

$$\mathcal{L}_\theta = \theta \frac{\alpha_s}{8\,\pi} \, \mathrm{Tr} \, F^{\mu\nu} \, \tilde{F}_{\mu\nu} \, , \tag{2.23}$$

so that θ is a physical parameter of the theory, *i.e.* the choice of the vacuum has a physical meaning. The fact that $\mathcal{L}_\theta$ is CP-odd leads to the so calles *strong-CP problem*: instanton effects induce flavour blind CP-violating couplings in a non-abelian gauge theory, controlled by the real parameter θ. CP is restored as a good symmetry of the theory only in the limit $\theta \to 0$.

This means that the complete QCD Lagrangian, the theory of strong interactions based on the gauge group $SU(3)_c$, also contains the CP-odd coupling:

$$\bar{\theta} \frac{\alpha_s}{8\pi} \, \mathrm{Tr} \, F^{\mu\nu} \tilde{F}_{\mu\nu} \, , \tag{2.24}$$

where $\bar{\theta}$ includes also the effect of a CP-odd phase coming from the quark masses. These coupling is responsible for the appearance of a non-vanishing electric dipole moment (EDM) of nucleons and nuclei. This are known to be small, so that from the neutron EDM one can deduce the experimental bound:

$$\bar{\theta} < 10^{-14} \, . \tag{2.25}$$

This severe fine-tuning problem received a number of possible solutions. For instance if one of the quark fields is massles, then the determinant of the product of the mass matrices of up and down quarks in $\bar{\theta}$ would vanish, and its phase would become a completly arbitrary, unphysical parameter, which could then be choosen in such a way to set $\bar{\theta} = 0$. Unfortunately, extimates of the masses of the lightest quarks, based on chiral theories, seem to exclude this possibility [27].

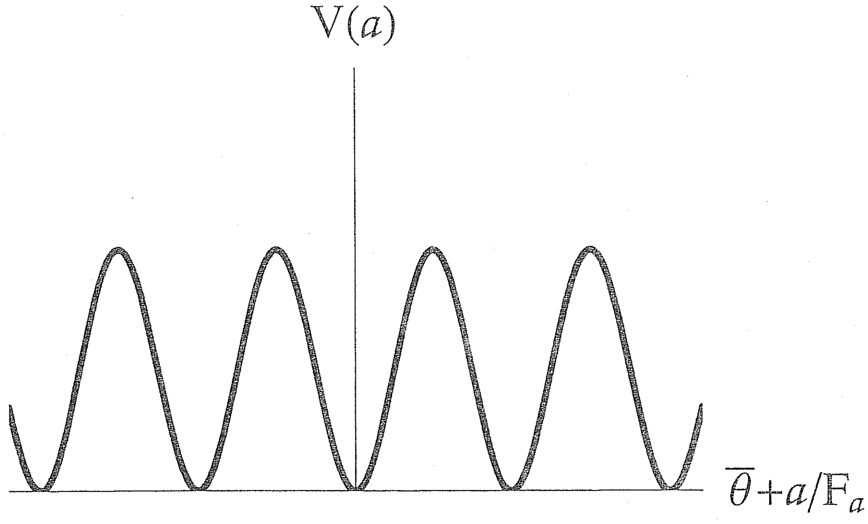

Figure 2.2: The effective potential of the axion field $V(a)$.

A more interesting dynamical solution has been proposed by Peccei and Quinn, and requires the existence of the so called axion field [21, 28, 29]. Let us introduce a pseudo-scalar, neutral field a, the Lagrangian of which is invariant under a translation $a \to a + k$ (this means there is no allowed tree level potential for a). This invariance is however violated by the axionic coupling to the $SU(3)_c$ field strength:

$$\frac{a}{F_a}\frac{\alpha_s}{8\pi}\,\mathrm{Tr}\,F^{\mu\nu}\tilde{F}_{\mu\nu}\,. \tag{2.26}$$

The effective potential, after functional integration over A_μ and fermions, is the periodic (non-flat) function in Fig. 2.2, with a minimum at [30, 31]:

$$\frac{\langle a\rangle}{F_a} + \bar{\theta} = 0 \bmod 2\pi\,. \tag{2.27}$$

In this way the effective CP-violating phase is dinamically set to zero.

In spite of the elegance of this solution, some problems arise because the axion field is a true physical degree of freedom, which has not been observed at the moment neither in direct searches (from accelerator physics) nor from indirect signals (from the study of supernovae explosion, red giant cooling and cosmology).

To discuss the bounds which can be derived on the axion couplings, one needs to build the low energy effective theory of electromagnetic and strong interactions, containing photons, nucleons, pions and axions only. The main couplings which are bounded from these studies are those to the electromagnetic field strength:

$$\mathcal{L}_{a\gamma\gamma} = -\frac{1}{4} g_{a\gamma}\, a\, F^{\mu\nu}\tilde{F}_{\mu\nu}\,, \tag{2.28}$$

and to the nucleons:

$$\mathcal{L}_{aNN} = -i g_{aN}\, a\, \bar{N}\gamma_5 N\,. \tag{2.29}$$

These couplings have the bounds [32, 33]

$$g_{a\gamma} < 10^{-10}\ \mathrm{GeV}^{-1}\ , \tag{2.30}$$

which is deduced by the study of the cooling of red-giants, and

$$g_{aN} < 3\times 10^{-10}\ , \tag{2.31}$$

computed by studying the explosion of the supernova SN1987A.

2.3 A $D = 6$ Standard Model

As an axample of the constraints deduced from the cancellation of irreducible and global anomalies, and of the phenomenological consequences of the application of the GS mechanism to reducible anomalies, we introduce the $D = 6$ extension of the SM proposed by Dobrescu and Poppitz in [22]. We will further discuss this example in the next Chapter, in order to show the difference between two of the most studied anomaly cancellation mechanisms in effective field theories.

2.3.1 Field content: anomaly cancellation

Let us write a non-supersymmetric [2] six-dimensional theory, based on the SM gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$; matter fields are assigned to chiral fermions (projected by $(1\pm\Gamma_7)/2$) in the usual representations of the gauge group Q, L, U, D and

[2] The procedure outlined below leads to the impossibility of cancelling simultaneously all the local and global anomalies in presence of Supersymmetry [22]: this is due to the fact that $(1,0)$ Supersymmetry forces to choose the same chirality for all matter fields, opposite to the one of gauginos, imposing quite severe restrictions on the allowed number of representations of the gauge group.

E, as shown in Tab. 2.2. Quantum numbers are fixed by the SM phenomenology, the only freedom is in the (six dimensional) chirality assignements. A scalar doublet h must be present for the Higgs mechanism to take place.

As explained in [22], cancellation of purely gravitational and irreducible gauge anomalies forces quark singlets to have opposite chirality with respect to doublets, and to introduce a SM singlet N in order to have the same number of fermions of both chiralities.

In fact if we consider the factorizable anomaly:

$$U(1)_Y \ [SU(3)_c] \ , \tag{2.32}$$

the $SU(3)$ anomaly deduced is proportional to the 4-form ω_4^1, while we are allowed to introduce only 2-forms as GS fields. This anomaly can be cancelled only if the matter field content is such to form a vector-like representation of $SU(3)$, *i.e.* choosing quark singlets to have opposite chirality with respect to doublets.

All the other gauge anomalies are reducible, we apply the GS mechanism to cancel them, and so no other constraints on the matter field content arise.

In $D = 6$ dimensions (and in general in $2n + 2$ dimensions) chiral fermions cause the presence of (pure) gravitational anomalies besides the mixed anomalies present also in $D = 4$. The only way to cancel gravitation anomalies in a non-supersymmetric theory is to have an equal number of fermions of the two chiralities [3] While quarks do not contribute (because we are forced to choose an equal number of chirality $+$ and chirality $-$ quarks), we need also to require that the field L and E have opposite chiralities, and also we need to add to the spectrum another (gauge) singlet N with the same chirality of E.

We are left with four possible combinations of the allowed chiralities. Our (conventional) choice is to assign positive chirality to doublets and negative to singlets. An alternative choice would be to assign opposite chiralities to leptons and hadrons: in what follows this would turn in a change of $\mathcal{O}(1)$ in the couplings introduced to

[3] In $D = 6$ the requirement of gravitational anomaly cancellation is quite restrictive: it requires 21 chiral fermions, 1 gravitino field and 8 self-adjoint antisymmetric tensor fields [15].

cancel anomalies, with minor modifications in the phenomenological consequences of the model.

Table 2.2: Fermionic field content for each family. The six-dimensional chirality is the eigenvalue of Γ_7.

	Chirality	$U(1)_Y$	$SU(2)_L$	$SU(3)_c$
Q	$+$	$1/6$	2	3
U	$-$	$2/3$	1	3
D	$-$	$-1/3$	1	3
L	$+$	$-1/2$	2	1
E	$-$	-1	1	1
N	$-$	0	1	1

Now, let as consider global anomalies. Dealing with only vector-like representations, the possible global anomalies of the $SU(3)_c$ and Lorentz group are automatically absent. For $SU(2)$, one has (see Table 1.2):

$$\pi_6\,(SU(2)) = \mathbf{Z}_{12}\,, \tag{2.33}$$

so that global anomalies are possible. The computation of the contribution to these anomalies of chiral fields (even in presence of GS fields) can be found in [34], and leads to the requirement that:

$$N(2_+) - N(2_-) = 0 \bmod 6, \tag{2.34}$$

where $N_\pm$ is the number of doublets with positive (negative) chirality.

This requirement is not fulfilled by the matter field content of Table 2.2, but can be obtained with $n_g = 0$ mod 3 copies of this matter content. In particular $n_g = 3$ is in agreement with our experimental knowledge [35]. This could be a theoretical explanation of the existence of three families of matter fields with the same quantum numbers.

We are thus left with Abelian and non-Abelian reducible gauge anomalies, that would spoil unitarity unless canceled by the GS mechanism. Accordingly, to recover non-Abelian gauge symmetries we must introduce two real antisymmetric tensors B^L_{MN} and B^c_{MN} whose couplings are tuned to exactly cancel the anomalous terms in the gauge transformations. The $B^{L,c}_{MN}$ fields transform in a non-standard way under gauge transformation:

$$\begin{aligned} \delta_L B^L &\equiv -\tfrac{1}{\Delta_L^2}\omega_2^L = -\tfrac{1}{\Delta_L^2}\mathrm{Tr}\,\{\beta\, \mathrm{d}W\}\,, \\ \delta_c B^c &\equiv -\tfrac{1}{\Delta_c^2}\omega_2^c = -\tfrac{1}{\Delta_c^2}\mathrm{Tr}\,\{\gamma\, \mathrm{d}G\}\,. \end{aligned} \tag{2.35}$$

In eq. (2.35), β and γ are the local parameters of $SU(2)_L$ and $SU(3)_c$ gauge transformations, while Δ_L and Δ_c are free mass parameters.

The six-dimensional anomaly free Lagrangian can be written as:

$$\mathcal{L}^{6D} = \mathcal{L}^{6D}_{SM} + \mathcal{L}^{6D}_{GS}\,, \tag{2.36}$$

where

$$\begin{aligned} \mathcal{L}^{6D}_{SM} &= \sum \bar{\psi} i\gamma^M D_M \psi - \frac{1}{4}F_{MN}F^{MN} \\ &\quad -\frac{1}{2}\mathrm{Tr}\, L_{MN}L^{MN} - \frac{1}{2}\mathrm{Tr}\, V_{MN}V^{MN} \\ &\quad +(D_M h)^\dagger D^M h \; + V(\phi^\dagger\phi) \\ &\quad +\Big(Y_l\, R\, \bar{L} h E + Y_d\, R\, \bar{Q} h D \\ &\quad +Y_u\, R\, \bar{Q}\tilde{h} U + Y_\nu\, R\, \bar{L}\tilde{h} N \; + \; \mathrm{h.c.}\Big) \end{aligned} \tag{2.37}$$

and

$$\begin{aligned} \mathcal{L}^{6D}_{GS} &= \frac{g'^{\,3}R^3}{16\pi^3}\sigma F_{MN}F_{RS}F_{PQ}\epsilon^{MNRSPQ} \\ &\quad +\frac{g^4R^4\Delta_L^2}{6\pi^3}B^L_{MN}\mathrm{Tr}\,\{L_{RS}L_{PQ}\}\epsilon^{MNRSPQ} \\ &\quad +\frac{g^2g'^{\,2}R^4\Delta_L^2}{144\pi^3}B^L_{MN}F_{RS}F_{PQ}\epsilon^{MNRSPQ} \\ &\quad -\frac{g^2g'\,R^3}{72\pi^3}\sigma F_{MN}\mathrm{Tr}\,\{L_{RS}L_{PQ}\}\epsilon^{MNRSPQ} \\ &\quad +\frac{g_s^2g'\,R^3}{48\pi^3}\sigma F_{MN}\mathrm{Tr}\,\{V_{RS}V_{PQ}\}\epsilon^{MNRSPQ} \\ &\quad -\frac{g_s^2g'^{\,2}R^4\Delta_c^{\,2}}{96\pi^3}B^c_{MN}F_{RS}F_{PQ}\epsilon^{MNRSPQ} \\ &\quad +\frac{g_s^2g^2R^4\Delta_L^2}{48\pi^3}B^L_{MN}\mathrm{Tr}\,\{V_{RS}V_{PQ}\}\epsilon^{MNRSPQ} \end{aligned} \tag{2.38}$$

$$+\frac{g_s^2 g^2 R^4 \Delta_c^2}{48\pi^3} B^c_{MN} \mathrm{Tr}\,\{L_{RS} L_{PQ}\} \epsilon^{MNRSPQ}$$
$$+\frac{1}{12} H^{L,c}_{MNS} H^{L,c\,MNS} .$$

In eq. (2.37), ψ is a generic fermionic chiral fields, D_M the covariant derivative on the associated gauge representation, V_{MN}, L_{MN} and F_{MN} are the field strength tensors of the gauge bosons G_M, W_M and A_M of $SU(3)_c$, $SU(2)_L$ and $U(1)_Y$ respectively. $\mathcal{L}^{6D}_{SM}$ is the usual SM Lagrangian, with Lorentz indexes in six dimensions, to which we are allowed to add a Yukawa interaction also for neutrinos ($\tilde{h} = i\sigma_2 h^*$). The scalar potential V is a power series in $\phi^\dagger\phi$. The coefficients of the GS terms in eq. (2.38) match the one-loop anomalous terms, computed for six space-time dimensions in [36].

The Lagrangian in eq. (2.38) contains the (gauge non-invariant) terms required for the cancellation of all reducible gauge anomalies, and the kinetic terms for the 2-forms $B^{L,c}_{MN}$, with:

$$H^{L,c} \equiv \mathrm{d}B^{L,c} - \frac{1}{\Delta^2_{L,c}} \omega_3^{L,c} , \tag{2.39}$$

and

$$\begin{aligned} \omega_3^c &= \mathrm{Tr}\,\{G \wedge V - \tfrac{1}{3} g_s R\, G \wedge G \wedge G\} , \\ \omega_3^L &= \mathrm{Tr}\,\{W \wedge L - \tfrac{1}{3} g R\, W \wedge W \wedge W\} . \end{aligned} \tag{2.40}$$

The Chern-Simons forms $\omega_3^{c,L}$ are needed to make $H^{L,c}_{MNS}$ invariant, and satisfy the relations $\delta_{L,c}\omega_3^{L,c} = -\mathrm{d}\omega_2^{L,c}$.

The presence of the scalar field h can be used to cancel the $U(1)$ anomalies. The Higgs field has been decomposed in a doublet ϕ with vanishing hypercharge and a $SU(2)_L$ singlet σ writing:

$$h = \phi\, e^{i\sigma} . \tag{2.41}$$

Under a $U(1)_Y$ gauge transformation $\delta\sigma = \frac{1}{2} g' R \alpha$, being α the parameter of the transformation. As already discussed, in the unitary gauge $\sigma = 0$, the gauge bosons acquire mass in the standard way and terms proportional to σ in eq. (2.38) vanish.

We have not written explicitly the terms that are needed to cancel mixed (gauge-gravitational) anomalies, because they are not relevant for the phenomenological discussion we are interested in. To achieve the cancellation it is enough to add for each gauge group couplings of the form

$$\omega_3\,\Omega_3 + B\,\mathrm{Tr}\,\mathcal{R}\wedge\mathcal{R}\,, \tag{2.42}$$

with appropriate coefficients, where B are the antisymmetric tensors B^L, B^c and $\sigma\,F$, Ω_3 is the gravitational Chern-Simon form defined by $\mathrm{d}\Omega_3 = \mathrm{Tr}\,\mathcal{R}\wedge\mathcal{R}$. $\mathcal{R}$ is the Ricci tensor. For details see [37].

We denote by R^2 the volume of the compact extra-dimensions, so that the Newton constant is related to the fundamental scale M_f of the theory by

$$M_{Pl} = RM_f^2\,. \tag{2.43}$$

By writing the dimensionful couplings as gR yields, after dimensional reduction, the (four dimensional) gauge couplings g_s, g and g'. The only remaining parameters in the Lagrangian are the mass parameters M_N, Δ^L, Δ^c and the couplings of the scalar potential.

