DEUTSCHES ELEKTRONEN-SYNCHROTRON DESY
DESY 88-109
August 1988

THE STRONG CP PROBLEM

by

R.D. Peccei

Deutsches Elektronen-Synchrotron DESY, Hamburg

ISSN 0418-9833

NOTKESTRASSE 85 · 2 HAMBURG 52

To be sure that your preprints are promptly included in the
HIGH ENERGY PHYSICS INDEX ,
send them to the following address (if possible by air mail) :

DESY
Bibliothek
Notkestrasse 85
2 Hamburg 52
Germany

DESY 88-109
August 1988

ISSN 0418-9833

The Strong CP Problem *

R.D. Peccei
Deutsches Elektronen Synchrotron DESY
Hamburg, Fed. Rep. Germany

Abstract

I discuss the origins of the strong CP problem and outline some of the suggestions put forward to solve this conundrum, paying specific attention to their physical consequences. In particular, the issues connected with solving the strong CP problem by imposing an additional chiral symmetry are addressed, and some of the phenomenology of axions is touched upon.

1 Introduction

It is often the case in physics that the solution of some problem engenders some other, totally unrelated and unexpected, problem. The $U(1)_A$ problem of QCD is a good case in point. In the limit in which the u and d quarks masses vanish, the QCD Lagrangian has a global $U(2)_L \times U(2)_R$ symmetry. The vectorial part of this symmetry $U(2)_{L+R}$, corresponding to isospin plus baryon number, is manifestly realized in nature. The axial part of this symmetry $U(2)_{L-R}$, is spontaneously broken and one naively would expect four Goldstone bosons, three pions plus the η. The $U(1)_A$ problem is that the η meson really does not fullfil this role: turning on the quark masses gives the Goldstone bosons associated with the $U(2)_{L-R}$ breaking a mass, but one can show [1] that $m_\eta \leq \sqrt{3}m_\pi$, in gross contradiction with reality!

The solution to the $U(1)_A$ problem was found by t'Hooft [2], who showed that as a result of the axial anomaly [3] of the $U(1)_A$ current and of the non trivial properties of the QCD vacuum. there is really no Goldstone boson coupled to the physical $U(1)_A$ current. Thus, even in the massless quark limit, the η has a mass and one expects only three light pseudoscalar states, the pions. Although the axial anomaly of the $U(1)_A$ current is important to the resolution of this problem, the crucial ingredient for the solution is the complex nature of the QCD vacuum [2]. One can show that transition matrix elements in QCD split up into a family of disconnected sectors, labeled by a vacuum angle Θ [4][5]. Furthermore, the computational rules in each sector are the usual ones, except that an additional term proportional to $\Theta F\tilde{F}$ is added to the Lagrangian density. Since no transitions are possible between different Θ vacua, effectively the nature of the strong

*To appear in the review volume CP **Violation**, ed. by C. Jarlskog (World Scientific Publishing Co., Singapore)

interaction vacuum state is reflected in this augmented Lagrangian:

$$\mathcal{L}^{eff}_{QCD} = \mathcal{L}_{QCD} + \Theta \frac{g^2}{32\pi^2} F^{\mu\nu}_a \tilde{F}_{a\mu\nu} \qquad (1)$$

The decoupling of $q^2 = 0$ singularities from the $U(1)_A$ current, necessary to solve the $U(1)_A$ problem, requires that non zero vacuum expectation values of chiral breaking operators have a precisely matched Θ-phase. Thus the presence of the additional Θ-term in Eq. (1) is crucial to the solution of the $U(1)_A$ problem. However, unless $\Theta = 0, \pi$, it is apparent that P, T and CP are violated in the strong interactions, contrary to common knowledge! The physics which solves the $U(1)_A$ problem has generated a more intriguing conundrum: why is CP only very weakly violated?

Actually, the situation is even more mysterious, when one considers also the effect of the electroweak interactions. In this more general case, since in the process of diagonalizing the quark mass matrix to render it γ_5-free one performs $U(1)_A$ transformations, there arises a further $F\tilde{F}$ contribution, because of the chiral anomaly [3]. In effect, in the full strong plus electroweak theory. the parameter Θ in Eq. (1) is replaced by

$$\bar{\Theta} = \Theta + ArgdetM \qquad (2)$$

with M being the quark mass matrix. One can show that $\bar{\Theta}$ is severely restricted experimentally, $\bar{\Theta} \leq 10^{-9}$. Such a near cancellation is difficult to countenance, given the distinct origin of the Θ and the $ArgdetM$ terms. This is the strong CP problem.

The plan of this article is as follows. In Sec 2 a brief discussion is given of Θ-vacua, the effect of chiral transformations on these states and how the existence of this more complicated vacuum solves the $U(1)_A$ problem. Sec 3 discusses critically various suggested solutions for the strong CP problem, focusing particularly on their physical consequences. In Sec 4, I examine in more detail some of the issues connected with imposing an additional global chiral symmetry to solve the strong CP problem. Here some of the phenomenological aspects of axions, including prospects for their detection, are addressed. Sec 5. finally, contains some concluding observations.

2 The Θ-Vacuum

To understand why a correct description of QCD requires the introduction of a further dimensionless parameter Θ - the vacuum angle - it is particularly convenient to study the theory in the temporal gauge $A^0_a = 0$ [4][5]. In this gauge, one has still at one's disposal a time independent gauge transformation $\Omega(\vec{x})$, so that the spatial pure gauge fields are time independent.

$$A^i(\vec{x}) \equiv \frac{1}{2}\lambda_a A^i_a(\vec{x}) = \frac{i}{g}\Omega(\vec{x})\nabla^i\Omega^{-1}(\vec{x}) \qquad (3)$$

A topological distinction among various possible vacuum states - which classically are associated with pure gauge configurations - is introduced by demanding that at spatial infinity the gauge fields should vanish. This restricts the gauge transformations $\Omega(\vec{x})$ in Eq(3) to those which obey [5]

$$\Omega(\vec{x}) \lim_{|\vec{x}|\to\infty} 1 \qquad (4)$$

These transformations define a mapping of the three dimensional space, with points at infinity identified, into the group space.

Let us restrict ourselves, for simplicity, to an $SU(2)$ subgroup of QCD. Then the mapping in Eq(4) is a mapping of S^3 into S^3, since the manifold of $SU(2)$ is also homomorphic to the three dimensional sphere S^3. It is known that these mappings fall into distinct classes (homotopy classes) that cannot be continuously distorted into each other and are classified by an integer n, the winding number. For the $SU(2)$ example we are considering, the matrices $\Omega(\vec{x})$ then can be written in general as $\Omega(\vec{x}) = \{\Omega_n(\vec{x})\}$, where the index n indicates to which homotopy class they belong. Hence also the gauge fields A^i in Eq(3), and thus the associated vacua, can be classified by the winding number n:

$$A^i_n(\vec{x}) = \frac{i}{g}\Omega_n(\vec{x})\nabla^i\Omega_n^{-1}(\vec{x}) \tag{5}$$

The winding number n can be expressed in terms of an integral over the gauge fields (5) and one finds that [6]

$$n = \frac{ig^3}{24\pi^2}\int d^3x Tr\epsilon_{ijk}A^i_n(\vec{x})A^j_n(\vec{x})A^k_n(\vec{x}) \tag{6}$$

This formula is easily understood since $\epsilon_{ijk}TrA^iA^jA^k$ is essentially the Jacobian of the transformation from S^3 to the hypersphere in group space [6], with n measuring the number of windings of S^3 on the group space. A representative of gauge transformations of homotopy one is, for instance [5]

$$\Omega_1(\vec{x}) = \frac{\vec{x}^2-\lambda^2}{\vec{x}^2+\lambda^2} + \frac{2i\lambda\vec{\sigma}\cdot\vec{x}}{\vec{x}^2+\lambda^2} \tag{7}$$

with $\lambda > 0$, but otherwise arbitrary [1]. Representatives for $\Omega_n(x)$ are easily constructed by compounding Ω_1:

$$\Omega_n(x) = (\Omega_1(x))^n \tag{8}$$

It is possible by means of a gauge transformation to change the gauge field $A^i_n(\vec{x})$ into $A^i_{n+1}(\vec{x})$. Indeed, in view of Eq(8), such a transformation is induced precisely by $\Omega_1(x)$

$$\begin{aligned} A^i_{n+1}(\vec{x}) &= \frac{i}{g}\Omega_{n+1}(\vec{x})\nabla^i\Omega^{-1}_{n+1}(\vec{x}) = \frac{i}{g}[\Omega_1(\vec{x})\Omega_n(\vec{x})]\nabla^i[\Omega_n^{-1}(\vec{x})\Omega_1^{-1}(\vec{x})] \\ &= \frac{i}{g}\Omega_1(\vec{x})\nabla^i\Omega_1^{-1}(\vec{x}) + \Omega_1(\vec{x})A^i_n(\vec{x})\Omega_1^{-1}(\vec{x}) \end{aligned} \tag{9}$$

Obviously, although one can contemplate a vacuum state associated with each of the gauge field classes A^i_n, the correct vacuum state of QCD cannot be any one of these $|n>$ vacua, since they are not invariant under gauge transformations:

$$\Omega_1|n> = |n+1> \tag{10}$$

It is easy, however, to construct a gauge invariant vacuum state by superposing the various $|n>$ vacua. This is the Θ-vacuum:

$$|\Theta> = \sum_n e^{-in\Theta}|n> \tag{11}$$

[1]One can readily check that the insertion of $A^i_1 = \frac{i}{g}\Omega_1\nabla^i\Omega_1^{-1}$ into Eq(6), with Ω_1 given by Eq(7), gives unity

which is clearly gauge invariant

$$\Omega_1|\Theta> = \sum_n e^{-in\Theta}|n+1> = e^{i\Theta}|\Theta> \tag{12}$$

Each value of Θ, in effect, labels a separate theory. To see this, consider an "off-diagonal" Green's function. That is , the time ordered product of some gauge invariant operators 0_α taken between Θ and Θ'.

$$\begin{aligned} G_{off} &= <\Theta|T(0_1...0_p)|\Theta'> \\ &= \sum_{mn} e^{i(m\Theta-n\Theta')} <m|T(0_1...0_p)|n> \end{aligned} \tag{13}$$

Since the 0_α are gauge invariant operators, one has that

$$\Omega_1^{-1}T(0_1...0_p)\Omega_1 = T(0_1...0_p) \tag{14}$$

which implies that the matrix element in (13) only depends on the difference $\nu = m - n$

$$<m|T(0_1...0_p)|n> = F(\nu) \tag{15}$$

Thus one can rewrite G_{off} as

$$\begin{aligned} G_{off} &= \sum_m e^{im(\Theta-\Theta')}\sum_\nu e^{i\frac{\nu}{2}(\Theta+\Theta')}F(\nu) \\ &= 2\pi\delta(\Theta-\Theta')\sum_\nu e^{i\nu\Theta}F(\nu) \end{aligned} \tag{16}$$

Eq(16) implies that Θ cannot be changed by any gauge invariant perturbations.

To generate the Green's functions of QCD one usually considers the vacuum to vacuum amplitude in the presence of external sources: $<\Theta_+|\Theta_->^J$ [2]. This vacuum functional has a path integral representation, in which one integrates over all field configurations weighted by the action:

$$<\Theta_+|\Theta_->^J = \int d\mu_{fields}\, e^{iS_J} \tag{17}$$

In the $A^0_a = 0$ gauge we are considering, the functional paths in (17) connect pure gauge configurations of the type $A^i_n(\vec{x})$, at $t = -\infty$, to analogous pure gauge configurations, at $t = +\infty$. However, in general, these paths need not start and end with configurations of the same winding number n, as is shown schematically in Fig 1. Indeed, using Eq(11), one has

$$\begin{aligned} <\Theta_+|\Theta_->^J &= \sum_{mn} e^{im\Theta}e^{-in\Theta} <m_+|n_->^J \\ &= \sum_\nu e^{i\nu\Theta}\{\sum_n <(n+\nu)_+|n_->^J\} \end{aligned} \tag{18}$$

So the vacuum amplitude is really a sum over different vacuum transition amplitudes, in which the net change in the winding number between $t = +\infty$ and $t = -\infty$ is ν. Furthermore, in the sum, each of the amplitudes is weighted by phase factor $e^{i\nu\Theta}$.

[2]Here the $\pm$ indicate the vacuum states at $t = \pm\infty$, which need not be the same due to the external sources.

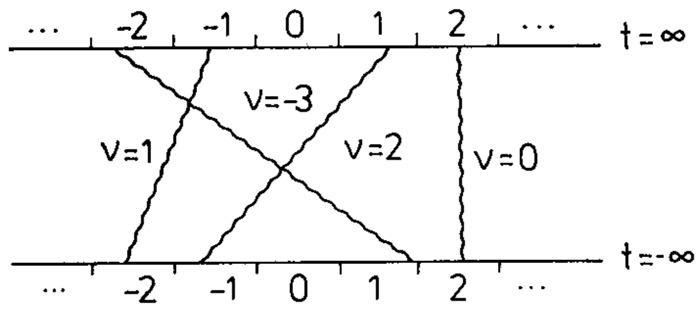


Figure 1: Some paths contributing to the vacuum to vacuum amplitude

The change $\nu = n_+ - n_-$ in the winding number between $t = +\infty$ and $t = -\infty$, characterizing the paths of Fig 1, has a simple interpretation in terms of the gauge field strengths. In fact, ν is related to an integral over all space time of $F\tilde{F}$ for these paths. To show this, one needs Bardeen's identity [7] expressing $F\tilde{F}$ as a total divergence:

$$F_a^{\mu\nu}\tilde{F}_{a\mu\nu} = \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}F_a^{\mu\nu}F_a^{\alpha\beta} = \partial^\mu K_\mu \tag{19}$$

where, for $SU(2)$,

$$K_\mu = \epsilon_{\mu\alpha\beta\gamma}A_a^\alpha[F_a^{\beta\gamma} - \frac{g}{3}\epsilon_{abc}A_b^\beta A_c^\gamma] \tag{20}$$