2.3.2 From the GS mechanism to the axion

The two extra dimensions are assumed to be compact, and the underlying geometry flat. Chiral fermions in six dimensions correspond to Dirac fermions in four dimensions, but chirality is recovered by orbifold projection. We assume space-time [4] to be

$$\mathcal{M}_4 \times \frac{S^1\times S^1}{Z_2}\,, \tag{2.44}$$

the product of four-dimensional Minkowski and a torus with orbifold Z_2 which impose a symmetry under the parity transformation

$$Z_2 : (y,z) \to (-y,-z)\,, \tag{2.45}$$

where (y,z) are the coordinates on the torus $S^1\times S^1$.

[4] Every smooth 2-dimensional manifold $\mathcal{K}$ has $b_2(\mathcal{K}) = 1$, so that the main discussion of this Section is not affected by the choice of toroidal compactification.

In what follows, we assume that we are allowed to work in the limit of dimensional reduction, in which the low-energy Lagrangian contains only the zero modes of the fields—while higher modes decouple because of their large masses, proportional to $2\pi/R$. The effects of this simplifying assumption should be checked at the end for consistency to make sure that the large number of these heavier states do not enhance potentially dangerous operators.

A consistent assignment of Z_2-parities makes it possible to have a single massless chiral field out of each ψ; the projection gives a factor 1/2 in the GS gauge non-invariant terms in eq. (2.38). The reduced Lagrangian also contains the zero modes of h and of the gauge bosons, together with two anti-symmetric tensors $B^{L,c}_{\mu\nu}$, and two pseudo-scalars $b^{L,c}$, coming, respectively, from the $\{0123\}$ and $\{56\}$ sectors, of the decomposed tensors

$$B^{L,c}_{MN} \to b^{L,c} \equiv \sqrt{3}/6\, \epsilon^{\hat{M}\hat{N}}\, B^{L,c}_{\hat{M}\hat{N}} \quad \hat{M}\hat{N} = 5, 6\,. \tag{2.46}$$

The $b^{L,c}$ are the model-dependent axions, whose existence is ensured by the second Betti number $b_2(\mathcal{S}^1 \times \mathcal{S}^1) = 1$. There is no zero mode for the $\{56\}$ part of the field strength tensors. In four dimensions an antisymmetric tensor is equivalent to a pseudo-scalar (what we called the model-independent axion), and we redefine:

$$\partial_\mu c^{L,c} \equiv i\frac{1}{6}\epsilon_\mu{}^{\nu\rho\sigma} H^{L,c}_{\nu\rho\sigma}\,. \tag{2.47}$$

We use Greek indexes for 4-dimensional quantities.

The spectrum, after integrating out the compact dimensions, is the same as in SM, with four additional pseudo-scalar fields c^L, c^c, b^L and b^c [25], and the Lagrangian becomes:

$$\begin{aligned} \mathcal{L}^{4D} = {} & \mathcal{L}^{SM} + (Y_\nu \bar{L}\tilde{H}N + \text{h.c.}) \\ & + \frac{g^4 R^3 \Delta_L^2}{6\sqrt{3}\pi^3} b^L \operatorname{Tr} L\tilde{L} + \frac{g_s^2 g^2 R^3 \Delta_c^2}{48\sqrt{3}\pi^3} b^c \operatorname{Tr} L\tilde{L} \\ & + \frac{g^2 g'^2 R^3 \Delta_L^2}{144\sqrt{3}\pi^3} b^L\, F\tilde{F} - \frac{g_s^2 g'^2 R^3 \Delta_c^2}{96\sqrt{3}\pi^3} b^c\, F\tilde{F} \\ & + \frac{g_s^2 g^2 R^3 \Delta_L^2}{48\sqrt{3}\pi^3} b^L \operatorname{Tr} V\tilde{V} \end{aligned} \tag{2.48}$$

$$
\begin{aligned}
&+\frac{1}{2}\partial_\mu b^L\,\partial^\mu b^L + \frac{1}{2}\partial_\mu b^c\,\partial^\mu b^c \\
&+\frac{1}{2}\partial_\mu c^L\,\partial^\mu c^L + \frac{1}{2}\partial_\mu c^c\,\partial^\mu c^c \\
&-\frac{c^L}{3\Delta_L^2 R}\mathrm{Tr}\,L\tilde{L} - \frac{c^c}{3\Delta_c^2 R}\mathrm{Tr}\,V\tilde{V}\,.
\end{aligned}
$$

The Lagrangian $\mathcal{L}^{SM}$ in eq. (2.48) is the SM Lagrangian (in the unitary gauge). The last terms have been obtained using the (six dimensional) Bianchi identity

$$
\mathrm{d}H^{L,c} = -\frac{1}{\Delta_{L,c}^2}\mathrm{d}\omega_3^{L,c}\,, \tag{2.49}
$$

that in $D = 4$ gives the equation of motion for the fields c^L and c^c:

$$
\begin{aligned}
\Box c^L &= -\frac{1}{3\,\Delta_L^2 R}\mathrm{Tr}\,L\tilde{L}\,, \\
\Box c^c &= -\frac{1}{3\,\Delta_c^2 R}\mathrm{Tr}\,V\tilde{V}\,.
\end{aligned} \tag{2.50}
$$

Only three linear combinations φ_1, φ_2 and φ_3 of the four pseudo-scalars are coupled to gauge fields by means of axion-like terms, while the orthogonal combination gives rise to a massless free field with no phenomenological consequence.

2.3.3 Axions

After removing the decoupled scalar from the Lagrangian in eq. (2.48), we obtain

$$
\begin{aligned}
\mathcal{L}^{4D} = \; & \mathcal{L}^{SM} + (Y_\nu \bar{L}\tilde{H}N + \mathrm{h.c.}) \\
&+\frac{1}{2}\partial_\mu\varphi_1\,\partial^\mu\varphi_1 + \frac{1}{2}\partial_\mu\varphi_2\,\partial^\mu\varphi_2 + \frac{1}{2}\partial_\mu\varphi_3\,\partial^\mu\varphi_3 \\
&+\varphi_1\left[\frac{1}{F_1^F}F\tilde{F} + \frac{1}{F_1^L}\mathrm{Tr}\,L\tilde{L} + \frac{1}{F_1^V}\mathrm{Tr}\,V\tilde{V}\right] \\
&+\varphi_2\left[\frac{1}{F_2^F}F\tilde{F} + \frac{1}{F_2^V}\mathrm{Tr}\,V\tilde{V}\right] + \varphi_3\,\frac{1}{F_3^V}\mathrm{Tr}\,V\tilde{V}\,,
\end{aligned} \tag{2.51}
$$

where the constant $F_i^{F,V,L}$ are functions of the coefficients in front of the scalar-gauge fields coupling terms in eq. (2.48). The fields φ_i have the same couplings of the PQ axion: they are invariant under translations but for the coupling to the gauge fields.

As we have seen, there are experimental constraints on the axion couplings coming from the combination of cosmological, astrophysical and accelerator searches [38].

In order to perform the comparison with the experimental constraints, it is necessary to write down the low-energy effective theory in terms of photons, pions, nucleons and axions only.

The most stringent bounds on the axion come from astrophysical observations, *i.e.* from phenomena involving very low energies: we can safely neglect interactions with Z and W bosons, and extract only the electromagnetic couplings

$$\left\{ \begin{array}{rcl} \mathrm{Tr}\, L\tilde{L} & = & \frac{1}{2}\sin^2\theta_W \; F_{em}\tilde{F}_{em} + \cdots \\ F\tilde{F} & = & \cos^2\theta_W \; F_{em}\tilde{F}_{em} + \cdots , \end{array} \right. \tag{2.52}$$

after the rotation of neutral bosons by the weak angle θ_W. Accordingly, only two combinations out of the three φ_i fields couple to the massless gauge fields: one to photons and gluons, the other to photons only. The former has the correct couplings and transformation properties to be identified with the PQ axion. Its presence is a consequence of the anomaly cancellation, and therefore of the choice of writing a six-dimensional gauge theory.

Now we turn the coupling to gluons into a coupling to quarks. This can be achieved by a chiral transformation. Then, using the methods of current algebra, we rewrite the theory in terms of pions, and eliminate quarks and gluons (for more details, see for instance [11]).

Adding the coupling of pions to photons, responsible for the decay $\pi^0 \to 2\gamma$, yields the interaction terms needed to compute all the contributions to the mass matrix of pions and axions. After all of these manipulations we can write

$$\begin{array}{rcl} \mathcal{L}^\pi & = & \frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 + \frac{1}{2}\partial_\mu a\,\partial^\mu a + \frac{1}{2}\partial_\mu\, a'\partial^\mu a' \\ & - & \frac{1}{2}\begin{pmatrix} \pi^0 & a & a' \end{pmatrix} \mathcal{M}^2 \begin{pmatrix} \pi^0 \\ a \\ a' \end{pmatrix} \\ & + & \left(\frac{\pi^0}{f_\pi} + \frac{a}{f_a} + \frac{a'}{f_{a'}}\right) \frac{\alpha}{8\,\pi} F^{\mu\nu}_{e.m.} F^{\rho\sigma}_{e.m.}\, \epsilon_{\mu\nu\rho\sigma} \end{array} \tag{2.53}$$

where

$$\mathcal{M}^2 = \Delta m_\pi^2 \begin{pmatrix} m_{\pi^0}^2/\Delta m_\pi^2 & f_\pi/f_a & k + f_\pi/f_{a'} \\ f_\pi/f_a & (f_\pi/f_a)^2 & f_\pi^2/(f_a f_{a'}) \\ k + f_\pi/f_{a'} & f_\pi^2/(f_a f_{a'}) & m_{\pi^+}^2/\Delta m_\pi^2 (f_\pi/2m)^2 + (f_\pi/f_{a'})^2 \end{pmatrix} . \tag{2.54}$$

The parameter $m \equiv F_3^V(\alpha_s/4\pi)$ and

$$k \equiv \frac{m_{\pi^+}^2}{\Delta m_\pi^2} \frac{m_d - m_u}{m_d + m_u} \frac{f_\pi}{2m} . \tag{2.55}$$

The masses m_u and m_d are those of up and down quarks, $f_\pi \simeq 93$ MeV is the pion decay constant, m_{π^0} and m_{π^+} are the masses of the pions, while $\Delta m_\pi^2 \equiv m_{\pi^0}^2 - m_{\pi^+}^2$. The decay constants are normalized in such a way that the partial decay rate of neutral pions into photons is

$$\Gamma(\pi^0 \to 2\gamma) = \frac{\alpha^2 \, m_\pi^3}{64\pi^3 \, f_\pi^2} \simeq 7.6 \text{ eV} . \tag{2.56}$$

The partial diagonalization of this matrix, performed assuming $f_{a,a'} \gg f_\pi$, makes it possible to identify the physical pion field and the couplings of the two remaining light pseudo-scalars a and a'. The coupling of the axions to the photon comes both from eq. (2.52) and pion-axion mixing. A stringent experimental bound to consider comes from helium burning lifetimes of red giants, and imposes an upper limit to the coupling axion-photon [32]

$$g_{a\gamma} < 10^{-10} \text{ GeV}^{-1} \tag{2.57}$$

with

$$\mathcal{L} = -\frac{1}{4} g_{a\gamma} \, a \, F^{\mu\nu} \tilde{F}_{\mu\nu} . \tag{2.58}$$

The limit of vanishing masses for axions can be used. In our case, we have that

$$g_{a\gamma} \equiv \frac{\alpha}{\pi} \sqrt{\left[\frac{1}{f_a}\left(1 + \frac{\Delta m_\pi^2}{m_{\pi^0}^2}\right)\right]^2 + \left[\frac{1}{f_{a'}}\left(1 + \frac{\Delta m_\pi^2}{m_{\pi^0}^2} + k\frac{f_{a'}}{f_\pi}\frac{\Delta m_\pi^2}{m_{\pi^0}^2}\right)\right]^2} . \tag{2.59}$$

The coupling $g_{a\gamma}$ depends both on positive and negative powers of Δ_c and Δ_L—through the parameters m, f_a and $f_{a'}$ in eq. (2.59), which, in turns, come from the couplings in eq. (2.48). For a fixed value of the radius R, there exists a minimum of

$g_{a\gamma}$ as a function of these free parameters. Taking this minimum and comparing it with the bound in eq. (2.57), yields a constraint on the possible values of R.

A similar bound is obtained by considering the coupling of axions to nucleons

$$\mathcal{L} = -i g_{aN}\, \bar{N}\gamma_5 N\, a\,, \tag{2.60}$$

where, in our case

$$g_{aN} = \frac{g_A m_N}{2}\sqrt{\left[\frac{1}{f_a}\frac{\Delta m_\pi^2}{m_{\pi^0}^2}\right]^2 + \left[\frac{1}{f_{a'}}\left(\frac{\Delta m_\pi^2}{m_{\pi^0}^2} + k\,\frac{f_{a'}}{f_\pi}\frac{\Delta m_\pi^2}{m_{\pi^0}^2}\right)\right]^2}\,, \tag{2.61}$$

and g_A is the axial nucleon coupling, whereas m_N is the nucleon mass. Equation (2.61) is obtained by including only the mixing between the neutral pion and the axion.

Limits coming from supernova SN1987a [33] impose

$$g_{aN} < 3 \times 10^{-10}\,. \tag{2.62}$$

The two bounds eq. (2.57) and eq. (2.62) give

$$\frac{1}{R} > 10^6 \text{ TeV}\,, \tag{2.63}$$

which, applying eq. (2.43), corresponds to

$$M_f > 10^{11} \text{ TeV}\,. \tag{2.64}$$

We have thus obtained an explicit lower limit on the fundamental scale from the experimental bounds on axion couplings.

2.3.4 Higher-order operators

The model under consideration is non-renormalizable; it must be understood as the low-energy limit of a more fundamental theory which gives additional interactions above the cut-off scale M_f. These interactions give rise to operators suppressed by powers of $1/M_f$ that violate the global symmetries of the low-energy theory. However, because of the limit we obtained for M_f, these effects are less worrisome than in models with large extra-dimensions in which the typical scale of such operators is in the TeV range. Nevertheless, some potentially dangerous operators must be checked. In particular, operators like

$$\mathcal{L} \sim \frac{1}{M_P^2} QQQL \,, \tag{2.65}$$

could lead to too fast a proton decay unless M_P is taken of the order of 10^{16} GeV. They are, however, excluded by the residual discrete symmetries that remain after compactification from the $SO(5,1)$ Lorentz symmetry in six dimensions [39].

Operators compatible with these discrete symmetries could, for an arbitrary phase in the coupling, lead to potentially dangerous electric dipole moments

$$\mathcal{L} = i\, e \frac{m_\psi}{M_d^2} \bar{\psi} \sigma^{\mu\nu} \psi F_{\mu\nu} \,. \tag{2.66}$$

Comparing $d \equiv e\, m_\psi / M_d^2$ with the experimental bound [40]

$$d_e < 2 \times 10^{-27}\, \mathrm{e\, cm} \,, \tag{2.67}$$

we find

$$M_d^2 > 10^4\, \mathrm{TeV}^2 \,, \tag{2.68}$$

which is satisfied by several orders of magnitude for $M_d \sim M_f$ imposing the bounds of eq. (2.63) and eq. (2.64). The similar flavor violating operator

$$\mathcal{L} = i\, e \frac{m_\mu}{M_\mu^2} \bar{e} \sigma^{\mu\nu} \mu F_{\mu\nu} \,, \tag{2.69}$$

would induce the decay $\mu \to e\gamma$ with the partial rate

$$\Gamma(\mu \to e\gamma) = \frac{\alpha m_\mu^5}{M_\mu^4} \,, \tag{2.70}$$

where m_μ is the muon mass. Comparing this with the experimental constraint [27],

$$\Gamma(\mu \to e\gamma) < 4 \times 10^{-33}\, \mathrm{TeV} \,, \tag{2.71}$$

yields

$$M_\mu^2 > 10^5\, \mathrm{TeV}^2 \,. \tag{2.72}$$

Another class of potentially dangerous corrections comes from Kaluza-Klein states. The bounds we obtain for the extra-dimensional volume justifies the approach of working with only the zero modes of the theory: the first Kaluza-Klein excitations are at a scale much larger than that experimentally relevant and they can be safely neglected in the computation of observable quantities. Those processes that take place in the SM only at the one-loop level could be an exception to this conclusion. However, no relevant effect is expected for our value of the compactification radius (see, for instance, [41]).

2.3.5 Summary

We have discussed in details the phenomenological consequences of a six-dimensional realization of the SM, in which global and irreducible local anomalies to be absent due to the choice of the fermionic field content, while reducible local anomalies are cancelled by the GS mechanism.

The cancellation of global anomalies imposes the presence of three generations [22]. Local anomaly cancellation requires that the four-dimensional spectrum contains, besides the usual fields of the SM, right-handed neutrinos and axion fields. The axion fields solve the strong CP problem.

The fundamental scale of the theory is related to the decay constant of the axion; therefore, its value must be large enough to evade experimental bounds. The fundamental scale is thus bounded. A problem of naturalness remains because of the large scale M_f of the theory: the Higgs sector requires fine-tuning in order to protect the weak scale. A dynamical explanation of the large difference between the electroweak symmetry breaking scale and the fundamental scale is required because a Supersymmetric version of the model has been shown to contain irreducible anomalies [22].

The large scale of the theory protects the phenomenology from the effect of higher order operators and loop dominated amplitudes, which could otherwise give potentially large contributions to electroweak precision observables, flavor changing processes [41] and CP violating quantities as electric dipole moments (see eq. (2.68)).

We will see in next chapter that a different approach, derived from effective field theory techniques, allows to achieve anomaly cancellation without the introduction of any new degrees of freedom.

Chapter 3

From anomaly matching to anomaly cancellation

When constructing a low-energy action by integrating out some heavy degrees of freedom, the effective theory must reproduce correctly all the Infra-red phenomena of the underlying fundamental theory. This is not automatically guaranteed, as shown, for instance, by the fact that the Appelquist-Carazzone decoupling theorem [42] does not apply to gauge theories with chiral representations, in which fermions acquire a mass, after symmetry breaking, through Yakawa interactions.

As seen in Chapter 1, anomalies are intrinsically low-energy effects. In particular, the contribution to the anomalies of heavy chiral fermions cannot disappear when they have been integrated out. This is the so-called 't Hooft *anomaly matching* argument [43]: the anomalies of the fundamental theory, being long-distance effects, must be equal to those of its low-energy effective action. Indeed, it can be seen that, after integrating out some heavy degrees of freedom, the effective theory acquires these gauge non-invariant couplings [44].

When the phenomenon of spontaneous symmetry breaking takes place, so that some fermions acquire a mass and can be decoupled from the low-energy effective theory, the Goldstone bosons of the theory acquire gauge non-invariant higher order couplings such as to reproduce the anomaly given by the integrated out fermions. These couplings are contained in the WZ effective action [5]: its addition to the effective action is required by the 't Hooft anomaly matching argument.

We describe how to compute the WZ action, for a chiral theory containing local anomalies of a group G broken down to a generic subgroup H, clarifying what are the conditions under which this is possible, and explicitly showing its expression in the simplest cases.

The use of the WZ action makes possible a better understanding of global anomalies [17, 45], and leads to the development of a powerful algebraic technique for computing these anomalies [18], in generic dimension D, and for generic representations of a gauge group G.

The WZ term can also be viewed from a different perspective: an anomalous theory with gauge group G can be rendered gauge invariant by coupling it to G-valued scalar fields that transform non-linearly under the corresponding gauge symmetry, and adding the WZ term to the action [46]. Thus, instead of *reproducing* the anomalies of an underlying theory, one can use the WZ term to *cancel* them. These G-valued fields can either be introduced as additional degrees of freedom (as it happens in the effective quantum field theory approach to the GS mechanism) or be the would-be Goldstone bosons of the spontaneous breaking of the gauge symmetry [47]. The GS mechanism has some intrinsic limitations, for instance it applies only to (local) reducible anomalies, while the WZ term can always be added and every kind of gauge anomalies canceled.

We discuss in some details the case when both local and global anomalies are present [48] in D-dimensional effective theories of a gauge group G spontaneously broken to H. We show how to construct local operators that cancel both local and global anomalies using the would-be Goldstone bosons of the spontaneously broken theory without introducing additional degrees of freedom. We establish the set of conditions under which this cancellation works.

We then apply this technique to the $D = 6$ dimensional model already introduced in the previous Chapter, substituting the GS mechanism with the use of the WZ action. We show that it is possible to cancel all the local anomalies of the theory without introducing any new degree of freedom. The WZ action cancels also the global anomalies, so that the prediction on the number of matter families is lost.

3.1 WZ term and anomaly matching

If $W[A]$ is the effective action for the gauge field A obtained by integrating out fermions, the WZ term is defined in general by [1]

$$\Gamma_{WZ}\,(A,g) \;\equiv\; W[A^g] - W[A]\,, \tag{3.1}$$

where $A^g = g^{-1}Ag + g^{-1}dg$ is the gauge transformed connection. From the definition, it satisfies the following (cocycle) condition

$$\Gamma_{WZ}\,(A^g, U^g) - \Gamma_{WZ}\,(A, U) = -\Gamma_{WZ}\,(A, g)\,, \tag{3.2}$$

where U are the G-valued scalar fields and $U^g \equiv g^{-1}U$.

Although the fermionic effective action is nonlocal, the WZ action is local, as discussed in Chapter 1. The standard way of deriving an explicit expression for the WZ action is through the dimensional descent. Assume that the boundary conditions on the fields are such that space-time can be compactified to a D-dimensional sphere. We can think of this sphere as the boundary of a $D+1$-dimensional ball B. Assume for simplicity that the gauge fields are topologically trivial (no instantons) and that the scalar fields are homotopically trivial. (This will automatically be the case if the homotopy group $\pi_D(G) = 0$, but we need not assume this). Then the fields can be extended to a gauge field $\hat{A}$ and a scalar field $\hat{U}$ defined on all of B. Let $\Omega(\hat{A})$ be the Chern-Simons functional, the integral of the $(D+1)$-form $\omega^0_{D+1}(\hat{A})$ on B. Its gauge variation under a gauge transformation g is the WZ action:

$$\Gamma_{WZ}\,(A, g) = \Omega(\hat{A}^{\hat{g}}) - \Omega(\hat{A})\,, \tag{3.3}$$

where $\hat{g}$ is a gauge transformation in B which coincides with g at the boundaries. Even though the right-hand side of eq. (3.3) is written in terms of the fields on B,

[1] In order to avoid confusion between gauge transformations and scalar fields with values in the group G, here and in the following we slightly modify the notation in respect to the previous Chapters. We refer to gauge transformations as $g(x)$, while the $U(x)$ are G-valued scalars. We also drop everywhere the explicit indication of the gauge coupling.

it is the integral of an exact form (to see this, see eq. (1.103)) and therefore can be rewritten as an integral on the D-dimensional space-time, depending only on the boundary values of the fields $\hat{A}$ and $\hat{U}$. This gives an explicit, local formula for the WZ term.

Since there is usually a nontrivial unbroken group H, one has to generalize the WZ term to the case of G/H-valued, rather than G-valued, scalars φ. Let us review more explicitly how this construction works, following closely the analysis in [49, 50]. Consider a given D-dimensional action with a local gauge symmetry G spontaneously broken to a subgroup H. We denote by T^A the whole set of generators of the Lie algebra of G, whereas T^i and T^α are, respectively, the generators of the Lie algebra of H and of the coset G/H. We assume that G/H is a reductive space, i.e. that $[T^i, T^\alpha] = i\, f_{i\alpha\beta} T^\beta$. Assume also that the fermion content of the corresponding action gives rise to the following anomaly:

$$\begin{aligned} \delta_\alpha W[A] &= -\int d^D x\, \mathcal{G}^\alpha[A(x)]\,, \\ \delta_i W[A] &= 0, \end{aligned} \tag{3.4}$$

where $\mathcal{G}^\alpha$ denotes the usual one-loop gauge anomaly. In other words, eq. (3.4) implies that any potentially anomalous fermionic one-loop amplitude vanishes as soon as one of the external gauge fields belongs to H.

To write a WZ term, we now assume that there exists a unitary gauge, i.e. that there is a globally defined gauge transformation $U(x)$ that transforms the G/H-valued field $\varphi(x)$ to a constant. [2] This is a purely topological restriction on φ; it is less restrictive than the assumption of being homotopic to a constant.

This gauge transformation U is not unique: two G-valued maps U and U' correspond to the same φ if and only if they differ by a right-H transformation: $U'(x) = U(x)h(x)$. Thus we can use U as a dynamical variable instead of φ, but in doing so, we introduce an additional H gauge freedom. Having reformulated the theory in terms of U, we can add to the action a WZ term $\Gamma_{WZ}(A, U)$. Because of

[2] This is always true whenever the gauge field A and its components in H are topologically trivial. More general cases can be however similarly analyzed.

eq. (3.4), $\Gamma_{WZ}(A, Uh) = \Gamma_{WZ}(A, U)$, so it depends only on the coset $\varphi(x) = U(x)H$, i.e. on the would-be Goldstone boson fields. In this case we will therefore write

$$\Gamma_{WZ}(A, \varphi) \equiv \Gamma_{WZ}(A, U) . \tag{3.5}$$

We can give a concrete expression for the WZ term, if it is possible to write $U(x) = e^{i\xi^\alpha(x)T^\alpha}$. This amounts to choosing a specific coordinate system on the coset space and can generally be valid only locally. The fields $\xi^\alpha(x)$ can be identified with the Goldstone bosons. In this notation

$$\Gamma_{WZ}(A, \xi) = i \int_0^1 dt \int d^D x \, \xi^\alpha(x) \, \mathcal{G}^\alpha[A_t] , \tag{3.6}$$

where

$$A_t = e^{-it\xi} A e^{it\xi} + i e^{-it\xi} d e^{it\xi} . \tag{3.7}$$

In particular, in the simple case of $G = U(1)$ and $H = \{1\}$ one can easily check that this reduces to the form

$$\Gamma_{WZ}(A, \xi) \sim \int \xi F \tilde{F} , \tag{3.8}$$

We already met two times this coupling: in Chapter 1, when we reviewed the argument of Gross and Jackiw to discuss the correlation between anomalies and non-renormalizability [6], and in Chapter 2 in the context of the GS mechanism in $D = 4$ [26].

This procedure can be used also when the second condition in eq. (3.4) is replaced by the weaker condition

$$\delta_i W[A_H] = 0 , \tag{3.9}$$

where A_H denotes the component of the gauge field in the sub-algebra of H, i.e. when the fermion representations, restricted to the subgroup H, are free of local anomalies.

The two requirements can be expressed in terms of the D-symbols $D^{a_1 \cdots a_{1+n}}$, defined generalizing eq. (1.20). Respectively, eq. (3.4) is equivalent to the vanishing of $D^{a_1 \cdots a_{1+n}}$ whenever at least one of the indexes labels an element of H, while eq. (3.9) requires that the D-symbol is zero when all of the indexes label generators of H.

We already stressed that the anomaly is defined up to the addition of polynomial (local) terms: as shown in [49, 50], if eq. (3.9) holds, one can add to the action a local functional $B_D(A_H, A)$ such that the second relation in eq. (3.4) is satisfied (see [50] and eq. (3.30) below for an explicit expression of $B_D(A_H, A)$). Thus, in this more general setting, all local anomalies are reproduced by the presence in the low energy action of the modified WZ term

$$\Gamma'_{WZ}(A, U) = W'[A^U] - W'[A], \tag{3.10}$$

where

$$W'[A] = W[A] + B_D(A_H, A). \tag{3.11}$$

3.2 Computing global anomalies

Besides the construction of the low energy effective field theory of chiral fermions, the WZ action enters also a very important technical issue: the computation of global anomalies. We saw in Chapter 1 how Witten discovered the $SU(2)$ global anomaly, by making use of the index theorems. We want here to illustrate a different algebraic technique, which allows a more systematic exploration of the possible global anomalies arising from a generic chiral representation of a group G, in any even dimensions.

In [17], Witten considered the following question. Let us build the effective action of the chiral model based on the group $K = SU(3)_L \times SU(3)_R$, which is broken down to the vectorial $H = SU(3)$. The effective action containes gauge bosons and Golstone bosons only. Certainly there are local anomalies of K, but it is perfectly meaningful to think of the chiral symmetry as a global symmetry of the action, whose anomalies enter the low-energy phenomenology. The subgroup H is anomaly free, and $\pi_4(K) = 0$, so that one can build the WZ term: there are no global anomalies nor topological obstructions to the construction developed in the previous Section.

But what are the conditions under which it is possible to quantize a subgroup G of K? These conditions amount essentially to the requirement of G being anomaly free.

If one tries to gauge $G = SU(2)_L$, there are no local anomalies, because, in $D = 4$, $SU(2)$ admits only real representations. But in $D = 4$ there are global $SU(2)$ anomalies. Let us consider a topologically non-trivial map W in G: it can be written as a map $\bar{W}$ in K, which now can be continuously deformed (in K) to the identity 1 (of K). This deformation requires to act with gauge transformations that cannot belong to G, and therefore are, in general, anomalous, because K has local anomalies. Witten showed explicitly [45, 17] that the total phase shift of the WZ action, along the deformation from $\bar{W}$ to 1 is:

$$\Gamma_{WZ}(\bar{W}\mathrm{d}\bar{W}^{-1}, \bar{W}) - \Gamma_{WZ}(0, 1) = \pi \,. \tag{3.12}$$

This is precisely what we found in Chapter 1, when we computed the global anomaly of $SU(2)$ by making use of the index theorem.

So, we discover that the WZ term reproduces also global anomalies. This is quite a general result, and we will discuss in more detail the construction of such WZ terms in Section 3.3, although with a different aim.

Here we want to stress another important consequence of this observation. There is in fact another (technical) application of the WZ action, to the computation of global anomalies. The global anomaly of the subgroup G, which is free from local anomalies, can be computed as the total phase shift of the WZ action due to the local anomalies of the larger (homotopically trivial) group K.

Exploiting this idea, Elitzur and Nair in [18] constructed a general procedure, based on the algebraic structure given by the existence of the exact sequence of homotopy groups: [3]

$$\pi_{D+1}(K) \to \pi_{D+1}(K/G) \to \pi_D(G) \to \pi_D(K) = 0 \,. \tag{3.14}$$

[3] A sequence is a set of groups G_n and of homomorphisms $f_n : G_n \to G_{n+1}$. It is said to be

Assume we have chiral representations r of the homotopically non-trivial group G such as not to have perturbative anomalies. Suppose there exists a group $K \supset G$ such that $\pi_D(K) = 0$. We embed r in a representation R of K such that $R = r + 1$'s upon reduction to G.

One can compute the variation (global anomaly) of the WZ action obtained integrating a certain chiral representation r of G, deducing from eq. (3.14) exact relations between the generators of the various homotopy groups, and then using the known results about the local anomalies of K due to R.

For instance, consider the exact sequence:

$$\begin{array}{ccccccc} \pi_{2n+1}\left(SU(n+1)\right) & \to & \pi_{2n+1}\left(\dfrac{SU(n+1)}{SU(n)}\right) & \to & \pi_{2n}\left(SU(n)\right) & \to & \pi_{2n}\left(SU(n+1)\right) \, , \\ \mathbf{Z} & \to & \mathbf{Z} & \to & \mathbf{Z}_{n!} & \to & 0 \, . \end{array} \tag{3.15}$$

One uses this exact sequence in order to see that the generator $\tilde{k}$ of $\pi_{2n+1}(SU(n+1))$ gets mapped onto $\tilde{f}^{n!}$, where $\tilde{f}$ is the generator of $\pi_{2n+1}(SU(n+1)/SU(n))$. Then one sees that the image of $\tilde{f}$ is the generator $\tilde{g}$ of $\pi_{2n}(SU(n))$ (written as a $SU(n+1)$ transformation extended to $\mathcal{S}^{2n+1}$).

In this way, knowing that the generator of $\pi_{2n+1}(SU(n+1))$ causes a 2π shift of WZ action [51] defined integrating out a fermion on the fundamental representation of $SU(n+1)$, one obtains simple algebraic expressions for the phase shift ϕ of the WZ action due to the global anomaly of $SU(n)$:

$$\phi = \frac{2\,\pi\,A_R}{n!} \, . \tag{3.16}$$

The factor A_R takes into account that the normalization of the anomaly depends on the choice of R, and in terms of the trace over the fundamental one defines:

$$\mathrm{Tr}_R\, F^{n+1} = A_R\, \mathrm{Tr}\, F^{n+1} + \cdots \, , \tag{3.17}$$

exact if for each n one has

$$\mathrm{Im}\, f_{n-1} = \mathrm{Ker}\, f_n \, . \tag{3.13}$$

where we omitted factorized terms which correspond to lower dimensional traces, that can be seen to define exact forms on $\mathcal{S}^{D+1}$, and then give no contribution to the phase shift we are discussing.

This technique also shows that the global anomaly of G is connected to irreducible (local) anomalies of the group K. This property is relevant in extending the construction to the case in which the fermion field content produces local anomalies which can be canceled by the GS mechanism [34]: any GS terms, having to do with factorized anomalies, cannot contribute to global anomalies.

In the case discussed by Witten, $SU(2)$ in $D = 4$, this procedure works: after building the embedding onto the homotopically trivial group $SU(3)$—considering the fundamental representation of $SU(3)$ decomposed into the fundamental of $SU(2)$ and a singlet—from eq. (3.16) one gets again the known result $\phi = \pi$.

3.3 Canceling local and global anomalies

We have seen that the WZ term reproduces in the low-energy effective theory the effects of both local and global anomalies. We also have seen that one can make use of the WZ action as a technical tool in order to compute the global anomalies of a group G, provided an exact sequence of homotopy groups can be built relating the global anomaly of G to known results on local anomalies of a larger group K in which to embed G.

In this Section we show that the WZ action can have also a completely different interpretation: it is possible to cancel both local and global anomalies of a gauge theory of the group G spontaneously broken to a subgroup H, by the addition of generalized WZ terms to the action. This can be done provided that: [4]

1. the coset space G/H is reductive;

2. the fermion representations are free of *local* anomalies when restricted to the group H;

[4] A possible limitation of this procedure arises whenever we wish to preserve further symmetries of the original action that are explicitly broken by the introduction of these operators.

3. the fermion representations are free of *global* anomalies when restricted to the group H;

4. G can be embedded in a group K such that its homotopy group $\pi_D(K) = 0$ and the fermion representations can be extended to K without generating further anomalies of G.

Conditions 1 and 2 have been derived for local gauge anomalies [49, 50], as discussed already in Section 3.1, while conditions 3 and 4 are closely related to the results of [17, 45, 18], as discussed in Section 3.2. We need to generalize the idea of Elitzur and Nair of computing global anomalies as local anomalies of a larger group K [18]: we argue that the operators constructed in [49, 50] can be used to cancel local anomalies of G, but can also be properly defined globally and used to cancel global anomalies of G, provided the above four conditions are satisfied.

This result is of interest in model building in so far as the fermion content of an effective theory is only restricted by the cancellation of the anomalies of the unbroken group H.