For pure gauge fields, and in the $A_a^0 = 0$ gauge, only K^0 is non vanishing, and one has

$$K^0 = -\frac{g}{3}\epsilon_{ijk}\epsilon_{abc}A_a^i A_b^j A_c^k = \frac{4}{3}ig\epsilon_{ijk}TrA^iA^jA^k \tag{21}$$

Thus the integral over all space time of $F\tilde{F}$ is related to the difference between the integrals at $t = +\infty$ and $t = -\infty$ of K^0 over all space. Specifically

$$\frac{g^2}{32\pi^2}\int d^4x F_a^{\mu\nu}\tilde{F}_{a\mu\nu} = \frac{g^2}{32\pi^2}\int d\sigma_\mu K^\mu = \frac{g^2}{32\pi^2}\int d^3x K^0|_{t=-\infty}^{t=\infty} \tag{22}$$

since on the surfaces only pure gauge fields contribute, $F^{\mu\nu} = 0$. Using Eq(21), and recalling the definition of the winding number n from Eq(6), one sees that

$$\nu = n_+ - n_- = \frac{g^2}{32\pi^2}\int d^3x K^0|_{t=-\infty}^{t=\infty} = \frac{g^2}{32\pi^2}\int d^4x F_a^{\mu\nu}\tilde{F}_{a\mu\nu} \tag{23}$$

Eq.(23) allows one to recast the phase factor $e^{i\nu\Theta}$ in Eq(18) as an effective <u>additional</u> contribution to the QCD Lagrangian. The transition amplitudes $<(n+\nu)_+|n_->$ in

Eq(18) are given by a path integral in which the space-time integral of $F\tilde{F}$ is fixed by (23). The extra factor of $e^{i\nu\Theta}$ corresponds simply to an additional piece to the QCD action

$$S_\Theta = \Theta\frac{g^2}{32\pi^2}\int d^4x F_a^{\mu\nu}\tilde{F}_{a\mu\nu} \tag{24}$$

Hence

$$< \Theta_+|\Theta_- >^J = \sum_\nu \int d\mu_{fields}\, e^{iS_J^{eff}}\delta(\nu - \frac{g^2}{32\pi^2}\int F\tilde{F}) \tag{25}$$

where the effective action S_J^{eff} is the usual QCD action with external sources, but now augmented by the contribution of Eq(24). We see, therefore, that the nontrivial nature of the QCD vacuum leads directly to the potentially P,T and CP violating effective Lagrangian for QCD, given in Eq(1).

The quantity $F\tilde{F}$ plays another role in QCD, which is connected with the anomalous divergence of the $U(1)_A$ current [3]. The current

$$J_5^\mu = \sum_{i=1}^{N_f}\bar{q}_i\gamma^\mu\gamma_5 q_i \tag{26}$$

is classically conserved in QCD, in the limit of vanishing quark mass terms. However, at the loop level, there is an anomaly [3]

$$\partial_\mu J_5^\mu = 2N_f(\frac{g^2}{32\pi^2}F_a^{\mu\nu}\tilde{F}_{a\mu\nu}) \tag{27}$$

Because of this anomaly one cannot really use the charge

$$Q_5 = \int d^3x J_5^0 \tag{28}$$

as the generator of chiral transformations, since $\dot{Q}_5 \neq 0$. However, in the limit of zero quark masses, Bardeen's identity provides one readily with a conserved chiral current

$$\tilde{J}_5^\mu = J_5^\mu - 2N_f(\frac{g^2}{32\pi^2}K^\mu) \tag{29}$$

whose associated charge

$$\tilde{Q}_5 = \int d^3x \tilde{J}_5^0 \tag{30}$$

is time independent.

Although $\tilde{Q}_5$ is a perfectly good generator for the chiral $U(1)_A$ transformations, since it involves K^0, it is <u>not</u> invariant under gauge transformations. Indeed, since under the gauge transformations $\Omega_1(\vec{x})$

$$\Omega_1(\frac{g^2}{32\pi^2}\int d^3x K^0)\Omega_1^{-1} = \frac{g^2}{32\pi^2}\int d^3x K^0 - 1 \tag{31}$$

it follows that [5]

$$\Omega_1\tilde{Q}_5\Omega_1^{-1} = \tilde{Q}_5 + 2N_f \tag{32}$$

This equation is very important, since it tells us that the Θ-vacuum is <u>not</u> invariant under chiral transformations. To see this, consider the action of Ω_1 on the state $e^{i\alpha\hat{Q}_5}|\Theta>$:

$$\begin{aligned}\Omega_1[e^{i\alpha\hat{Q}_5}|\Theta>] &= \Omega_1 e^{i\alpha\hat{Q}_5}\Omega_1^{-1}\Omega_1|\Theta> \\ &= e^{i(\Theta+2N_f\alpha)}[e^{i\alpha\hat{Q}_5}|\Theta>]\end{aligned} \tag{33}$$

Comparing the above with Eq(12), one sees that a chiral rotation by α changes the $|\Theta>$ vacuum state into the $|\Theta + 2N_f\alpha>$ state:

$$e^{i\alpha\hat{Q}_5}|\Theta> = |\Theta + 2N_f\alpha> \tag{34}$$

Eq(34) has two interesting corollaries:

1. In a world with massless quarks, the results of the theory are actually independent of the vacuum angle Θ. In this limit, the action is chiral invariant and one can pass from a given Θ-vacuum to another by a chiral transformation. Thus the additional CP violating term in the QCD Lagrangian of Eq(1) has no physical significance. No physical parameters can depend on Θ, in the limit that the quark masses vanish [3].

2. In the real world, where quarks acquire their mass via the breakdown of the electroweak $SU(2)\times U(1)$ group, the parameter Θ, associated with the QCD vacuum angle in Eq(1), gets an additional contribution. The quark mass matrix emerging from the electroweak breakdown is, in general, both off-diagonal and non Hermitian. With q_L, q_R corresponding to the chiral projections $\frac{1}{2}(1-\gamma_5)q$ and $\frac{1}{2}(1+\gamma_5)q$, one has

$$\mathcal{L}_{Mass} = -\bar{q}_{Ri}M_{ij}q_{Lj} - \bar{q}_{Li}(M^\dagger)_{ij}q_{Rj} \tag{35}$$

The mass matrix in (35) can be diagonalized by separate unitary transformations on the left and right quark fields. These transformations will also involve a chiral $U(1)_A$ transformation if M is non Hermitian, so that $Arg\det M \neq 0$. In fact, it is easy to check that the amount of $U(1)_A$ transformation needed to obtain a purely γ_5-free quark mass term

$$\mathcal{L}_{Mass} = -\sum_{i=1}^{N_f} m_i\bar{q}_i q_i \tag{36}$$

is just $Arg\det M$. Thus, for quarks fields defined to have the conventional mass term (36), the effective Θ-parameter in the theory is that of Eq(2): $\bar{\Theta} = \Theta + Arg\det M$, and the P, T and CP violating addition to the full standard model Lagrangian is

$$\mathcal{L}_{SM}^{eff} = \mathcal{L}_{SM} + \bar{\Theta}\frac{g^2}{32\pi^2}F_a^{\mu\nu}\tilde{F}_{a\mu\nu} \tag{37}$$