If $\pi_D(G) = 0 = \pi_D(G/H)$, so that no global anomalies can arise, and the low-energy theory contains both fermions and scalars, if we add to the scalar action the WZ term, then the effective action $\tilde{\Gamma}(A, U) = W[A] + \Gamma'_{WZ}\ (A, U)$ is seen to be gauge invariant by means of eq. (3.1) and eq. (3.2).[5]

Once gauge invariance is restored, it is possible to shift the fields in such a way as to decouple the Goldstone bosons φ, and give mass to the gauge bosons of G/H (unitary gauge). The presence of the term defined as in eq. (3.10) in the generic gauge signals however the non-renormalizability of the theory. An explicit proof of non-renormalizability in the 't Hooft-Landau gauge is given in [47]. This can also be understood, as anticipated in Chapter 1, by looking at the diagram giving rise to the anomaly: it contains off-shell non-analytic contributions [1] that cannot be canceled by any local counter-term.

[5]It can be shown that the addition of the WZ term cures also the problems arising in the canonical formulation of the theory [53], so that the quantum theory, aside from renormalizability issues, is well defined.

As long as local gauge anomalies are concerned, effective theories, for which renormalizability is not a requirement, can be made anomaly free by the addition of an appropriate WZ term. No restriction on the fermion content of the theory is needed, provided that the first two conditions listed above are satisfied.

So far, our analysis has been purely local without considering possible topological obstructions or possible global gauge anomalies. In fact, a potential problem may arise in the above construction whenever $\pi_D(G/H) \neq 0$ [50], because all the construction of the WZ term outlined in Section 3.1 requires the extension of all the G-valued maps to the $D+1$-dimensional ball B. On the other hand, the same condition of having a non-trivial homotopy group can give rise to global anomalies. Let us see how to generalize the above procedure in order to take into account these global issues and remove the above topological condition.

To begin with, we assume that no local anomalies are present. As we have seen, even a theory which is free of local anomalies—i.e. invariant under infinitesimal gauge transformations—can still be anomalous under gauge transformations that are not homotopic to the identity. This can occur whenever $\pi_D(G) \neq 0$. [6]

Extending the results of [46], it is always possible to cancel any global anomaly by coupling the theory to a G-valued scalar field $U(x)$ and adding a suitable WZ term γ to the action [54]. The absence of local anomalies means that γ, defined as in eq. (3.1), is zero, independently of A, when g is homotopic to a constant. Every g in a certain homotopy class can be written as $g_1 g'$ where g_1 is a fixed (representative) map in that same homotopy class and g' is homotopically trivial. Then from the cocycle condition (3.2) one sees that

$$\gamma(A, g_1 g) = \gamma(A, g_1)\,, \tag{3.18}$$

so γ is invariant under continuous deformations of g. In conclusion, γ depends only on the homotopy class $[g]$ and therefore can be seen as a topological term. From

[6] When space-time is not flat, or instantons are present, the conditions can become more complicated. We do not consider these cases.

eq. (3.2) we also see that it must define a representation of $\pi_D(G)$:

$$\gamma(g_1 \cdot g_2) = \gamma(g_1) + \gamma(g_2)\,. \tag{3.19}$$

We are especially interested in the case where the scalar fields do not have to be introduced *ad hoc*, but are already present in the theory. We are thus led to ask: if the G gauge symmetry is spontaneously broken to H due to a Higgs mechanism, can one cancel the global anomalies of the low-energy effective theory by means of a WZ term written as a functional of the would-be Goldstone bosons? The answer is that this is possible if the fermion representations, restricted to H, are free of global anomalies.

To show this, we start again by replacing the coset-valued field φ by the G-valued field U. Having reformulated the theory in terms of U, we can write WZ terms $\gamma(U)$ cancelling any global anomaly, using the method described above. Of course, as in Section 3.1, we want these terms to depend only on the physical scalar fields φ. This will automatically be the case if the unbroken group H is free of global anomalies. Indeed, in this case it follows from (3.19) that

$$\gamma(U \cdot h) = \gamma(U) + \gamma(h) = \gamma(U)\,, \tag{3.20}$$

and therefore γ really only depends on the (homotopy class of the) coset-valued field $\varphi(x)$. Global H anomalies will certainly be absent if $\pi_D(H) = 0$, but even if $\pi_D(H)$ is nontrivial the theory may be free of global H anomalies provided the fermion representations are chosen appropriately. In this case, global anomalies can be present only when $\pi_D(G/H) \neq 0$.

One may be interested in a general method for calculating the topological term γ: following the discussion in [45] and [18], as we discussed in the previous Section, it can be written as a WZ term, albeit for a larger group. Let us get back to this construction, and discuss it in our context of anomaly cancellation. The construction of the WZ term given in eq. (3.3) demands that the map U be homotopic to a constant. If U is not homotopic to a constant one can still proceed by embedding G in a larger group K such that $\pi_D(K) = 0$. One then defines a map $\bar{U}$ by composing

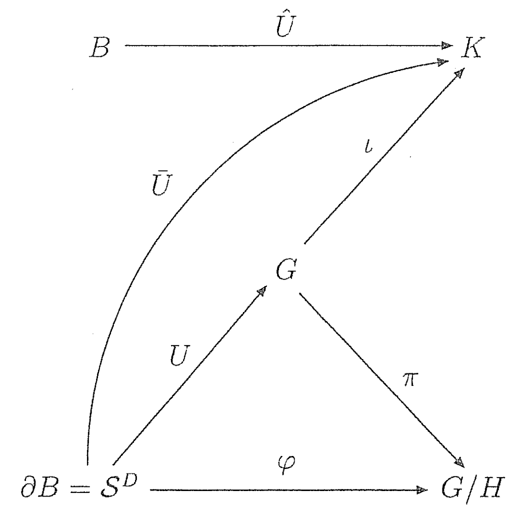

Figure 3.1: Diagrammatic representation of the algebraic construction described in the text. The coset valued scalar fields φ are lifted through π^{-1} to the fields U, and the homotopically non-trivial group G is embedded in K by ι. The maps $\bar{U} \equiv \iota \circ U$ are homotopically trivial in D dimensions, so that they can be extended to $\bar{U}$, defined in the $D+1$-dimensional ball B, whose boundary is the D-dimensional space-time $\partial B = \mathcal{S}^D$ (compactification to a sphere is assumed). In the simpler case when $\pi_D\left(G\right) = 0$, then $K = G$, ι is the identity in G, and $\bar{U} = U$.

U with this embedding ι, and a gauge field $\bar{A}$ by the corresponding embedding of the Lie algebras. One can for instance write:

$$\bar{U} = \begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \bar{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} . \tag{3.21}$$

The fermion content of K is so chosen that upon reduction to G it gives rise to the required G representations and a number of G singlets (or anomaly-free representations of G). The diagram of this construction in depicted in Fig. 3.1.

The field $\bar{U}$ is homotopic to a constant (in K) and one can explicitly write the WZ term $\Gamma^K_{WZ}\left(\bar{A}, \bar{U}\right)$ as in eq. (3.3), by defining the extension $\hat{U}$ to the $D+1$ dimensional ball B. As described above, if the theory is free of perturbative anomalies we can take simply

$$\gamma(g) \equiv \Gamma_{WZ}^{K}\,(\bar{A},\bar{g})\,, \tag{3.22}$$

where $\bar{g} = \iota \circ g$. Thus, the WZ term canceling the non-perturbative G-anomaly can be calculated as a perturbative WZ term for the larger group K, along the lines of [18].

Finally, we are ready to consider the generic case when the theory has perturbative anomalies, with only $\delta_i W[A_H] = 0$, and at the same time $\pi_D(G/H) \neq 0$. In Section 3.1 we have seen, following [49, 50], that it is always possible to add to the action a local operator $B_D(A_H, A)$ such as to construct a WZ term as in eq. (3.10), that cancels the perturbative anomalies. This WZ term is well defined when the map U is homotopic to a constant. We have also described a way of writing a WZ term for a larger group K that, when restricted to the subgroup G, makes sense for all maps U, irrespective of their homotopy class. If there are local anomalies, this WZ term, now denoted $\Gamma'^{K}_{WZ}\,(\bar{A},\bar{U})$, is no longer topological. In fact, one can see by means of eq. (3.21) that when U is homotopically trivial, it agrees with $\Gamma'_{WZ}(A,U)$ defined in eq. (3.10). There follows that, defining

$$\Gamma'_{WZ}\,(A,U) \equiv \Gamma'^{K}_{WZ}\,(\bar{A},\bar{U})\,, \tag{3.23}$$

the resulting effective action, obtained with the addition of this local term, is defined for all maps U, whether trivial or not, and free of both local and global gauge anomalies for the group G spontaneously broken to H. [7]

A physical interpretation of the group K can be given (although not necessary), along the lines of [44]. One can imagine a microscopic theory with a gauge group K and a completely anomaly free (local and global) fermion spectrum. The fermions are in representations of K such that, upon the spontaneous breaking of K to G, they give rise to the required fermions in representations of G. The massive fermions that are integrated out produce the WZ term $\Gamma'^{K}_{WZ}(\bar{A},\bar{U})$ above. There are other effects due to the heavy fermions, for instance they may enter the radiative

[7]Even in the absence of global anomalies, this construction generalizes that of [50], which is limited by the condition $\pi_D\,(G/H) = 0$.

corrections to the masses of the scalar fields, but, for what anomalies are concerned, Γ'^{K}_{WZ}, satisfying the 't Hooft argument, reproduces all the physics of the fundamental theory. On the other hand, it is not necessary to invoke the existence of these heavy fermionic fields: in general gauge non-invariant effective couplings could have other origins. For instance, we have seen that in the simplest case of an Abelian Higgs model, the same coupling Γ'^{K}_{WZ} can be deduced from the compactification to four dimensions of Superstring theories [26].

3.4 The Standard Model in Six Dimensions

Let us go back to the six-dimensional version of the Standard Model [22] we discussed in Chapter 2, and try to apply the mechanism discussed here. We keep the same field content for each single family as in Table 2.2, which was restricted by the requirement of having a vector-like theory of strong and gravitational interactions

As we saw, the application of the GS mechanism to the reducible gauge anomalies, with the introduction of two GS (2-form) fields, leaves some pseudo-scalar remnants in the low-energy $D = 4$ effective theory (after compactification of the two extra dimensions). These fields behave as axions: they solve the strong-CP problem, but impose severe constraints on the volume of the compact extra-dimensions, requiring a very heavy fundamental scale in the theory, in the range of GUT theories. This fact produces a naturalness problem in the scalar sector: the Higgs mass must be many orders of magnitude lower than the fundamental scale, in order to reproduce correctly the mass of the electroweak gauge bosons.

As we discussed in this chapter, there is also another way, besides the GS mechanism, to cancel gauge anomalies: instead of introducing new degrees of freedom transforming non-linearly under the action of the gauge symmetry, one can use the non-linear realization provided by the Goldstone bosons already present in the spectrum, by the addition of a WZ term to the action. In this way no axionic remnants are left in the $D = 4$ effective theory.

Let us verify that this idea works correctly, and what are its main effects. The choice of fermion content guarantees that the model has no pure gravitational anomaly. The addition of a local Chern-Simons term shifts mixed gravitational anomalies into gauge anomalies.

We can thus identify the groups of the previous sections with those in [22]

$$\begin{cases} K &= SU(4)_c \times SU(4)_L \times U(1)_Y \,, \\ G &= SU(3)_c \times SU(2)_L \times U(1)_Y \,, \\ H &= SU(3)_c \times U(1)_{\text{e.m.}} \,. \end{cases} \tag{3.24}$$

Since

$$\pi_6\,(G/H) = \pi_6\,(SU(2)) = \mathbf{Z}_{12}\,, \tag{3.25}$$

we have enlarged G to K for which $\pi_6\,(K) = 0$.

Each $SU(2)$ doublet in Table 2.2 goes into the fundamental representation of $SU(4)$ (to be decomposed into a $SU(2)$ doublet plus singlets). In addition to the already-present singlets, more singlets are necessary in order to preserve the $U(1)_Y^4$ anomaly. Notice that the choice of vector-like representations of the group $SU(3)_c$ makes the theory free from global and irreducible $SU(3)_c$ anomalies, but does not prevent from the presence of anomalies in the form

$$\left[SU(2)_L\right]^2 \left[SU(3)_c\right]^2 \,, \tag{3.26}$$

and similar, so that the subgroup H is not anomaly free, but it is simple (using the definition of electric charge $Q \equiv Y + T^3$) to verify that, with the field content in Table 2.2, fermions form representations of group H such that the weaker condition leading to the B_6 term is satisfied: there are no anomalies in the form H^4, which has only vector-like representations. One rearranges the generators Y and T^i of $U(1)_Y$ and $SU(2)_L$, respectively, in the following way:

$$\begin{cases} Q &\equiv Y + T^3 \\ Z &\equiv Y - T^3 \\ T^{\pm} &\equiv T^1 \pm iT^2 \end{cases} . \tag{3.27}$$

the set $\{Q, \lambda^a\}$, where λ^a are the generators of $SU(3)_c$ (Gell-Mann matrices), form a complete set of generators of H, while $T^\alpha = \{Z, T^{\pm}\}$ generate the coset G/H.

All the λ^a commutes with the T^α generators, while:

$$\left\{ \begin{array}{lcl} [Z\,,\,Q] & = & 0 \\ [T^{\pm}\,,\,Q] & = & \mp T^{\pm} \end{array} \right. , \tag{3.28}$$

which demonstrate that G/H is reductive. So all the conditions required in order to write the WZ term hold, and therefore the WZ term can be used to cancel all anomalies, as explained in the previous Section. This term can be explicitly written as

$$\Gamma'^{K}_{WZ} = -\int_0^1 \mathrm{d}t\,\xi^\alpha \mathcal{G}'^{\alpha}(\bar{A}_t)\,, \tag{3.29}$$

where $\mathcal{G}'^{\alpha}(A_t) = \mathcal{G}^\alpha(A_t) + \delta_\alpha B_6(A_H, A)$, with

$$B_6(A_0, A_1) = 12\int_\Delta \mathrm{d}\mu\,\mathrm{d}\lambda\,\mathrm{Str}[A_0, A_1, F^2_{\mu,\lambda}]\,, \tag{3.30}$$

and A_t is defined in eq. (3.7). Following the notation of [50], we defined:

$$\begin{array}{rcl} \mathrm{Str}[C_1, \cdots, C_N] & = & \sum_P \frac{(-1)^{f_P}}{N!}\,\mathrm{Tr}\,[C_{P_1}\cdots C_{P_N}]\,, \\ F_{\mu,\lambda} & = & \mathrm{d}A_{\mu,\lambda} + A^2_{\mu,\lambda}, \\ A_{\mu,\lambda} & = & \mu A_0 + \lambda A_1\,. \end{array} \tag{3.31}$$

In eq. (3.29), the fields ξ^α are the longitudinal components of the massive gauge bosons W and Z. $\mathcal{G}^\alpha$ is the (one-loop) anomaly. All fields must be thought as K-valued, as in eq. (3.21). The integration region Δ is a triangle in the (μ, λ) plane with vertices $(0,1)$, $(1,0)$ and the origin; f_P is the number of times the permutation P permutes two odd objects.

The term B_6 in eq. (3.29) remains, even in the unitary gauge, as a Chern-Simons coupling between gauge bosons, and gives rise to dimension-six operators after compactification to $D=4$. These operators are suppressed by a coefficient proportional to the compact volume R^2. Even though they modify, for instance, the photon couplings, we expect their effects to be negligible because $R \sim$ few TeV^{-1} [41].

Thanks to this construction, no GS fields are needed; since there are no axions in the low energy $D=4$ theory, their experimental bounds do not apply. However, all global anomalies are canceled as well, and therefore the interesting prediction on the number of families is lost. A similar procedure might be relevant also for other six-dimensional extensions discussed in [55].

Bibliography

[1] J. S. Bell and R. Jackiw, Nuovo Cim. A **60** (1969) 47;
S. L. Adler, Phys. Rev. **177** (1969) 2426.

[2] R. Jackiw, *In *Treiman, S.b. (Ed.) Et Al.: Current Algebra and Anomalies*, 81-210.*

[3] K. Fujikawa, Phys. Rev. Lett. **42**, 1195 (1979), Phys. Rev. D **21**, 2848 (1980) [Erratum-ibid. D **22**, 1499 (1980)].

[4] R. Stora, CPT-76/P.854 *Lectures given at Summer Inst. for Theoretical Physics, Cargese, France, Jul 12-31, 1976*, LAPP-TH-94 *Seminar given at Cargese Summer Inst.: Progress in Gauge Field Theory, Cargese, France, Sep 1-15, 1983*; B. Zumino, UCB-PTH-83/16 *Lectures given at Les Houches Summer School on Theoretical Physics, Les Houches, France, Aug 8 - Sep 2, 1983*; J. Manes, R. Stora and B. Zumino, Commun. Math. Phys. **102**, 157 (1985).

[5] J. Wess and B. Zumino, Phys. Lett. B **37**, 95 (1971).

[6] D. J. Gross and R. Jackiw, Phys. Rev. D **6**, 477 (1972).

[7] E. Witten, Phys. Lett. B **117**, 324 (1982).

[8] S. B. Treiman, E. Witten, R. Jackiw and B. Zumino, *Singapore, Singapore: World Scientific (1985) 537p.*

[9] R. A. Bertlmann, *Oxford, UK: Clarendon (1996) 566 p. (International series of monographs on physics: 91).*

[10] M. Nakahara, *Bristol, UK: Hilger (1990) 505 p. (Graduate student series in physics).*

[11] S. Weinberg, *Cambridge, UK: Univ. Pr. (1996) 489 p.*

[12] S. L. Adler and W. A. Bardeen, Phys. Rev. **182**, 1517 (1969).

[13] see for instance S. Okubo and J. Patera, Phys. Rev. D **31**, 2669 (1985) and ref. therein.

[14] P. H. Frampton and T. W. Kephart, UTTG-9-83.

[15] L. Alvarez-Gaume and E. Witten, Nucl. Phys. B **234**, 269 (1984).

[16] W. A. Bardeen and B. Zumino, Nucl. Phys. B **244**, 421 (1984).

[17] E. Witten, Nucl. Phys. B **223**, 422 (1983).

[18] S. Elitzur and V. P. Nair, Nucl. Phys. B **243**, 205 (1984).

[19] M. B. Green and J. H. Schwarz, Nucl. Phys. B **255**, 93 (1985).

[20] M. B. Green and J. H. Schwarz, Phys. Lett. B **149**, 117 (1984); see also A. Sagnotti, Phys. Lett. B **294**, 196 (1992) [arXiv:hep-th/9210127].

[21] R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977); Phys. Rev. D **16**, 1791 (1977).

[22] B. A. Dobrescu and E. Poppitz, Phys. Rev. Lett. **87**, 031801 (2001) [hep-ph/0102010].

[23] M. Fabbrichesi, M. Piai and G. Tasinato, Phys. Rev. D **64**, 116006 (2001) [arXiv:hep-ph/0108039].

[24] M. B. Green, J. H. Schwarz and E. Witten, *Cambridge, Uk: Univ. Pr. (1987) 596 P. (Cambridge Monographs On Mathematical Physics).*

[25] E. Witten, Phys. Lett. B **149**, 351 (1984). Nucl. Phys. B **268**, 79 (1986).

[26] M. Dine, N. Seiberg, X. G. Wen and E. Witten, Nucl. Phys. B **278**, 769 (1986); Nucl. Phys. B **289**, 319 (1987).

[27] J. Heitger, R. Sommer and H. Wittig [ALPHA Collaboration], Nucl. Phys. B **588**, 377 (2000) [arXiv:hep-lat/0006026]; A. C. Irving, C. McNeile, C. Michael, K. J. Sharkey and H. Wittig [UKQCD Collaboration], Phys. Lett. B **518**, 243 (2001) [arXiv:hep-lat/0107023]; D. R. Nelson, G. T. Fleming and G. W. Kilcup, arXiv:hep-lat/0112029.

[28] F. Wilczek, Phys. Rev. Lett. **40**, 279 (1978);
S. Weinberg, Phys. Rev. Lett. **40**, 223 (1978).

[29] J. E. Kim, Phys. Rev. Lett. **43**, 103 (1979);
M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **166**, 493 (1980);
A. R. Zhitnitsky, Sov. J. Nucl. Phys. **31**, 260 (1980) [Yad. Fiz. **31**, 497 (1980)];
M. Dine, W. Fischler and M. Srednicki, Phys. Lett. B **104**, 199 (1981).

[30] D. J. Gross, R. D. Pisarski and L. G. Yaffe, Rev. Mod. Phys. **53**, 43 (1981).

[31] C. Vafa and E. Witten, Nucl. Phys. B **234**, 173 (1984); Phys. Rev. Lett. **53**, 535 (1984).

[32] G. G. Raffelt, Phys. Rev. D **33**, 897 (1986).

[33] G. G. Raffelt, Phys. Rept. **198**, 1 (1990).

[34] M. Bershadsky and C. Vafa, arXiv:hep-th/9703167.

[35] B. Adeva *et al.* [L3 Collaboration], Phys. Lett. B **275**, 209 (1992);
D. Decamp *et al.* [ALEPH Collaboration], Phys. Lett. B **235**, 399 (1990).

[36] P. H. Frampton and T. W. Kephart, Phys. Rev. Lett. **50**, 1343 (1983) [Erratum-ibid. **51**, 232 (1983)]; Phys. Rev. D **28**, 1010 (1983).

[37] P. H. Frampton and K. Yamamoto, Phys. Lett. B **156**, 345 (1985).

[38] see for instance P. Sikivie, Nucl. Phys. Proc. Suppl. **87**, 41 (2000) [arXiv:hep-ph/0002154]; and references therein.

[39] T. Appelquist, B. A. Dobrescu, E. Ponton and H. Yee, hep-ph/0107056.

[40] E. D. Commins, S. B. Ross, D. DeMille and B. C. Regan, Phys. Rev. A **50**, 2960 (1994).

[41] A. Delgado, A. Pomarol and M. Quiros, JHEP **0001**, 030 (2000) [hep-ph/9911252];
T. Appelquist and B. A. Dobrescu, hep-ph/0106140;

[42] T. Appelquist and J. Carazzone, Phys. Rev. D **11**, 2856 (1975).

[43] G. 't Hooft, lecture given at the Cargèse Summer Institute, 1979, in *Recent Developments in Gauge Theories*, eds. G. 't Hooft *et al.* (Plenum, New York, 1980).

[44] T. Sterling and M. J. Veltman, Nucl. Phys. B **189**, 557 (1981); M. S. Chanowitz, M. A. Furman and I. Hinchliffe, Nucl. Phys. B **153** (1979) 402; E. D'Hoker and E. Farhi, Nucl. Phys. B **248**, 59 (1984).

[45] E. Witten, Nucl. Phys. B **223** (1983) 433.

[46] L. D. Faddeev and S. L. Shatashvili, Phys. Lett. B **167**, 225 (1986); N. V. Krasnikov, JETP Lett. **41**, 586 (1985) [Pisma Zh. Eksp. Teor. Fiz. **41**, 481 (1985)]; O. Babelon, F. A. Schaposnik and C. M. Viallet, Phys. Lett. B **177**, 385 (1986); K. Harada and I. Tsutsui, Phys. Lett. B **183**, 311 (1987).

[47] J. Preskill, Ann. Phys. **210**, 323 (1991).

[48] M. Fabbrichesi, R. Percacci, M. Piai and M. Serone, arXiv:hep-th/0207013.

[49] L. Alvarez-Gaumé and P. Ginsparg, Nucl. Phys. B **262** (1985) 439; Y.-S. Wu, Phys. Lett. B **153** (1985) 70.

[50] C. S. Chu, P. M. Ho and B. Zumino, Nucl. Phys. B **475**, 484 (1996) [arXiv:hep-th/9602093].

[51] see, for instance, B. Zumino in 1.3

[52] C. Bouchiat, J. Iliopoulos and P. Meyer, Phys. Lett. B **38**, 519 (1972).

[53] R. Jackiw and R. Rajaraman, Phys. Rev. Lett. **54**, 1219 (1985) [Erratum-ibid. **54**, 2060 (1985)]; R. Rajaraman, Phys. Lett. B **184**, 369 (1987); R. Percacci and R. Rajaraman, Int. J. Mod. Phys. A **4** (1989) 4177.

[54] R. Percacci, Mod. Phys. Lett. A **12** (1987) 977.

[55] N. Borghini, Y. Gouverneur and M. H. Tytgat, Phys. Rev. D **65**, 025017 (2002) [arXiv:hep-ph/0108094].

Part II

OSCILLATION OF NEUTRINOS FROM NUCLEAR REACTORS

Introduction

In recent years the experiments with solar and atmospheric neutrinos have provided strong evidences in favor of the existence of oscillations between the flavor neutrinos, ν_e, ν_μ and ν_τ. Further progress in our understanding of the neutrino mixing and oscillations requires precise measurements of the parameters entering the oscillation probabilities: the neutrino mass-squared differences and mixing angles, and the reconstruction of the neutrino mass spectrum.

Reactor experiments give the possibility to test the Large Mixing Angle (LMA) Mikheev-Smirnov-Wolfenstein (MSW) solution of the solar neutrino problem. While the KamLAND experiment should be able to test this solution, a new experiment with a shorter baseline might be required to determine the mass difference $\Delta m^2_\odot$ entering solar neutrino oscillations with high precision if the results of the KamLAND experiment show that $\Delta m^2_\odot > 10^{-4}$ eV^2. Performing a three-neutrino oscillation analysis of both the total event rate suppression and the e^+-energy spectrum distortion caused by the $\bar{\nu}_e-$oscillations in vacuum, we show that a value of $\Delta m^2_\odot$ from the interval 10^{-4} eV^2 $< \Delta m^2_\odot \lesssim 8.0 \times 10^{-4}$ eV^2 could be determined with a high precision in experiments with $L \cong (20 - 25)$ km if the e^+-energy spectrum is measured with a sufficiently good accuracy.

Furthermore, if $\Delta m^2_\odot \cong (1.0 - 5.0) \times 10^{-4}$ eV^2, such an experiment with $L \cong (20 - 25)$ km might also be able to distinguish between the cases of neutrino mass spectrum with normal and inverted hierarchy; for larger values of $\Delta m^2_\odot$ not exceeding 8.0×10^{-4} eV^2, a shorter baseline, $L \cong 10$ km, should be used for the purpose.

The indicated possibility poses remarkable challenges and might be realized for a limited range of values of the relevant parameters. The corresponding detector must have a good energy resolution (allowing a binning in the positron energy with $\Delta E_e \lesssim 0.40$ MeV) and the observed event rate due to the reactor $\bar{\nu}_e$ must be sufficiently high to permit a high precision measurement of the e^+-spectrum. Furthermore, the mixing angle constrained by the CHOOZ and Palo Verde data θ must be sufficiently large ($\sin^2\theta \sim 0.03 - 0.05$), and the "solar" mixing angle $\theta_\odot$ should not be maximal ($\sin^2 2\theta_\odot \lesssim 0.9$). In addition, the value of Δm^2_{31}, which is responsible for the dominant $\nu_\mu \rightarrow \nu_\tau$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_\tau$ oscillations of the atmospheric neutrinos, should be known to high precision.

Chapter 4

LMA-MSW Solution of the Solar Neutrino Problem and Reactor Experiments

The atmospheric neutrino data can be explained by dominant $\nu_\mu \rightarrow \nu_\tau$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_\tau$ oscillations, characterized by large, possibly maximal, mixing, and a mass squared difference, Δm^2_{atm}, having a value in the range [1] (99% C.L.):

$$1.3 \times 10^{-3}\text{eV}^2 \lesssim |\Delta m^2_{atm}| \lesssim 5 \times 10^{-3}\text{eV}^2. \tag{4.1}$$

The first results from the Sudbury Neutrino Observatory (SNO) [2], combined with the mean event rate data from the Super-Kamiokande (SK) experiment [3], provide a very strong evidence for oscillations of the solar neutrinos [4] - [10]. Global analyses of the solar neutrino data, including the SNO results and the SK data on the e^-–spectrum and day-night asymmetry, show that the data favor the LMA MSW solution of the solar neutrino problem, with the corresponding neutrino mixing parameter $\sin^2 2\theta_\odot$ and mass-squared difference $\Delta m^2_\odot$ lying in the regions (99.73% C.L.):

$$2 \times 10^{-5}\text{eV}^2 \lesssim \Delta m^2_\odot \lesssim 8 \times 10^{-4}\text{eV}^2 \tag{4.2}$$

$$0.6 \leq \sin^2 2\theta_\odot \leq 1. \tag{4.3}$$

The best fit value of $\Delta m^2_{\odot}$ found in the independent analyses [5, 6, 7, 9] is spread in the interval $(4.3 - 6.3) \times 10^{-5}$ eV^2. The results obtained in [5, 6, 7, 9] show that values of $\Delta m^2_{\odot} > 10^{-4}$ eV^2 are allowed already at 90% C.L. Values of $\cos 2\theta_{\odot} < 0$ (for $\Delta m^2_{\odot} > 0$) are disfavored by the data.

Important constraints on the oscillations of electron (anti-)neutrinos, which play a significant role in our current understanding of the possible patterns of oscillations of the three flavor neutrinos and anti-neutrinos, were obtained in the CHOOZ and Palo Verde disappearance experiments with reactor $\bar{\nu}_e$ [11, 12]. The CHOOZ and Palo Verde experiments were sensitive to values of neutrino mass squared difference $\Delta m^2 \gtrsim 10^{-3}$ eV^2, which includes the corresponding atmospheric neutrino region, eq. (1). No disappearance of the reactor $\bar{\nu}_e$ was observed. Performing a two-neutrino oscillation analysis, the following rather stringent upper bound on the value of the corresponding mixing angle, θ, was obtained by the CHOOZ collaboration [1] [11] at 95% C.L. for $\Delta m^2 \geq 1.5 \times 10^{-3}$eV^2:

$$\sin^2\theta < 0.09. \tag{4.4}$$

The precise upper limit in eq. (4.4) is Δm^2-dependent: it is a decreasing function of Δm^2 as Δm^2 increases up to $\Delta m^2 \simeq 6{\cdot}10^{-3}$ eV^2 with a minimum value $\sin^2\theta \simeq 10^{-2}$. The upper limit becomes an increasing function of Δm^2 when the latter increases further up to $\Delta m^2 \simeq 8 \cdot 10^{-3}$ eV^2, where $\sin^2\theta < 2 \cdot 10^{-2}$. Somewhat weaker constraints on $\sin^2\theta$ have been obtained by the Palo Verde collaboration [12]. In the future, $\sin^2\theta$ might be further constrained or determined, e.g., in long baseline neutrino oscillation experiments [13].

The long baseline experiment with reactor $\bar{\nu}_e$ KamLAND [14] has been designed to test the LMA MSW solution of the solar neutrino problem. This experiment is planned to provide a rather precise measurement of $\Delta m^2_{\odot}$ and $\sin^2 2\theta_{\odot}$. Due to the long baseline of the experiment, $L \sim 180$ km, however, $\Delta m^2_{\odot}$ can be determined with a relatively good precision only if $\Delta m^2_{\odot} \lesssim 10^{-4}$ eV^2.

[1] The possibility of large $\sin^2\theta > 0.9$ which is admitted by the CHOOZ data alone is incompatible with the neutrino oscillation interpretation of the solar neutrino deficit (see, e.g., [15, 16])

The explanation of both the atmospheric and solar neutrino data in terms of neutrino oscillations requires, as is well-known, the existence of 3-neutrino mixing in the weak charged lepton current:

$$\nu_{l\mathrm{L}} = \sum_{j=1}^{3} U_{lj} \nu_{j\mathrm{L}}, \tag{4.5}$$

where ν_{lL}, $l = e, \mu, \tau$, are the three left-handed flavor neutrino fields, $\nu_{j\mathrm{L}}$ is the left-handed field of the neutrino ν_j having a mass $m_j > 0$ and U is a 3×3 unitary mixing matrix - the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) neutrino mixing matrix [17, 18]. The three neutrino masses $m_{1,2,3}$ can obey the so-called normal hierarchy (NH) relation $m_1 < m_2 < m_3$, or that of the inverted hierarchy (IH) type, $m_3 < m_1 < m_2$. Thus, in order to reconstruct the neutrino mass spectrum in the case of 3-neutrino mixing, it is necessary to establish, in particular, which of the two possible types of neutrino mass spectrum is actually realized. This information is particularly important for the studies of a number of fundamental issues related to lepton mixing, as like the possible Majorana nature of massive neutrinos, which can manifest itself in the existence of neutrino-less double β-decay (see, e.g., [19, 20]). It would also constitute a critical test for theoretical models of fermionic mass matrices and flavor physics in general.

It would be possible to determine whether the neutrino mass spectrum is with normal or inverted hierarchy in terrestrial neutrino oscillation experiments with a sufficiently long baseline, so that the neutrino oscillations take place in the Earth and the Earth matter effects in the oscillations are non-negligible [21, 22, 23]. The ambiguity regarding the type of the neutrino mass spectrum might be resolved by the MINOS experiment [13], although on the baseline of this experiment the matter effects are relatively small [21]. This might be done in an experiment with atmospheric neutrinos, utilizing a detector with a sufficiently good muon charge discrimination [24]. The experiments at neutrino factories would be particularly suitable for the indicated purpose [22, 23].

In the context of three-neutrino oscillations, we study [25] the possibility of using anti-neutrinos from nuclear reactors to explore the $\Delta m^2_{\odot} > 10^{-4}$ eV^2 region of the

LMA MSW solution. Such an experiment might be of considerable interest if, in particular, the results of the KamLAND experiment will confirm the validity of the LMA-MSW solution of the solar neutrino problem, but will allow to obtain only a lower bound on $\Delta m^2_\odot$ due to the fact that $\Delta m^2_\odot > 10^{-4}$ eV^2 [26, 27, 28]. We determine the optimal baseline of the possible experiment with reactor $\bar{\nu}_e$, which would provide a precise measurement of $\Delta m^2_\odot$ in the region 10^{-4} eV^2 $< \Delta m^2_\odot \lesssim 8 \times 10^{-4}$ eV^2. Furthermore, the same experiment might be used to try to distinguish between the two types of neutrino mass spectrum - with normal or with inverted hierarchy. This might be done by exploring the effect of interference between the amplitudes of neutrino oscillations, driven by the solar and atmospheric Δm^2, i.e., by $\Delta m^2_\odot$ and Δm^2_{atm}. For the optimal baseline found earlier, $L \cong (20 - 25)$ km, the indicated effect could be relevant for 10^{-4} eV^2 $< \Delta m^2_\odot \lesssim 5 \times 10^{-4}$ eV^2. For larger values of $\Delta m^2_\odot$ within the interval (4.2), the effect could be relevant at $L \cong 10$ km. Distinguishing between the two possible types of neutrino mass spectrum requires a relatively high precision measurement of the positron spectrum in the reaction $\bar{\nu}_e + p \to e^+ + n$ (i.e., a high statistics experiment with sufficiently good energy resolution), a measurement of Δm^2_{atm} with very high precision, $\sin^2 2\theta_\odot \neq 1.0$, e.g., $\sin^2 2\theta_\odot \lesssim 0.9$, and a sufficiently large value of the angle θ, which for $\Delta m^2_\odot \ll \Delta m^2_{atm}$ controls, e.g., the oscillations of the atmospheric ν_e and $\bar{\nu}_e$ and is constrained by the CHOOZ and Palo Verde data.