We shall see in the next Sections how difficult it is to find cogent and experimentally viable ways to make $\bar{\Theta}$ either vanish or be extremely small. Before embarking on this arduous task, it seems worthwhile to question more closely the theoretical basis for the appearance of the additional term in (37). Its raison d'être resides essentially on taking

[3] To get rid of any Θ-dependence, it suffices that one quark mass vanishes, since all that is needed is the freedom to perform <u>some</u> chiral transformation

seriously the topological structure of the QCD vacuum, typified in our discussion by the requirement of Eq(4) that the gauge transformations $\Omega(\vec{r})$ vanish at spatial infinity. Is this topological information really germane? Our arguments made use of semiclassical notions in which vacuum states are associated with classes of pure gauge field configurations and, at best, they have only some heuristic value. One can, of course, do better. For instance, by quantizing QCD in a finite volume, the topological arguments can be sharpened [8]. However, one still has to worry that, perhaps, on taking the infinite volume limit, the topological information washes out. Fortunately, the $U(1)_A$ problem comes to the rescue! We know that the η (or more precisely the η', if we consider also the s-quark as being light), does not behave as a Goldstone boson. This is understood if the QCD vacuum is the Θ-vacuum we discussed, but otherwise no other credible explanation exists. So, in some sense, experiment tells us to take the topological structure of QCD seriously. Without the Θ-vacua one cannot really understand why the η is not light!

2.1 Solving the $U(1)_A$ Problem

Before closing this Section, it behooves us to indicate in more detail why the existence of the Θ-vacua solves the $U(1)_A$ problem. [9] Consider an operator X which has non vanishing chirality χ_X, so that

$$[\hat{Q}_5, X] = -\chi_X X \tag{38}$$

and which has non vanishing vacuum expectation value

$$<\bar{\Theta}|X|\bar{\Theta}> \neq 0 \tag{39}$$

Then it is easy to show that, in the massless quark limit, the Green's function

$$\tilde{G}^\mu(q) = \int d^4x e^{-iqx} <\bar{\Theta}|T(\tilde{J}_5^\mu(x)X)|\bar{\Theta}> \tag{40}$$

has a q^{-2} singularity. Because $\tilde{J}_5^\mu$ is not gauge invariant, however, it is not immediately clear if this singularity is to be associated with the presence of a <u>physical</u> Goldstone boson. What one needs to know is if this q^{-2} term also appears in the Green's function $G^\mu(q)$, involving the gauge invariant current J_5^μ.

The residue of the q^2 pole in $\tilde{G}^\mu(q)$ is given by

$$\int d^4x\, \partial_\mu <\bar{\Theta}|T(\tilde{J}_5^\mu(x), X)|\bar{\Theta}> = -\chi_X <\bar{\Theta}|X|\bar{\Theta}> \tag{41}$$

Using the anomaly equation (27) and the definition (29), it follows that the residue of the q^2 pole in $G^\mu(q)$ is

$$\begin{aligned}\int d^4x\, \partial_\mu <\bar{\Theta}|T(J_5^\mu(x), X)|\bar{\Theta}> &= -\chi_X <\bar{\Theta}|X|\bar{\Theta}> \\ &+2N_f\frac{g^2}{32\pi^2}\int d^4x <\bar{\Theta}|T(F_a^{\mu\nu}\tilde{F}_{a\mu\nu}(x), X)|\bar{\Theta}>\end{aligned} \tag{42}$$

So, even though the residue (41) is non vanishing, it is possible that the residue (42) vanishes, if there is a cancellation among the two terms on the RHS. Since the space time integral of $F\tilde{F}$ explicitly enters in the QCD vacuum functional (cf Eq(24)), the second term in (42) is simply

$$2N_f\frac{g^2}{32\pi^2}\int d^4x < \bar{\Theta}|T(F_a^{\mu\nu}\tilde{F}_{a\mu\nu}(x),X)|\bar{\Theta} >= 2N_f\frac{1}{i}\frac{\partial}{\partial\bar{\Theta}} < \bar{\Theta}|X|\bar{\Theta} > \quad (43)$$

Thus the condition of having no physical Goldstone excitations associated with the $U(1)_A$ current requires that the non zero vacuum expectation value of X have a specific $\bar{\Theta}$-dependence:

$$[2N_f\frac{1}{i}\frac{\partial}{\partial\bar{\Theta}} - \chi_X] < \bar{\Theta}|X|\bar{\Theta} >= 0 \quad (44)$$

What 't Hooft [2] showed was that Eq(44) obtains identically for operators X which break $U(1)_A$, due to the presence of instanton fields [10]. In the presence of an instanton of index ν, there are $N_f\nu$ chiral eigenmodes in the theory [2], as a result of the Atiyah-Singer theorem [11]. Thus the operators which can have non vanishing expectation value have chirality $\chi_X = 2N_f\nu$. Furthermore, since their non vanishing vacuum expectation arises in a ν-sector, one has $< \bar{\Theta}|X|\bar{\Theta} >= Ce^{i\nu\bar{\Theta}}$. Clearly, in this case, Eq(44) is trivially satisfied.

Having found an instance where one can show that there are no $q^2 = 0$ singularities attached to the gauge invariant $U(1)_A$ current, it is not difficult to imagine that the way the $U(1)_A$ problem is solved is that in all instances this matching of $\bar{\Theta}$-phase with chirality obtains. Actually, one has to be a little careful when one deals with operators X which are not purely $SU(N_f)$ singlets [12]. In this case, since X can couple to some of the $SU(N_f)$ Goldstone bosons, in the limit in which the quark masses are restored, one has an additional contribution to Eq(44):

$$[2N_f\frac{1}{i}\frac{\partial}{\partial\bar{\Theta}} - \chi_X] < \bar{\Theta}|X|\bar{\Theta} >= -2i\sum_{i=1}^{N_f}\int d^4x < \bar{\Theta}|T(m_i\bar{q}_i\gamma_5 q_i(x),X)|\bar{\Theta} > \quad (45)$$

Although the RHS of Eq(45) appears to vanish as the quark masses vanish, this will not occur if X can couple to an $SU(N_f)$ Goldstone boson. In this case, the amplitude on the RHS actually is finite in the $m_i \to 0$ limit, since the Goldstone boson contributions are of $0(\frac{1}{m_i})$. Thus, for a general operator X it is Eq(45), not Eq(44), that determines the $\bar{\Theta}$-dependence of $< \bar{\Theta}|X|\bar{\Theta} >$ necessary to solve the $U(1)_A$ problem.

Let me consider briefly the case of two light flavors, keeping, however, $m_d \neq m_u$. In the $\bar{\Theta}$-vacuum, the vacuum expectation values of $\bar{u}_L u_R$ and $\bar{d}_L d_R$ will have the same magnitude, but different phases

$$\begin{aligned} < \bar{\Theta}|\bar{u}_L u_R|\bar{\Theta} > &= Ve^{i\phi_u(\bar{\Theta})} \\ < \bar{\Theta}|\bar{d}_L d_R|\bar{\Theta} > &= Ve^{i\phi_d(\bar{\Theta})} \end{aligned} \quad (46)$$

The phases in (46) show that the vacuum state is not CP invariant, since $\bar{\Theta}$ is non vanishing. We can make these vacuum expectation values real by rotating the quark fields appropriately. These rotations eliminate the $\bar{\Theta}$-term in the Lagrangian but transfer the phase information to the quark mass terms. Obviously, one has that

$$\phi_u(\bar{\Theta}) + \phi_d(\bar{\Theta}) = \bar{\Theta} \quad (47)$$

and the CP violating Lagrangian is now

$$\mathcal{L}_{CP\ viol} = -im_u sin\phi_u(\bar{\Theta})\bar{u}\gamma_5 u - im_d sin\phi_d(\bar{\Theta})\bar{d}\gamma_5 d \quad (48)$$

In principle one can determine ϕ_u and ϕ_d by using Eq(45). However, one can bypass this computation by using Dashen's theorem [13] on vacuum state alignment, in the presence of a perturbation. This theorem states that, in the correct vacuum, the perturbation has a local minimum under rotations of the global symmetry group. The perturbation in our case is just the mass Lagrangian (before the $U(1)_A$ rotations that led to Eq(48))

$$\mathcal{L}_{Mass} = -m_d\bar{d}d - m_u\bar{u}u \quad (49)$$

and Dashen's theorem guarantees that its first variations vanishes:

$$< \bar{\Theta}|[G_i,\mathcal{L}_{Mass}]|\bar{\Theta} >= 0 \quad (50)$$

for all $SU(2)\times SU(2)$ generators G_i. This result implies that the matrix element of $\mathcal{L}_{Mass}$ between the correct vacuum and an $SU(N_f)$ Goldstone boson is of $0(m_i^2)$. Indeed, to lowest order in m_i, this matrix element can be transformed into the commutator of the chiral generators with $\mathcal{L}_{Mass}$ and this vanishes by Eq(50). We see, therefore, that if we use for the operator X in Eq(45) the mass perturbation (49), the RHS of this equation is of $0(m_i^2)$ and can be dropped. Hence one has, for the case of two flavors,

$$[4\frac{1}{i}\frac{\partial}{\partial\bar{\Theta}} - 2] < \bar{\Theta}|m_u\bar{u}_L u_R + m_d\bar{d}_L d_R|\bar{\Theta} > +[4\frac{1}{i}\frac{\partial}{\partial\bar{\Theta}} + 2] < \bar{\Theta}|m_u\bar{u}_R u_L + m_d\bar{d}_R d_L|\bar{\Theta} >= 0 \quad (51)$$

Using Eq(46) this implies

$$m_u[2\frac{\partial\phi_u(\bar{\Theta})}{\partial\bar{\Theta}} - 1]sin\phi_u(\bar{\Theta}) + m_d[2\frac{\partial\phi_d(\bar{\Theta})}{\partial\bar{\Theta}} - 1]sin\phi_d(\bar{\Theta}) = 0 \quad (52)$$

which, in view of Eq(47), reduces to the constraint [4]

$$m_u sin\phi_u(\bar{\Theta}) = m_d sin\phi_d(\bar{\Theta}) \quad (53)$$

One can check that the solution of (47) and (53) is given by [14][15]

$$\begin{aligned} sin\phi_u(\bar{\Theta}) &= \frac{m_d sin\bar{\Theta}}{[m_u^2 + m_d^2 + 2m_u m_d cos\bar{\Theta}]^{\frac{1}{2}}} \\ sin\phi_d(\bar{\Theta}) &= \frac{m_u sin\bar{\Theta}}{[m_u^2 + m_d^2 + 2m_u m_d cos\bar{\Theta}]^{\frac{1}{2}}} \end{aligned} \quad (54)$$

The solution of the $U(1)_A$ problem requires that the $\bar{\Theta}$-dependence of the vacuum expectation values of $\bar{u}u$ and $\bar{d}d$ be as detailed in (54). However, contrary to the case of the operators X which broke $U(1)_A$ due to instanton fields, here there is no direct proof that Eq(54) obtains [16]. Nevertheless, as we shall explicitly show in Section 4, one can construct a low energy effective Lagrangian for the meson sector of QCD which precisely

[4] The solution $\phi_u = \phi_d = \frac{1}{2}\bar{\Theta}$ of Eq(52) is rejected by the requirement that the $\bar{\Theta}$-periodicity of the vacuum expectation values be 2π

reproduces this behaviour [14]. Thus, it is very plausible that the $\bar{\Theta}$ dependence of the quark vacuum expectation values is that given by Eq(54) and the $U(1)_A$ problem is solved [9]. In addition, one can argue that if Eq(54) holds, then the CP violating Lagrangian which arises from a non zero $\bar{\Theta}$ parameter [Eq(48)] has all the expected physical properties. To see this, consider for simplicity $\bar{\Theta}$ to be small - as it needs to be experimentally. Then

$$\phi_u(\bar{\Theta}) \simeq \frac{m_d\bar{\Theta}}{m_u + m_d} \qquad \phi_d(\bar{\Theta}) \simeq \frac{m_u\bar{\Theta}}{m_u + m_d} \tag{55}$$

and the CP violating Lagrangian becomes

$$\mathcal{L}_{CP\ viol} = -i\frac{m_u m_d}{m_u + m_d}\bar{\Theta}[\bar{u}\gamma_5 u + \bar{d}\gamma_5 d] \tag{56}$$

We see that there is no CP violation precisely in all the relevant limits we discussed previously: $\bar{\Theta} = 0$, or $m_d = 0$, or $m_u = 0$. Thus, although one cannot demonstrate with absolute rigor that the presence of the Θ-vacua and the chiral anomaly combine to solve the $U(1)_A$ problem, this result also makes this assertion more convincing.

As a last point, I should mention that there is a second constraint, besides Eq(45), that needs to be checked before one can really claim a solution to the the $U(1)_A$ problem. This constraint arises by considering a double insertion of the $F\tilde{F}$ operator [17]. One can show that the absence of a $U(1)_A$ Goldstone boson necessitates that the Green's function containing two $F\tilde{F}$ operators obey the constraint:

$$i(\frac{N_f g^2}{32\pi^2})^2 \int d^4x < \bar{\Theta}|T[F_a^{\mu\nu}\tilde{F}_{a\mu\nu}(x), F_b^{\lambda\kappa}\tilde{F}_{b\lambda\kappa}(0)]|\bar{\Theta} >= m_\pi^2 F_\pi^2 \tag{57}$$

This equation also has not really been proved to hold rigorously. However, in the large N limit of QCD, Witten [18] has shown that this equation is consistent with having a meson in the $U(1)_A$ channel whose mass does not vanish, as the quark masses vanish.