In this context, we will focus on theoretical and phenomenological considerations. For what concerns the actual feasibility of the proposed experiment, the search for an appropriate reactor site, the study of the backgrounds, and the discussion of the detector characteristics, the interested reader can find a detailed preparatory study in [29].

4.1 The $\bar{\nu}_e$ survival probability

We shall assume in what follows that the 3-neutrino mixing described by eq. (4.5) takes place. We shall number (without loss of generality) the neutrinos with definite mass in vacuum ν_j, $j = 1, 2, 3$, in such a way that their masses obey $m_1 < m_2 < m_3$. Then the cases of NH and IH neutrino mass spectrum differ, in particular, by the relation between the mixing matrix elements $|U_{ej}|$, $j = 1, 2, 3$, and the mixing angles $\theta_\odot$ and θ (see further). With the indicated choice one has $\Delta m^2_{jk} > 0$ for $j > k$. Let us emphasize that we do not assume any of the relations $m_1 \ll m_2 \ll m_3$, or $m_1 \lesssim m_2 \ll m_3$, or $m_1 \ll m_2 \cong m_3$, to be valid in what follows.

Under the conditions of the experiment we are going to discuss, which must have a baseline L considerably shorter than the baseline ~ 180 km of the KamLAND experiment, the reactor $\bar{\nu}_e$ oscillations will not be affected by Earth matter effects when the $\bar{\nu}_e$ travel between the source (reactor) and the detector. If 3-neutrino mixing takes place, eq. (4.5), the $\bar{\nu}_e$ would take part in 3-neutrino oscillations in vacuum on the way to the detector.

We shall obtain next the expressions for the reactor $\bar{\nu}_e$ survival probability of interest in terms of measurable quantities for the two types of neutrino mass spectrum. In the case of normal hierarchy between the neutrino masses we have:

$$\Delta m^2_\odot = \Delta m^2_{21}, \tag{4.6}$$

and

$$|U_{e1}| = \cos\theta_\odot \sqrt{1 - |U_{e3}|^2}, \quad |U_{e2}| = \sin\theta_\odot \sqrt{1 - |U_{e3}|^2}, \tag{4.7}$$

where

$$\theta_\odot = \theta_{12}, \quad |U_{e3}|^2 = \sin^2\theta \equiv \sin^2\theta_{13}, \tag{4.8}$$

θ_{12} and θ_{13} being two of the three mixing angles in the standard parameterization of the PMNS matrix (see, e.g., [16]). Note that $|U_{e3}|^2$ is constrained by the CHOOZ and Palo Verde results. It is not difficult to derive the expression for the $\bar{\nu}_e$ survival probability in the case under discussion:

$$\begin{aligned} & P_{NH}(\bar{\nu}_e \to \bar{\nu}_e) \\ = \;& 1 - 2\sin^2\theta\cos^2\theta \left(1 - \cos\frac{\Delta m^2_{31}\, L}{2\, E_\nu}\right) \\ - \;& \frac{1}{2}\cos^4\theta\,\sin^2 2\theta_\odot \left(1 - \cos\frac{\Delta m^2_\odot\, L}{2\, E_\nu}\right) \\ + \;& 2\sin^2\theta\,\cos^2\theta\,\sin^2\theta_\odot \left(\cos\left(\frac{\Delta m^2_{31}\, L}{2\, E_\nu} - \frac{\Delta m^2_\odot\, L}{2\, E_\nu}\right) - \cos\frac{\Delta m^2_{31}\, L}{2\, E_\nu}\right), \end{aligned} \tag{4.9}$$

where E_ν is the neutrino energy and we have made use of eqs. (4.6), (4.7) and (4.8).

If the neutrino mass spectrum is with inverted hierarchy one has (see, e.g., [30, 20, 16]):

$$\Delta m^2_\odot = \Delta m^2_{32}, \tag{4.10}$$

and

$$|U_{e2}| = \cos\theta_\odot \sqrt{1 - |U_{e1}|^2}, \quad |U_{e3}| = \sin\theta_\odot \sqrt{1 - |U_{e1}|^2}. \tag{4.11}$$

The mixing matrix element constrained by the CHOOZ and Palo Verde data is now $|U_{e1}|^2$:

$$|U_{e1}|^2 = \sin^2\theta. \tag{4.12}$$

The expression for the $\bar{\nu}_e$ survival probability can be written in the form [31]:

$$\begin{aligned} & P_{IH}(\bar{\nu}_e \to \bar{\nu}_e) \\ = \;& 1 - 2\sin^2\theta\cos^2\theta \left(1 - \cos\frac{\Delta m^2_{31}\, L}{2\, E_\nu}\right) \\ - \;& \frac{1}{2}\cos^4\theta \sin^2 2\theta_\odot \left(1 - \cos\frac{\Delta m^2_\odot\, L}{2\, E_\nu}\right) \\ + \;& 2\sin^2\theta\cos^2\theta\cos^2\theta_\odot \left(\cos\left(\frac{\Delta m^2_{31}\, L}{2\, E_\nu} - \frac{\Delta m^2_\odot\, L}{2\, E_\nu}\right) - \cos\frac{\Delta m^2_{31}\, L}{2\, E_\nu}\right). \end{aligned} \tag{4.13}$$

Several comments concerning the expressions for the $\bar{\nu}_e$ survival probability, eqs. (4.9) and (4.13), follow. In the first lines in the right-hand side of eqs. (4.9) and (4.13), the oscillations of the electron (anti-)neutrino driven by the "atmospheric" Δm^2_{31} are accounted for. The CHOOZ and Palo Verde experiments are primarily sensitive to this term and their results limit $\sin^2\theta$. The second lines in the expressions in eqs. (4.9) and (4.13) contain the solar neutrino oscillation parameters. This is the term KamLAND should be most sensitive to. For $\Delta m^2_\odot \ll \Delta m^2_{31} = \Delta m^2_{atm}$, $\Delta m^2_\odot \lesssim 10^{-4}$ eV^2, only one of the indicated two terms leads to an oscillatory dependence of the $\bar{\nu}_e$ survival probability for the ranges of L/E_ν characterizing the CHOOZ and Palo Verde, and the KamLAND experiments: on the source-detector distance L of the CHOOZ and Palo Verde experiments the oscillations due to $\Delta m^2_\odot$ cannot develop, while on the distance(s) traveled by the $\bar{\nu}_e$ in the KamLAND experiment Δm^2_{atm} causes fast oscillations which average out and are not predicted to lead, e.g., to specific spectrum distortions of the KamLAND event rate.

The terms in the third lines in eqs. (4.9) and (4.13) are not present in any two-neutrino oscillation analysis. They represent interference terms between the amplitudes of neutrino oscillations, driven by the solar and atmospheric neutrino mass squared differences. The term in eq. (4.9) is proportional to $\sin^2\theta_\odot$, while the corresponding term in eq. (4.13) is proportional to $\cos^2\theta_\odot$ [31]. This is the only difference between $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e)$ and $P_{IH}(\bar{\nu}_e \to \bar{\nu}_e)$, that can be used to distinguish between the two cases of neutrino mass spectrum in an experiment with reactor $\bar{\nu}_e$. Obviously, if $\cos 2\theta_\odot = 0$, we have $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e) = P_{IH}(\bar{\nu}_e \to \bar{\nu}_e)$ and the two types of spectrum would be indistinguishable in the experiments under discussion. For vanishing $\sin^2\theta$, only the terms in the second line of eqs. (4.9) and (4.13) survive, and the two-neutrino mixing formula for solar neutrino oscillations in vacuum is exactly reproduced.

Let us discuss next the ranges of values the different oscillation parameters, which enter into the expressions for the probabilities of interest $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e)$ and $P_{IH}(\bar{\nu}_e \to \bar{\nu}_e)$, can take. The allowed region of values of Δm^2_{31}, $\Delta m^2_\odot$, $\sin^2\theta_\odot$ and θ should be determined in a global 3-neutrino oscillation analysis of the solar, atmospheric and reactor neutrino oscillation data, in which, in particular, $\Delta m^2_\odot$ should be allowed to take values in the LMA solution region, including the interval $\Delta m^2_\odot \sim (1.0-6.0)\times 10^{-4}$ eV^2. Such an analysis is lacking in the literature. However, as was shown in [32], a global analysis of the indicated type would not change essentially the results for the LMA MSW solution we have quoted [2] in eqs. (4.2) and (4.3) as long as $\Delta m^2_{31} \gtrsim 1.5\times 10^{-3}$ eV^2. The reason is that for $\Delta m^2_{31} \gtrsim 1.5\times 10^{-3}$ eV^2 and $\Delta m^2_\odot \lesssim 6.0\times 10^{-4}$ eV^2, the solar ν_e survival probability, which determines the level of suppression of the solar neutrino flux and plays a major role in the analyses of the solar neutrino data, depends very weakly on (i.e., is practically independent of) Δm^2_{31}. Thus, $\Delta m^2_\odot$ and $\theta_\odot$ are uniquely determined by the solar

[2] Let us note that the LMA MSW solution values of $\Delta m^2_\odot$ and $\theta_\odot$ we quote in eqs. (4.2) and (4.3) were obtained by taking into account the CHOOZ and Palo Verde limits as well.

neutrino and CHOOZ and Palo Verde data, independently of the atmospheric neutrino data and of the type of the neutrino mass spectrum. The CHOOZ and Palo Verde data lead to an upper limit on $\Delta m^2_\odot$ in the LMA MSW solution region (see, e.g., [6, 33]): $\Delta m^2_\odot \lesssim 7.5 \times 10^{-4}$ eV^2. For $\Delta m^2_\odot \lesssim 1.0 \times 10^{-4}$ eV^2, the CHOOZ and solar neutrino data imply the upper limit on $\sin^2\theta$ given in eq. (4.4). For $\Delta m^2_\odot \sim (2.0 - 6.0) \times 10^{-4}$ eV^2 of interest, the upper limit on $\sin^2\theta$ as a function of $\Delta m^2_{31} \gtrsim 10^{-3}$ eV^2 for given $\Delta m^2_\odot$ and $\sin^2 2\theta_\odot$ is somewhat more stringent [31].

Would a global 3-neutrino oscillation analysis of the solar, atmospheric and reactor neutrino oscillation data lead to drastically different results for Δm^2_{31} in the two cases of normal and inverted neutrino mass hierarchy? Our preliminary analysis shows that given the existing atmospheric neutrino data from the Super-Kamiokande experiment, such an analysis i) would not be able to discriminate between the two cases of neutrino mass spectrum, and ii) would give essentially the same allowed region for Δm^2_{31} in the two cases of neutrino mass spectrum. We expect the regions of allowed values of the mixing angle θ_{atm}, which controls the dominant atmospheric $\nu_\mu \to \nu_\tau$ and $\bar{\nu}_\mu \to \bar{\nu}_\tau$ oscillations, to differ somewhat in the two cases. Note, however, that this mixing angle does not enter the expression for the $\bar{\nu}_e$ survival probability we are interested in.

For $\Delta m^2_\odot \lesssim 1.0 \times 10^{-4}$ eV^2 and sufficiently small values of $\sin^2\theta$, Δm^2_{31} coincides effectively with Δm^2_{atm} of the two-neutrino ν_μ and $\bar{\nu}_\mu$ oscillation analyses of the SK atmospheric neutrino data. If $\sin^2\theta > 0.01$, a three-neutrino oscillation analysis of the atmospheric neutrino and CHOOZ data, performed under the assumption of $\Delta m^2_\odot \lesssim 1.0 \times 10^{-4}$ eV^2 [33], gives regions of allowed values of $\Delta m^2_{atm} = \Delta m^2_{31}$, which are correlated with the value of $\sin^2\theta$. The latter must satisfy the CHOOZ and Palo Verde constraints.

At present, as we have already indicated, a complete three-neutrino oscillation analysis of the atmospheric neutrino and CHOOZ data with $\Delta m^2_\odot$ allowed to take values up to $\sim (6.0 - 7.0) \times 10^{-4}$ eV^2, i.e., in the region where deviations from the two-neutrino approximation could be non-negligible, is lacking in the literature. Therefore in what follows we will use representative values of Δm^2_{31} which lie in the region given by eq. (4.1).

4.1.1 The Difference between $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e)$ and $P_{IH}(\bar{\nu}_e \to \bar{\nu}_e)$

Let us discuss next in greater detail the difference between the $\bar{\nu}_e$ surviving probabilities in the two cases of neutrino mass spectrum of interest, $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e)$ and $P_{IH}(\bar{\nu}_e \to \bar{\nu}_e)$. While the terms in the first two lines in eqs. (4.9) and (4.13) describe oscillations in L/E_ν with frequencies $\Delta m^2_{31}/4\pi$ and $\Delta m^2_{\odot}/4\pi$, respectively, the third term has the shape of beats, being produced by the interference of two waves, with the same amplitude but slightly different frequencies:

$$\begin{aligned} \cos\left(\frac{\Delta m^2_{31}\, L}{2\, E_\nu} - \frac{\Delta m^2_{\odot}\, L}{2\, E_\nu}\right) - \cos\frac{\Delta m^2_{31}\, L}{2\, E_\nu} \\ &= 2\,\sin\frac{\Delta m^2_{\odot}\, L}{4\, E_\nu}\,\sin\left(\frac{\Delta m^2_{31}\, L}{2\, E_\nu} - \frac{\Delta m^2_{\odot}\, L}{4\, E_\nu}\right) \\ &\simeq 2\,\sin\frac{\Delta m^2_{\odot}\, L}{4\, E_\nu}\,\sin\left(\frac{\Delta m^2_{31}\, L}{2\, E_\nu}\right) \end{aligned} \tag{4.14}$$

This is a modulated oscillation with approximately the same frequency of the first term in eqs. (4.9) and (4.13) ($\Delta m^2_{31}/4\pi$) and amplitude oscillating between 0 and $2\sin^2\theta_{\odot}$ of the amplitude of the first term itself. The beat frequency is equal to the frequency of the dominant oscillation ($\Delta m^2_{\odot}/4\pi$). The modulation is exactly in phase with the $\Delta m^2_{\odot}$−driven dominant oscillation of interest, so that the maximum of the oscillation amplitude of the interference term (third lines in the expressions for $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e)$ and $P_{IH}(\bar{\nu}_e \to \bar{\nu}_e)$) is reached in coincidence with the points of maximal decreasing of the $\bar{\nu}_e$ survival probability, where $\Delta m^2_{\odot}\, L/4\, E = \pi/2$, and vice versa - this amplitude vanishes at the local maxima of the survival probability. At the minima of the $\bar{\nu}_e$ survival probability, for instance at $\Delta m^2_{\odot}\, L/4\, E_\nu = \pi/2$, $P^{NH(IH)}\,(\bar{\nu}_e \to \bar{\nu}_e)$ takes the value:

$$\begin{aligned} P^{NH(IH)}\,(\bar{\nu}_e \to \bar{\nu}_e)\Big|_{\frac{\Delta m^2_{\odot} L}{2\pi E_\nu}=1} &= 1 - 2\,\sin^2\theta\,\cos^2\theta - \cos^4\theta\,\sin^2 2\theta_{\odot} \\ &\quad \overset{(+)}{-}\cos 2\theta_{\odot}\, 2\,\sin^2\theta\,\cos^2\theta\,\cos\,\pi\frac{\Delta m^2_{31}}{\Delta m^2_{\odot}}\,. \end{aligned} \tag{4.15}$$

From eqs. (4.9), (4.13) and (4.15) one deduces that:

- for maximal mixing, $\cos 2\theta_\odot = 0$, the last term cancels, and $P^{NH} = P^{IH}$;

- for very small mixing angles, $\cos 2\theta_\odot \simeq 1$, the terms describing the oscillations driven by Δm^2_{31} in the NH and IH cases have opposite signs: the two waves are exactly out of phase.

- for intermediate values of $\cos 2\theta_\odot$ from the LMA MSW solution region, $\cos 2\theta_\odot \cong (0.3-0.6)$, the $\Delta m^2_{31}-$driven contributions in the cases of normal and inverted hierarchy have still opposite signs and the magnitude of the effect is proportional to $2\cos 2\theta_\odot \sin^2\theta$.

The net result of these properties is that in the region of the minima of the $\bar{\nu}_e$ survival probability due to $\Delta m^2_\odot$, where $\Delta m^2_\odot\, L/(2E) = \pi(2k+1)$, $k = 0, 1, \ldots$, the difference between $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e)$ and $P_{IH}(\bar{\nu}_e \to \bar{\nu}_e))$ is maximal. In contrast, at the maxima of $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e)$ and $P_{IH}(\bar{\nu}_e \to \bar{\nu}_e))$ determined by $\Delta m^2_\odot\, L/(2E) = 2\pi k$, we have, for any $\sin^2\theta_\odot$, $P_{NH}(\bar{\nu}_e \to \bar{\nu}_e) = P_{IH}(\bar{\nu}_e \to \bar{\nu}_e)$.