3 Theoretical Attempts to Solve the Strong CP Problem

The presence of the CP violating Lagrangian of Eq(56) will induce a non vanishing electric dipole moment for the neutron. Since this quantity is very well bounded experimentally [19]: $d_n \leq 6 \times 10^{-25}$ ecm, this provides a strong numerical constraint on the value of $\bar{\Theta}$. That this constraint is severe can be appreciated from a naive, dimensional analysis, estimate of the neutron dipole induced by Eq(56)

$$d_n \sim \frac{e}{M}(\frac{m}{M})\bar{\Theta} \simeq 10^{-16}\bar{\Theta}\ ecm \tag{58}$$

A more accurate estimate was obtained early on by Baluni [20] and by Crewther, Di Vecchia, Veneziano and Witten [21]. Baluni used a bag model to compute the contributions of excited resonances to the expectation of $T(J^\mu_{em}(x)\mathcal{L}_{CP\ viol})$ in a neutron, while Crewther et al computed d_n in chiral perturbation theory, where the dominant contribution comes from the $P\pi$ intermediate state. Their results

$$d_n = \begin{cases} 2.7 \times 10^{-16}\ ecm & Baluni \\ 5.2 \times 10^{-16}\ ecm & Crewther\ et\ al \end{cases} \tag{59}$$

are still quite representative of the present theoretical expectations for d_n [22], which range from $d_n = 4 \times 10^{-17}\bar{\Theta}$ ecm [23] to $d_n = 2 \times 10^{-15}\bar{\Theta}$ ecm [24]. Clearly one needs a value of $\bar{\Theta}$ below 10^{-9} to be on the safe side with experiment!

The near cancellation of the QCD vacuum angle contribution, Θ, with the phase coming from the quark matrix, $Arg\det M$, to yield $\bar{\Theta} \leq 10^{-9}$ is the crux of the strong CP problem. What physics fixes this physical parameter to be so small? In the standard model, where CP violation is due to having complex Yukawa couplings, the parameter $\bar{\Theta}$ receives infinite contributions in perturbation theory (albeit at very high order [25]) and must be renormalized. Thus its physical value is undeterminable theoretically. To explain its value - like that of any other small parameter in the standard model, like $\frac{m_u}{m_t}$ - requires imagining that there is some additional physics beyond the standard model. Granting this, then there appears to be three possible answers to why $\bar{\Theta} \leq 10^{-9}$:

1. There is really no strong CP problem

2. There is a dynamical reason which forces $\bar{\Theta} = 0$

3. There is a symmetry reason which allows $\bar{\Theta}$ to be calculable and (very) small.

I want to examine each of these suggestions in turn.

3.1 Is There a Strong CP Problem?

In the last Section I already broached the possibility that the strong CP problem could be solved simply by not taking the topological structure of the QCD vacuum seriously. However, I rejected this possibility because of the $U(1)_A$ problem. This point of view resurfaces periodically in the literature, with clever suggestions of how to avoid the topological structure of QCD. I believe, however, that these attempts are misguided. For instance, recently Khlebnikov and Shaposhnikov [26] have suggested that one could get rid of $\bar{\Theta}$ altogether by considering higher dimensional theories which do not compactify to manifolds of the type of $M^4\times$ compact space (with M^4 being the 4-dimensional Minkowski space). In such theories one never gets the $|n>$ vacua we discussed in Sec 2 and so the strong CP problem is absent. However, and this is the crucial point, these theories cannot explain away the $U(1)_A$ problem. One cannot simply get rid of the QCD topology, without providing an alternative explanation of why the η-meson is not light!

There is, of course, a way to preserve the nice features of the Θ-vacua and yet have no strong CP problem. If some quark really had zero mass, then, although the $U(1)_A$ problem is solved by the existence of Θ-vacua, all these vacua are equivalent and there is no CP violation at all [c.f. Eq(56), which shows that the CP-violating Lagrangian vanishes as m_u or m_d vanish]. The most natural quark to have a vanishing mass is the u-quark. However, although the mass of the u-quark is small, it does not appear to be zero. A careful study by Gasser and Leutwyler [27], of the influence of quark masses on the masses of baryons and mesons, gives a non vanishing value for m_u, with the running mass [28] at 1 GeV being

$$m_u(1GeV) = (5.1 \pm 1.5)\ MeV \tag{60}$$

It is difficult to vitiate this result. Nevertheless, recently Kaplan and Manohar [29] have argued, by using a general $SU(3)\times SU(3)$ chiral Lagrangian, that one can have sufficiently large 2^{nd} order corrections to the meson masses so that the solution $m_u = 0$ is allowed. Although Kaplan and Manohar are technically correct, my own impression, however, is that they are stretching a point.

A rather more interesting idea along these lines, has been suggested recently by Choi, Kim and Sze [30]. These authors speculate that the u quark mass in the QCD Lagrangian, m_u, actually vanishes, thereby solving the strong CP problem. However, they show that instanton induced terms, in the presence of nonvanishing d and s mass terms, yield an effective u quark mass which contributes to the meson and baryon masses. It is this effective mass which is fixed by the analysis of Gasser and Leutwyler [27]. Thus the suggestion of Choi et al does not question the validity of the current algebra and QCD sum rule calculations which yielded Eq(60). However, their argument is that the mass m_u measured by these calculations is not the intrinsic quark mass, but a parameter which arises through the instanton induced chiral symmetry breaking, in conjunction with the explicit breaking provided by m_d and m_s. This is quite a nice way to solve the strong CP problem, except for an unexplained fact: what physics forces one of the eigenvalues of the quark mass matrix M to vanish? To my mind, unless one finds an explicit reason why this is so, this solution of the strong CP problem remains as unsatisfactory as postulating that $\bar{\Theta}$ just happens to be small! At any rate, this is my prejudice.

3.2 Adjusting $\bar{\Theta}$ to Zero Dynamically

Having disposed of the first alternative, that the solution of the strong CP problem is due to the fact that there is no strong CP problem, let me now turn to the second possibility: namely, that there is a dynamical reason which forces $\bar{\Theta}$ to vanish. This suggestion for solving the strong CP problem was put forward, about 10 years ago, by Helen Quinn and me. We showed [31] that, if the standard model was augmented by a global chiral $U(1)$ symmetry (known in the literature now as a $U(1)_{PQ}$ symmetry), then the dynamics of the theory is such that $\bar{\Theta}$ is dynamically set to zero. I do not want to repeat the arguments of [31] here, which made considerable use of properties of instantons. I prefer instead, and I think it is more useful, to expose the salient points of the idea in a different, more direct, way.

Obviously, if the standard model has an additional global chiral symmetry, and this symmetry is unbroken, there is never any strong CP problem. This is like having a massless quark, because by a chiral transformation one can rotate $\bar{\Theta}$ away (c.f. (Eq34)). The interesting physical case to consider, therefore, is the case when the standard model is augmented by a $U(1)_{PQ}$ symmetry and this symmetry is spontaneously broken. If the $U(1)_{PQ}$ symmetry is spontaneously broken, there must exist in the theory a spin zero excitation, with zero mass at the Lagrangian level. This excitation, the Goldstone boson of the broken $U(1)_{PQ}$ symmetry, is the famous (or infamous!) axion, first discussed in this context by Weinberg and Wilczek [32]. Under the $U(1)_{PQ}$ symmetry, the axion field, being the field of a Goldstone boson, translates:

$$a(x) \to a(x) + \delta\alpha v_{PQ} \tag{61}$$

where v_{PQ} is a scale parameter associated with the breaking, and $\delta\alpha$ is the infinitesimal parameter of the transformation.

For a normal spontaneously broken symmetry, the fact that the Goldstone bosons translate under the symmetry, implies that the effective Lagrangian containing the Goldstone bosons can only involve derivatives of these fields. The $U(1)_{PQ}$ symmetry, however, because it is chiral, suffers from an Adler Bell Jackiw anomaly [3]. This anomalous behaviour of the Lagrangian of the theory is reproduced, simply, by having a term linear in the axion field coupled to the anomaly. Hence, if the standard model is augmented by a spontaneously broken $U(1)_{PQ}$ symmetry, it will be described by an effective Lagrangian

$$\begin{aligned}\mathcal{L}_{eff} &= \mathcal{L}_{SM} + \bar{\Theta}\frac{g^2}{32\pi^2}F_a^{\mu\nu}\tilde{F}_{a\mu\nu} + \frac{a}{v_{PQ}}\xi\frac{g^2}{32\pi^2}F_a^{\mu\nu}\tilde{F}_{a\mu\nu} \\ &\quad -\frac{1}{2}\partial_\mu a\partial^\mu a + \mathcal{L}(\partial_\mu a, \psi)\end{aligned} \tag{62}$$

Here ξ and $\mathcal{L}(\partial_\mu a, \psi)$ are model dependent quantities related to how one assigns $U(1)_{PQ}$ transformations to the fermions in the theory.

The existence of the $U(1)_{PQ}$ breaking interaction $aF\tilde{F}$, due to the chiral anomaly, actually provides a potential for the axion field. It is no longer true that in the vacuum all values of $< a >$ are allowed, reflecting the naive $U(1)_{PQ}$ symmetry of the Lagrangian. Including the anomaly contribution, one finds that the vacuum expectation value of the axion field is fixed to be [31]

$$< \bar{\Theta}|a|\bar{\Theta} > = -\bar{\Theta}\frac{1}{\xi}v_{PQ} \tag{63}$$

The physical axion field, of course, is the excitation with this vacuum expectation removed

$$a_{phys} = a - < \bar{\Theta}|a|\bar{\Theta} > \tag{64}$$

Thus, in terms of this field one has

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \mathcal{L}(\partial_\mu a_{phys}, \psi) - \frac{1}{2}\partial^\mu a_{phys}\partial_\mu a_{phys} + \frac{a_{phys}}{v_{PQ}}\xi\frac{g^2}{32\pi^2}F_a^{\mu\nu}\tilde{F}_{a\mu\nu} \tag{65}$$

The presence of the extra $U(1)_{PQ}$ symmetry has eliminated the offending P, T and CP violating $\bar{\Theta}$ parameter, replacing it by a dynamical field: the axion!

How can one understand Eq(63)? This is easily seen by examining the equations of motion of the axion field. From (62) one has

$$-\partial^2 a + \partial^\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu a} = \frac{\xi}{v_{PQ}}\frac{g^2}{32\pi^2}F_a^{\mu\nu}\tilde{F}_{a\mu\nu} \tag{66}$$

If it were not for the anomaly term, it is clear that any constant value for the axion field in the vacuum would be allowed. Eq(66) informs us, however, that the axion field settles in the vacuum at the value where

$$< \frac{\partial V_{eff}}{\partial a} > \equiv -\frac{\xi}{v_{PQ}}\frac{g^2}{32\pi^2} < F_a^{\mu\nu}\tilde{F}_{a\mu\nu} > = 0 \tag{67}$$

The expectation value of the P, T and CP violating density $F\tilde{F}$ is periodic in the relevant Θ-parameter of the theory. From Eq(62) this is simply just $\bar{\Theta} + \frac{<a>}{v_{PQ}}\xi$ and the expectation of $F\tilde{F}$ in the vacuum vanishes precisely when Eq(63) is fulfilled [5].

[5] For instance, in the one instanton approximation, $< F_a^{\mu\nu}\tilde{F}_{a\mu\nu} >$ is proportional to $sin(\bar{\Theta} + \frac{<a>}{v_{PQ}}\xi)$ [31]

Although the introduction of the spontaneously broken $U(1)_{PQ}$ symmetry solves the strong CP problem, it necessitates the presence of a new dynamical field in the theory: the axion [32]. This excitation, which is nominally massless because it is a Goldstone boson, acquires a mass as a result of the chiral anomaly. Indeed,

$$m_a^2 = < \frac{\partial^2 V_{eff}}{\partial a^2} > = -\frac{\xi}{v_{PQ}} \frac{g^2}{32\pi^2} \frac{\partial}{\partial a} < F\tilde{F} > |_{<a>=-\frac{\Theta}{\xi} v_{PQ}} \tag{68}$$

The axion mass is proportional to the curvature of the effective potential induced by the anomaly. We shall have occasion later on to evaluate (68) more precisely. Here we remark only that, purely on dimensional grounds, the axion mass is of order

$$m_a^2 \sim \frac{\Lambda_{QCD}^4}{v_{PQ}^2} \tag{69}$$

where Λ_{QCD}^4 typifies the scale of the expectation of $< F\tilde{F} >$. If $v_{PQ} >> \Lambda_{QCD}$, one sees that one predicts a very light boson, as the price for resolving the strong CP problem dynamically.

3.3 The Soft CP Option for the Strong CP Problem

I shall examine in much more detail in the next Section the properties of axions and the unsuccessful quest, so far, of finding these excitations experimentally. Before doing this, however, I consider the last logical possibility for an understanding of the strong CP problem. Namely, that the full theory which contains the standard model, is such that $\bar{\Theta}$ is calculable and is very tiny. For $\bar{\Theta}$ to be calculable, it is necessary that the overlaying theory require no $F\tilde{F}$ counterterm. This can only be guaranteed if there is a symmetry reason to prevent such a counterterm from arising. Obviously, the right symmetry to invoke is CP invariance! However, since we have evidence for CP violation in nature, one must assume that CP is not quite a full symmetry of the theory.

More technically, we want a theory where $Arg\det M$ is calculable (i.e. finite in all orders of perturbation theory). Such a theory requires that CP be broken softly, by operators of dimension less than 4. This soft CP breaking can arise in two ways:

1. CP can be broken explicitly in the Lagrangian, but only by operators of low dimensionality, $d < 4$. For instance, it could be broken by having complex mass mixing in the Higgs sector, among different Higgs fields.

2. CP can be broken spontaneously by non trivial phases (i.e. phases which cannot be rotated away) appearing in the vacuum expectation values of the Higgs fields in the theory.

To my mind, only the second type of soft CP violation makes sense in connection to the strong CP problem. Recall that to solve this problem we need to make $\bar{\Theta} = \Theta + Arg\det M$ small, and not only $Arg\det M$ small. It suffices to make $Arg\det M$ small, if we can set the QCD angle Θ to zero. This condition can be imposed as a CP requirement on the QCD part of the total Lagrangian. However, this requirement makes only sense if CP is a symmetry of the full Lagrangian! So soft CP breaking, in what follows, is to be

understood only in the context of spontaneous CP breaking by the vacuum, a phenomenon first considered by T.D. Lee [33] and P. Sikivie [34].