The two-neutrino oscillation approximation used in the analysis of the CHOOZ and Palo Verde data is rather accurate as long as $\Delta m^2_\odot$ is sufficiently small [31]: for $\Delta m^2_\odot \lesssim 10^{-4}$ eV^2, the L/E_ν values characterizing these experiments, chosen to ensure maximal sensitivity to $\Delta m^2_{31} \gtrsim 10^{-3}$ eV^2, are much smaller than the value at which the first minimum of $P_{NH(IH)}(\bar{\nu}_e \to \bar{\nu}_e)$ due to the $\Delta m^2_\odot$-dependent oscillating term occurs. Correspondingly, the effect of the interference term is strongly suppressed by the beats. For $\Delta m^2_\odot \gtrsim 2\times 10^{-4}$ eV^2 this is no longer valid and the interference term under discussion has to be taken into account in the analyses of the CHOOZ and Palo Verde data [31].

4.2 Measuring $\Delta m^2_\odot$ at Reactor Facilities

As is well-known, nuclear reactors are intense sources of low energy $\bar{\nu}_e$ ($E_\nu \lesssim 8$ MeV), emitted isotropically in the β-decays of fission products with high neutron density [34]. Anti-neutrinos can then be detected through the positrons produced by inverse β-decay on nucleons. The reactor $\bar{\nu}_e$ energy spectrum has been accurately measured and is theoretically well understood [3] [35]: it essentially consists of a bell-shaped distribution in energy centered around $E_\nu \sim 4$ MeV, having a width of approximately 3 MeV. CHOOZ, Palo Verde and KamLAND are examples of experiments with reactor $\bar{\nu}_e$, the main difference being the distance between the source and the detector explored ($L \sim 1$ km for CHOOZ and Palo Verde, and $L \sim 180$ km for KamLAND).

The best sensitivity to a given value of $\Delta m^2_\odot$ of the experiment of interest is at L at which the maximum reduction of the survival probability is realized. As can be seen from eqs. (4.9) - (4.13), this happens for L around $L^* \equiv 2\pi\, E_\nu/\Delta m^2_\odot$. This implies that for $E_\nu = 4$ MeV, the optimal length to test neutrino oscillations with reactor experiments is:

$$L^* \cong \frac{5 \times 10^{-3}}{(\Delta m^2_\odot/\mathrm{eV}^2)} \ \mathrm{km} \tag{4.16}$$

The best sensitivity of KamLAND, for instance, is in the range of $2 \div 3 \times 10^{-5}$ eV^2. We will discuss next in greater detail the distances L which could be used to probe the LMA MSW solution region at $\Delta m^2_\odot > 10^{-4}$ eV^2, in order to extract $\Delta m^2_\odot$ from these oscillation experiments.

4.2.1 Total Event Rate Analysis

One of the signatures of the $\bar{\nu}_e$−oscillations would be a substantial reduction of the measured total event rate due to the reactor $\bar{\nu}_e$ in comparison with the predicted one in the absence of oscillations. In order to compute the expected total event rate one has to integrate the $\bar{\nu}_e$ survival probability multiplied by the $\bar{\nu}_e$ energy spectrum over E_ν. In Fig. 4.1 we show this averaged survival probability for different values

[3] By reactor $\bar{\nu}_e$ energy spectrum we mean here and in what follows the product of the $\bar{\nu}_e$ production spectrum and the inverse β-decay cross-section, which gives the "detected" neutrino spectrum in the no oscillation case. The $\bar{\nu}_e$ production spectrum is known with larger uncertainties at $\bar{\nu}_e$ energies $E_\nu \lesssim 2$ MeV, but this range is not of interest due to the threshold energy $E^{th}_\nu \cong 1.8$ MeV of the inverse β-decay reaction [36]. Certain known time dependence at the level of a few percent is also present up to 3.5 MeV [37] and should possibly be taken into account in the analysis of the experimental data.

of L as a function of $\Delta m^2_{\odot}$, using the "best fit" values [1, 5, 6, 7] for Δm^2_{31} and $\sin^2 2\theta_{\odot}$.

When averaging over the $\bar{\nu}_e$ energy spectrum, oscillatory effects with too short a period are washed out, and the experiment is sensitive only to the average amplitude. This happens when the width δE_{ν} of the energy spectrum is such that the integration runs over more than one period, i.e., approximately for:

$$\delta E_{\nu} \gtrsim \frac{4\pi\, E^2_{\nu}}{\Delta m^2\, L} \simeq \frac{4 \times 10^4\, \mathrm{eV}^3}{\Delta m^2\, (L/Km)}. \tag{4.17}$$

Since $\delta E_{\nu} \sim 3$ MeV, at KamLAND this happens approximately for $\Delta m^2_{\odot} \gtrsim 7 \times 10^{-5}$ eV^2. The corresponding curve in Fig. 4.1 indicates that the actual sensitivity extends to somewhat larger values of $\Delta m^2_{\odot}$ than what is expected on the basis on the above estimate, but the total event rate becomes flat for $\Delta m^2_{\odot} \gtrsim 10^{-4}$ eV^2. This means that KamLAND will be able, through the measurement of the total even rate, to test all the region of the LMA MSW solution and determine whether the latter is the correct solution of the solar neutrino problem, but will provide a precise measurement of $\Delta m^2_{\odot}$ only if $\Delta m^2_{\odot} \lesssim 10^{-4}$ eV^2.

If $\Delta m^2_{\odot} \gtrsim 2 \times 10^{-4}$ eV^2, it would be possible to obtain only a lower bound on $\Delta m^2_{\odot}$ and a new experiment might be required to determine $\Delta m^2_{\odot}$. This observation have been confirmed also by a detailed numerical analysis simulating the KamLAND results performed in [38], in the two neutrino approximation, illustrated by Fig. 4.2.

Fig. 4.1 shows that as L decreases, the sensitivity region moves to larger $\Delta m^2_{\odot}$. These results imply that a reactor $\bar{\nu}_e$ experiment with $L \cong (20 - 25)$ km can probe the range $0.8 \times 10^{-4}\,\mathrm{eV}^2 < \Delta m^2_{\odot} \lesssim 6 \times 10^{-4}\,\mathrm{eV}^2$. One finds that for $\Delta m^2_{\odot} \cong 2 \times 10^{-4}$ eV^2 and $\Delta m^2_{31} \cong 2.5 \times 10^{-3}$ eV^2, the best sensitivity is at $L \cong 20$ km. Moreover, with $L \cong (20 - 25)$ km, the predicted total event rate deviates from being flat (in $\Delta m^2_{\odot}$) actually for $\Delta m^2_{\odot}$ as large as $\sim (5 - 6) \times 10^{-4}$ eV^2. In order to have a precise determination of $\Delta m^2_{\odot}$ with $L \cong (20 - 25)$ km for the largest values given in eq. (4.2), $\Delta m^2_{\odot} \cong (7 \div 8) \times 10^{-4}$ eV^2, one should use the information about the e^+−spectrum distortion due to the $\bar{\nu}_e$−oscillations. By measuring the e^+−spectrum

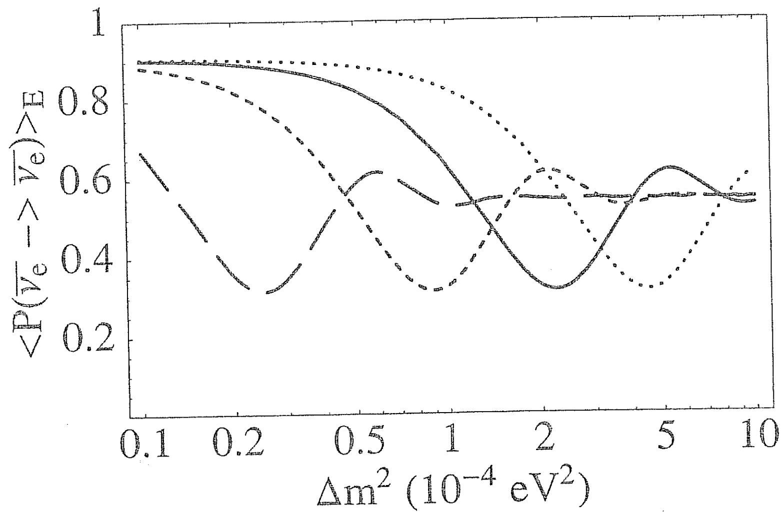

Figure 4.1: The reactor $\bar{\nu}_e$ survival probability, averaged over the $\bar{\nu}_e$ energy spectrum, for $\Delta m^2_{31} = 2.5 \times 10^{-3}$ eV^2, $\sin^2 2\theta_\odot = 0.8$, $\sin^2\theta = 0.05$, as a function of $\Delta m^2_\odot$. The curves correspond to $L = 180$ km (long dashed), $L = 50$ km (dashed), $L = 20$ km (thick) and $L = 10$ km (dotted), respectively.

with a sufficient precision it would be possible to cover the whole interval

$$1.0 \times 10^{-4}\,\text{eV}^2 \lesssim \Delta m^2_\odot \lesssim 8.0 \times 10^{-4}\,\text{eV}^2 \,, \tag{4.18}$$

i.e., to determine $\Delta m^2_\odot$ if it lies in this interval, by performing an experiment at $L \cong (20 - 25)$ km from the reactor(s) [4] (see the next sub-section).

Applying eq. (17) with $\Delta m^2 = \Delta m^2_{31}$, one sees that for the ranges of L which allow to probe $\Delta m^2_\odot$ from the LMA MSW solution region, the total event rate is not sensitive to the oscillations driven by $\Delta m^2_{31} \gtrsim 1.5 \times 10^{-3}$ eV^2. Thus, the total event rate analysis would determine $\Delta m^2_\odot$ which would be the same for both the normal and inverted hierarchy neutrino mass spectrum.

[4]The fact that if $\Delta m^2_\odot \cong 3.2 \times 10^{-4}$ eV^2, a reactor $\bar{\nu}_e$ experiment with $L \cong 20$ km would allow to measure $\Delta m^2_\odot$ with a high precision was also noticed recently in [28].

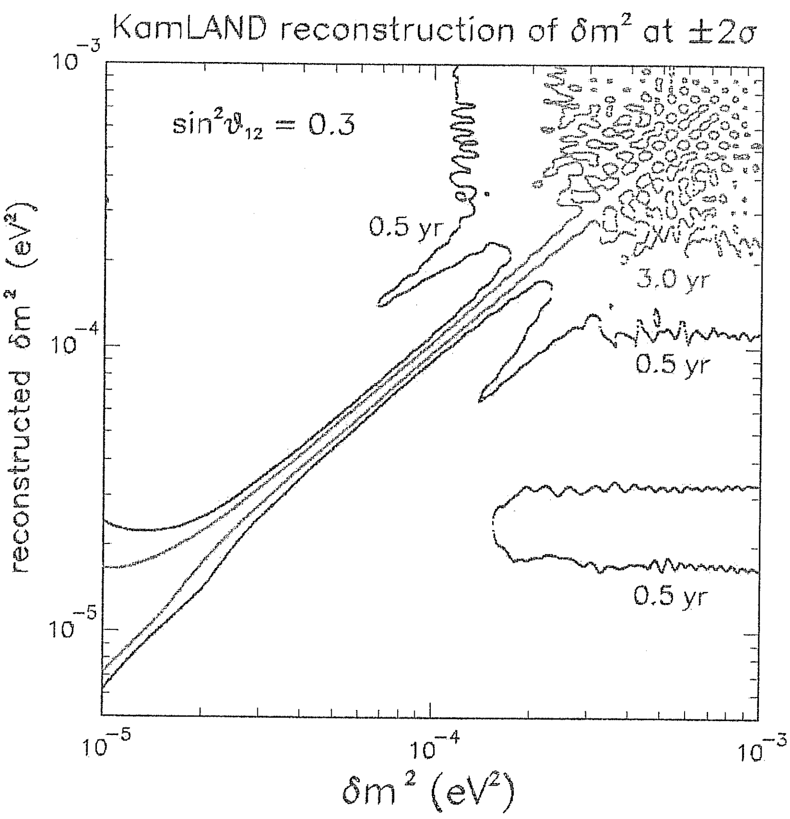

Figure 4.2: Simulation of KamLAND results from the analysis in [38]. Reconstructed range of δm^2 at $\pm 2\sigma$, as a function of the true value of δm^2 in abscissa. The true value of $\sin^2\theta_{12}$ is fixed at 0.3. The curves refer to detector exposures of 0.5 and 3 years.

4.2.2 Energy Spectrum Distortions

An unambiguous evidence of neutrino oscillations would be the characteristic distortion of the $\bar{\nu}_e$ energy spectrum. This is caused by the fact that, at fixed L, neutrinos with different energies reach the detector in a different oscillation phase, so that some parts of the spectrum would be suppressed more strongly by the oscillations than other parts. The search for distortions of the $\bar{\nu}_e$ energy spectrum is essentially a direct test of the $\bar{\nu}_e$ oscillations. It is more effective than the total rate analysis since it is not affected, e.g., by the overall normalization of the reactor $\bar{\nu}_e$ flux. However, such a test requires a sufficiently high statistics and sufficiently good energy resolution of the detector used.

Energy spectrum distortions can be studied, in principle, in an experiment with $L \cong (20-25)$ km. In Fig. 4.3 we show the comparison between the $\bar{\nu}_e$ spectrum expected for $\Delta m^2_\odot = 2 \times 10^{-4}$ eV^2 and $\Delta m^2_\odot = 6 \times 10^{-4}$ eV^2 and the spectrum in the absence of $\bar{\nu}_e$ oscillations. No averaging has been performed and the possible detector resolution is not taken into account. The curves show the product of the probabilities given by eqs. (4.9) and (4.13) and the predicted reactor $\bar{\nu}_e$ spectrum [39].

As Fig. 4.3 illustrates, the $\bar{\nu}_e$ spectrum in the case of oscillation is well distinguishable from that in the absence of oscillations. Moreover, for $\Delta m^2_\odot$ lying in the interval 10^{-4} eV^2 $< \Delta m^2_\odot \lesssim 8.0 \times 10^{-4}$ eV^2, the shape of the spectrum exhibits a very strong dependence on the value of $\Delta m^2_\odot$. A likelihood analysis of the data would be able to determine the value of $\Delta m^2_\odot$ from the indicated interval with a rather good precision. This would require a precision in the measurement of the e^+–spectrum, which should be just not worse than the precision achieved in the CHOOZ experiment and that planned to be reached in the KamLAND experiment. If the energy bins used in the measurement of the spectrum are sufficiently large, the value of $\Delta m^2_\odot$ thus determined should coincide with value obtained from the analysis of the total event rate and should be independent of Δm^2_{31}.

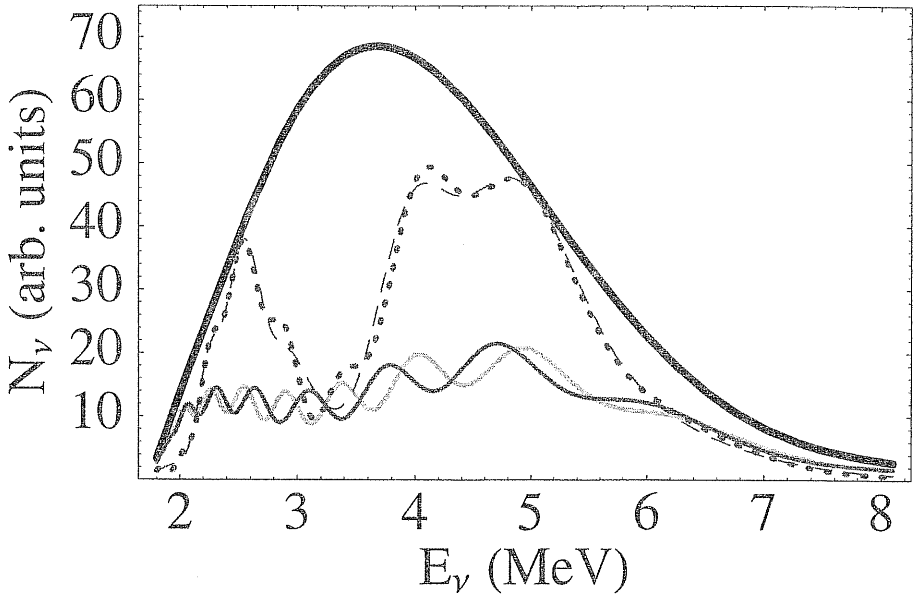

Figure 4.3: The reactor $\bar{\nu}_e$ energy spectrum at distance $L = 20$ km from the source, in the absence of $\bar{\nu}_e$ oscillations (double-thick solid line) and in the case of $\bar{\nu}_e$ oscillations characterized by $\Delta m^2_{31} = 2.5 \times 10^{-3}$ eV^2, $\sin^2 2\theta_\odot = 0.8$ and $\sin^2\theta = 0.05$. The thick lines are obtained for $\Delta m^2_\odot = 2 \times 10^{-4}$ eV^2 and correspond to NH (light grey) and IH (dark grey) neutrino mass spectrum. Shown is also the spectrum for $\Delta m^2_\odot = 6 \times 10^{-4}$ eV^2 in the NH (dotted) and IH (dashed) cases.

4.3 Normal vs. Inverted Hierarchy

In Fig. 4.3 we show the deformation of the reactor $\bar{\nu}_e$ spectrum both for the normal and inverted hierarchy neutrino mass spectrum: as long as no integration over the energy is performed, the deformations in the two cases of neutrino mass spectrum can be considerable, and the sub-leading oscillatory effects driven by the atmospheric mass squared difference (see the first and the third line of eqs. (4.9) - (4.13)) can, in principle, be observed. They could be used to distinguish between the two hierarchical patterns, provided the solar mixing is not maximal [5], $\sin^2\theta$ is not too small and Δm^2_{31} is known with high precision. It should be clear that the possibility we will be discussing next poses remarkable challenges.

The experiment under discussion could be in principle an alternative to the measurement of the sign of Δm^2_{31} in long (very long) baseline neutrino oscillation experiments [21, 22, 23] or in the experiments with atmospheric neutrinos (see, e.g., [24]).

The magnitude of the effect of interest depends, in particular, on three factors, as we have already pointed out:

- the value of the solar mixing angle $\theta_\odot$: the different behavior of the two survival probabilities is due to the difference between $\sin^2\theta_\odot$ and $\cos^2\theta_\odot$; correspondingly, the effect vanishes for maximal mixing; thus, the more the mixing deviates from the maximal the larger the effect;
- the value of $\sin^2\theta$, which controls the magnitude of the sub-leading effects due to Δm^2_{31} on the $\Delta m^2_\odot$−driven oscillations: the effect of interest vanishes in the decoupling limit of $\sin^2\theta \to 0$;

[5] It would be impossible to distinguish between the normal and inverted hierarchy neutrino mass spectrum if for given $\Delta m^2_\odot > 10^{-4}$ eV^2 and $\sin^2 2\theta_\odot \neq 1$, the LMA solution region is symmetric with respect to the change $\theta_\odot \to \pi/2 - \theta_\odot$ ($\cos 2\theta_\odot \to -\cos 2\theta_\odot$). While the value of $\sin^2 2\theta_\odot$ is expected to be measured with a relatively high precision by the KamLAND experiment, the sign of $\cos 2\theta_\odot$ will not be fixed by this experiment. However, the $\theta_\odot - (\pi/2 - \theta_\odot)$ ambiguity can be resolved by the solar neutrino data. Note also that the current solar neutrino data disfavor values of $\cos 2\theta_\odot < 0$ in the LMA solution region (see, e.g., [5, 6, 10]).

- the value of $\Delta m^2_\odot$ (see Fig. 4.1): for given L and $\Delta m^2_\odot$ the difference between the spectrum in the cases of normal and inverted hierarchy is maximal at the minima of the survival probability, and vanishes at the maxima.

A rough estimate of the possible difference between the predictions of the event rate spectrum for the two hierarchical patterns, is provided by the ratio between the difference and the sum of the two corresponding probabilities at $\Delta m^2_\odot L = 2\pi E_\nu$:

$$\frac{P_{NH} - P_{IH}}{P_{NH} + P_{IH}} = \frac{2\,\cos 2\theta_\odot\,\sin^2\theta\,\cos^2\theta}{1 - 2\sin^2\theta\,\cos^2\theta - \cos^4\theta\,\sin^2 2\theta_\odot}\,\cos\pi\,\frac{\Delta m^2_{31}}{\Delta m^2_\odot}\,. \tag{4.19}$$

The ratio could be rather large: the factor in front of the $\cos\pi\,\Delta m^2_{31}/\Delta m^2_\odot$ is about 25% for $\sin^2 2\theta_\odot = 0.8$ and $\sin^2\theta = 0.05$.

The actual feasibility of the study under discussion depends crucially on the integration over (i.e., the binning in) the energy: for the effect not to be strongly suppressed, the energy resolution of the detector ΔE_ν must satisfy:

$$\Delta E_\nu \lesssim \frac{4\pi\,E_\nu^2}{\Delta m^2_{31}\,L} \simeq \frac{2 \div 6 \times 10^4\,\mathrm{eV}^3}{\Delta m^2_{31}\,(L/km)}. \tag{4.20}$$

For $L \sim 1$ km this condition could be satisfied for $\delta E_\nu \simeq \Delta E_\nu$, but at $L \cong (15-20)$ Km, for $\Delta m^2_{31} = 2.5 \times 10^{-3}$ eV^2 and E_ν in the interval $(3-5)$ MeV, one should have $\Delta E_\nu \lesssim 0.5$ MeV.

Our discussion so far was performed for simplicity in terms of the reactor $\bar{\nu}_e$ energy spectrum, while in the experiments of interest one measures the energy of the positron emitted in the inverse β-decay, E_e. The relation between E_e and E_ν is well known (see for instance [39]), and, up to corrections of at most few per cent, consists just in a shift due to the threshold energy of the process: $E_\nu \cong E_e + (E_\nu^{th} - m_e)$. The maximal ΔE_ν allowed in order to make the effect observable can be then directly compared to the experimental positron energy resolution ΔE_e [6].

For $\Delta m^2_\odot \lesssim 10^{-4}$ eV^2, the first (most significant) minimum of the survival probability can be explored if $L \sim 180$ km. In this case, due to the bigger distance L,

[6]In the CHOOZ experiment, for instance, the binning in E_e was $\Delta E_e \simeq 0.40$ MeV [11]. KamLAND is expected to have a resolution better than $\Delta E_e/E_e = 10\%/\sqrt{E_e}$, where E_e is in MeV [40]
.

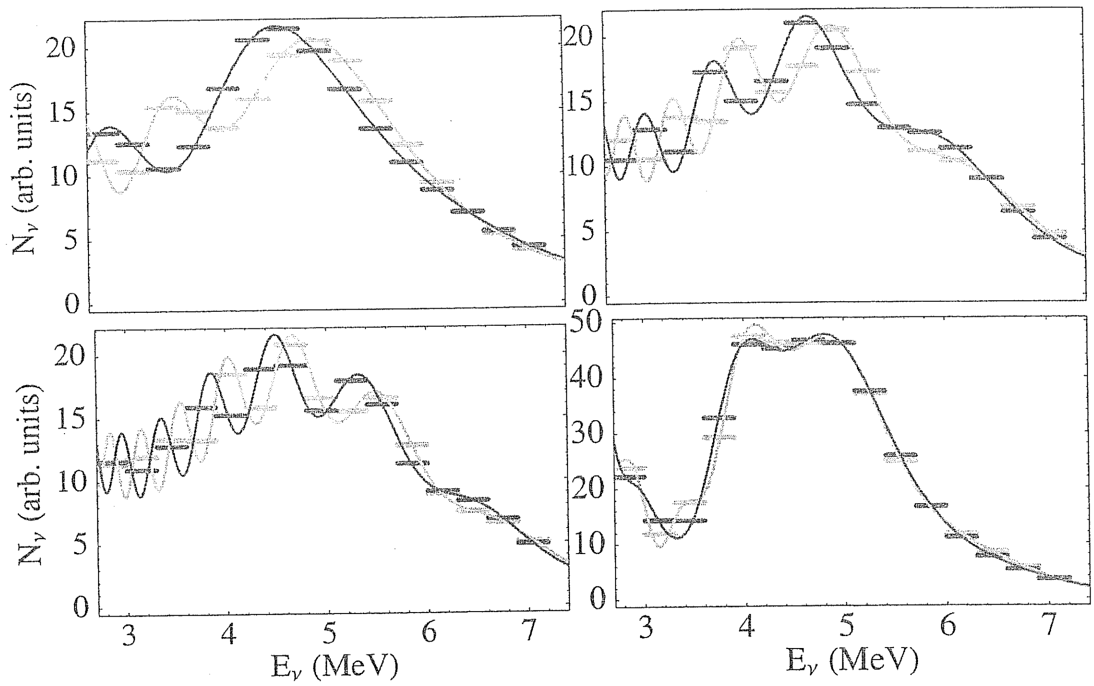

Figure 4.4: Comparison between the predicted event rate spectrum at $L = 20$ km, measured in energy bins having a width of $\Delta E_\nu = 0.3$ MeV in the cases of normal (light grey) and inverted (dark grey) neutrino mass hierarchy. The two upper and the lower left figures are for $\Delta m^2_\odot = 2 \times 10^{-4}$ eV^2, $\sin^2 2\theta_\odot = 0.8$, $\sin^2\theta = 0.05$, and $\Delta m^2_{31} = 1.3;\ 2.5;\ 3.5 \times 10^{-3}$ eV^2, respectively. The lower right figure was obtained for $\Delta m^2_\odot = 6 \times 10^{-4}$ eV^2 and $\Delta m^2_{31} = 2.5 \times 10^{-3}$ eV^2.

the energy resolution required would be by a factor of ten smaller. This means that for $\Delta m^2_\odot \ll \Delta m^2_{31}$, it is practically impossible to realize the condition of maximization of the difference between the survival probabilities in the two cases of neutrino mass spectrum without strongly suppressing the magnitude of the difference by the binning of the energy spectrum.

In order to illustrate what are the concrete possibilities in the case of the experiment under discussion, we have divided the energy interval 2.7 MeV $< E_\nu <$ 7.2 MeV into 15 bins, with $\Delta E_\nu = 0.3$ MeV, and calculated the value of the product of the survival probability and the energy spectrum in each of the bins. The results are shown in Fig. 4.4.

As our results show and Fig. 4.4 indicates, for $\Delta m^2_{31} \cong (1.5 - 3.0) \times 10^{-3}$ eV^2,

$$\Delta m^2_{\odot} \cong (2.0 - 5.0) \times 10^{-4} \text{ eV}^2, \tag{4.21}$$

$\sin^2 2\theta_{\odot} \cong 0.8$ and $\sin^2\theta \cong (0.02 - 0.05)$, it might be possible to distinguish the two cases of neutrino mass spectrum by a high precision measurement of the positron energy spectrum in an experiment with reactor $\bar{\nu}_e$ with a baseline of $L \cong (20 - 25)$ km. This should be a high statistics experiment (not less than about 2000 $\bar{\nu}_e-$induced events per year) with a sufficiently good energy resolution [7]. For larger values of $\Delta m^2_{\odot}$ not exceeding 8.0×10^{-4} eV^2, and $\Delta m^2_{31} \cong (1.5 - 3.0) \times 10^{-3}$ eV^2, the experiment should be done with a smaller baseline, $L \cong 10$ km. If, however, $\sin^2\theta \lesssim 0.01$, and/or $\sin^2 2\theta_{\odot} \gtrsim 0.9$, and/or $\sin^2 2\theta_{\odot} \lesssim 0.9$ but the LMA solution admits equally positive and negative values of $\cos 2\theta_{\odot}$, the difference between the spectra in the two cases becomes hardly observable. Further, in obtaining Fig. 4.4 we have implicitly assumed that Δm^2_{31} is known with negligible uncertainty. Actually, for the difference between the spectra under discussion to be observable, Δm^2_{31} has to be determined, according to our estimates, with a precision of $\sim 10\%$ or better [8]: given the values of $\Delta m^2_{\odot}$, $\sin^2 2\theta_{\odot}$ and $\sin^2\theta$, a spectrum in the NH case corresponding to a given Δm^2_{31} can be rather close in shape to the spectrum in the IH case for a different value of Δm^2_{31}. There is no similar effect when varying $\Delta m^2_{\odot}$.

---

[7] Preliminary estimates show that a detector of the type of KamLAND and a system of nuclear reactors with a total power of approximately 5 - 6 GW might produce the required statistics and precision in the measurement of the positron spectrum.

[8] The analysis, e.g, of the MINOS potential for a high precision determination of Δm^2_{31} in the case of $\Delta m^2_{\odot} \lesssim 10^{-4}$ eV^2 yields very encouraging results (see, e.g., [41]). For $\Delta m^2_{\odot} \cong (2.0 - 8.0) \times 10^{-4}$ eV^2, such analysis is lacking in the literature.

Bibliography

[1] Super-Kamiokande Collaboration, Y. Fukuda *et al.*, Phys. Rev. Lett. **81** (1998) 1562.

[2] Q. R. Ahmad *et al.* [SNO Collaboration], Phys. Rev. Lett. **87** (2001) 071301 (nucl-ex/0106015).

[3] S. Fukuda *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **86** (2001) 5651 (hep-ex/0103032).

[4] V. Barger, D. Marfatia and K. Whisnant, hep-ph/0106207.

[5] G. L. Fogli et al., hep-ph/0106247.

[6] J. N. Bahcall, M. C. Gonzalez-Garcia and C. Pena-Garay, JHEP **0108** (2001) 014 (hep-ph/0106258), and hep-ph/0111150.

[7] A. Bandyopadhyay et al., Phys. Lett. B**519** (2001) 83. (hep-ph/0106264).

[8] P. Creminelli, G. Signorelli and A. Strumia, hep-ph/0102234, version 2 from 10 July, 2001.

[9] M.V. Garzelli and C. Giunti, hep-ph/0108191.

[10] P. I. Krastev and A. Y. Smirnov, hep-ph/0108177.

[11] M. Apollonio *et al.* [CHOOZ Collaboration], Phys. Lett. B**466** (1999) 415 (hep-ex/9907037).

[12] F. Boehm et al., Phys. Rev. Lett. **84** (2000) 3764, Phys. Rev. D**62** (2000) 072002 and Phys. Rev. D**64** (2001) 112001.

[13] MINOS Collaboration, P. Adamson *et al.*, NuMI-L-476 (March 1999); Web page: http://www.hep.anl.gov/ndk/hypertext/numi.html.

[14] A. Piepke [KamLAND Collaboration], Nucl. Phys. Proc. Suppl. **91** (2001) 99.

[15] S.M. Bilenky et al., Phys. Lett. B **356** (1995) 273 and Phys. Rev. D **54** (1996) 1881.

[16] S. M. Bilenky, C. Giunti and W. Grimus, Prog. Part. Nucl. Phys. **43** (1999) 1 (hep-ph/9812360).

[17] B. Pontecorvo, Zh. Eksp. Teor. Fiz. **33** (1957) 549, and **34** (1958) 247.

[18] Z. Maki, M. Nakagawa and S. Sakata, Prog. Theor. Phys. 28 (1962) 870.

[19] S.M. Bilenky and S.T. Petcov, Rev. Mod. Phys. 59 (1987) 671.

[20] S. M. Bilenky, S. Pascoli and S. T. Petcov, Phys. Rev. D **64** (2001) 053010 (hep-ph/0102265).

[21] P. Lipari, Phys. Rev. D**61**, 113004 (2000).

[22] V. Barger et al., Phys. Rev. D**62** (2000) 013004.

[23] M. Freund et al., Nucl. Phys. **B578** (2000) 27.

[24] M.C. Banuls, G. Barenboim and J. Bernabeu, Phys. Lett. B531 (2001) 391; J. Bernabeu et al., hep-ph/0110071

[25] S. T. Petcov and M. Piai, Phys. Lett. B **533**, 94 (2002) [arXiv:hep-ph/0112074].

[26] R. Barbieri and A. Strumia, hep-ph/0011307.

[27] S. Schoenert, Talk given at the TAUP2001 International Workshop, September 8 - 12, 2001, Laboratori Nazionali del Gran Sasso, Italy.

[28] A. Strumia and F. Vissani, hep-ph/0109172.

[29] S. Schonert, T. Lasserre and L. Oberauer, experiment," arXiv:hep-ex/0203013.

[30] S.M. Bilenky et al., Phys. Rev. D **54** (1996) 4432 (hep-ph/9604364).

[31] S. M. Bilenky, D. Nicolo and S. T. Petcov, hep-ph/0112216.

[32] G. L. Fogli. E. Lisi and A. Palazzo, hep-ph/0105080.

[33] M. C. Gonzalez-Garcia et al., Phys. Rev. D **63** (2001) 033005 (hep-ph/0009350).

[34] B. Pontecorvo, Chalk River Lab. Report PD-205 (1946).

[35] For reviews see: G. Gratta, Nucl. Phys. Proc. Suppl. **85**, 72 (2000) (hep-ex/9905011), and
C. Bemporad, G. Gratta and P. Vogel, hep-ph/0107277.

[36] V. I. Kopeikin, hep-ph/0110030.

[37] V. I. Kopeikin, L. A. Mikaelyan and V. V. Sinev, Phys. Atom. Nucl. **64**, 849 (2001) [Yad. Fiz. **64**, 914 (2001)] (hep-ph/0110290).

[38] G. L. Fogli, G. Lettera, E. Lisi, A. Marrone, A. Palazzo and A. Rotunno, arXiv:hep-ph/0208026.

[39] See, e.g., C. Bemporad, G. Gratta and P. Vogel in [35]

[40] M. C. Gonzalez-Garcia and C. Pena-Garay, hep-ph/0111432, and references therein.

[41] V. Barger et al., hep-ph/0110393.

Acknowledgments

I acknowledge my Supervisors, Marco Fabbrichesi, Antonio Masiero and Serguey T. Petcov, for guiding and stimolating my research activity with useful suggestions and constant encouragement.

I thank for fruitful collaboration Matteo Marsili, Roberto Percacci, Andrea Romanino, Marco Serone, Luca Silvestrini, Gianmassimo Tasinato and Oscar Vives, and for useful discussions Stefano Bertolini, Loriano Bonora, Gia Dvali, Enrico Lunghi, Adam Schwimmer and Alexei Yu. Smirnov.

I thank all of my colleagues and friends in Sissa, in particular Silvia Pascoli, for the beautiful atmosphere they created during my Ph. D. course in Trieste.

Special thanks to my parents and my sister, for always supporting and encouraging me with love and patience.