There have been many attempts in the literature to solve the strong CP problem by appealing to soft CP breaking [35]. However, most of the models constructed are plagued by a variety of technical and physical difficulties. Broadly speaking one can classify these models into two classes:

1. Models where the spontaneous breaking of CP is done at or near the weak scale

2. Models where CP is spontaneously broken at a large scale - usually assumed to be a scale connected with some grand unified theory.

Models of the first kind, which quite naturally were among the first to be suggested, have fallen into disrepute, because they lead to problems with flavor changing neutral currents and are, furthermore, cosmologically not viable. The second type of models are in somewhat better shape. Nevertheless, also here the prospects do not seem to be bright - at least to me. In particular, the observation of a non zero value for the CP violating parameter ϵ' in the Kaon system [36] severely restricts the available options.

If CP is broken spontaneously one needs to have two or more Higgs fields in the theory [33]. If the additional Higgs field is also an $SU(2)$ doublet, then one immediately has difficulty preventing flavor changing neutral currents - induced by Higgs exchange - to arise. Even if one, somehow, manages to avoid these effects being too big, it is difficult to guarantee that the soft CP phase does not enter in $Arg\det M$. Indeed, in general, this will happen, unless there are some extra symmetries in the theory which force $Arg\det M$ at tree level to vanish. Even in this case, one must then check that higher order corrections do not induce an unmanagable strong CP phase. Typically, this requires further adjustments to get rid of all the one loop effects. Thus models where CP is broken softly at the weak scale [37] require a barrage of extra assumptions, which very few people find palatable.

These technical pecadillos might be forgivable, if it were not for the domain wall problem. CP is a discrete symmetry and, if it is spontaneously broken by the vacuum, as the Universe cools down to the temperature below where this symmetry is spontaneously broken, domains of different CP phases form. Zeldovich Kobzarev and Okun [38] showed that the surface energy density in the walls separating the various domains is sizable. Furthermore, because domain walls are two dimensional objects, the rate of total energy decrease as the Universe cools goes down only like T. Hence, eventually, the energy density in the domain walls far exceeds the closure density of the Universe, unless processes exist to annihilate the domain walls. If no such processes exist, typically, one expects [38]

$$\rho_{wall} \sim < \phi >^3 T \sim 10^{-7} GeV^4 \tag{70}$$

which is enormously bigger than the closure density of the Universe today ($\rho_{closure} \sim 10^{-46} GeV^4$)!

The only sure cure against the domain wall problem, if CP is spontaneously broken, is to imagine that the CP phase transition occurs before an inflationary period. In that case, the domains get enormously stretched and the domain wall energy is irrelevant, since we live in one domain. From this point of view, therefore, a solution of the strong CP problem based on the soft CP scenario makes sense only if the spontaneous CP breaking occurs at a very high scale. This scale, for instance, could be that of a grand unified theory.

Even if CP is broken spontaneously at a GUT scale, it is not immediately obvious that the strong CP problem is solved. However, it is clear that things are considerably better, since the objects whose vacuum expectation values break CP spontaneously must be $SU(2) \times U(1)$ singlets. Hence they have no direct coupling to the quarks (and leptons) which can give rise to a tree level contribution to $ArgdetM$. The difficulty here is not so much of getting rid of the strong CP problem, but of finding a way to generate some CP violation in the kaon system! Once one details how the ordinary CP violation comes about, then one can proceed to estimate the loop contributions to $ArgdetM$ and check if they are below the neutron dipole moment bound of 10^{-9}.

Barr and Zee [39] have emphasized that there are two possible strategies one can follow to build these soft CP models. To wit:

1. One can try to arrange the GUT soft CP breaking so that, although $ArgdetM = 0$ at tree level, the mass matrix for the quarks is in fact complex. In this case, the low energy CP violation can effectively be of the Kobayashi Maskawa type.

2. One has a GUT soft CP breaking model leading to a quark mass matrix which is real at tree level. In this case, the low energy CP violation must arise from interactions which are not directly connected with the quark mass matrix.

Models of the first type were discovered by Nelson [40] and further explored by Barr [41] and Nelson [42]. As will be seen below, they are examples of perfectly consistent solutions to the strong CP problem, provided that one admits a hierarchy between the masses of the superheavy fermions in the model and the GUT scale. Apart from this blemish, these models suffer from a lack of testable consequences. The strong CP problem is solved by some (obscure) physics at a high scale, with no direct connection with the low energy theory! In this respect, the second class of models suggested above are more interesting, since they require an alternation explanation of the observed low energy CP violation. However, the recent experimental observation of a non vanishing value of ϵ' [36], with a value in the range predicted by the standard model, renders most of these alternative models moot. Given these facts, in my opinion, it is unlikely that the strong CP problem is solved via soft CP breaking at the GUT scale. Nevertheless, it is worthwhile at least to sketch some of the basic ingredients of these soft CP models, since they are quite nice.

The models suggested by Nelson [40] make use of mixing between the quarks and leptons and some heavy fermions, to introduce the CP phase arising in the GUT theory into the low energy theory. The general structure of these models has been analyzed by Barr [41] and is quite natural, (although it requires a bit of magic!) Basically one makes use of two ingredients:

1. The heavy fermions come in complex conjugate representations C and $\bar{C}$ and have only real mass terms connecting C with $\bar{C}$. Similarly, there are only Yukawa induced real mass term for the light fermions f (the quarks and leptons) among themselves:

2. There is a complex effective mass term between the C fermions and the f fermions, but no such term between f and $\bar{C}$. It is here that the model dependence enters, since it requires allowing only particular heavy fermions and Higgs representations to exist in the theory. It is here also that the soft CP breaking is introduced, via some singlet field, or fields, with complex vacuum expectation values.

A nice example has been provided by Barr [41], using an $SO(10)$ GUT. The heavy fermions are chosen to lie in the 126 and $\overline{126}$ representations, so that

$$f = N_g\, 16; \quad C = \overline{126}; \quad \bar{C} = 126 \tag{71}$$

The Higgs fields include a 10_H, to give mass to the light fermions, and a 45_H and (several) 16_H to break $SO(10) \to SU(3) \times SU(2) \times U(1)$. One must assume that the $SU(3) \times SU(2) \times U(1)$ singlet component of one (or more) of the 16_H is complex, so that CP is softly broken. The complex mass term joining f with C is provided by the coupling with the 16_H, but there is no $f\bar{C}$, CC or $\bar{C}\bar{C}$ mass terms, with the Higgs representations chosen. Furthermore, with only one 10_H and 45_H the mass term for ff and $C\bar{C}$ are purely real since, by the assumption of CP conservation of the Lagrangian, the Yukawa couplings must be real.

The general structure of the mass matrix for the Barr model [43] illustrates how one can have $ArgdetM = 0$ (at tree level) and yet, effectively, the mixing matrix of the light fermions is of the Kobayashi-Maskawa type. According to the above discussion, the mass matrix for the fermions reads, schematically:

$$(16\ \overline{126}\ 126) \left| \begin{matrix} m & 0 & M_1 e^{i\phi} \\ M_1 e^{i\phi} & M_2 & 0 \\ 0 & 0 & M_2 \end{matrix} \right| \begin{pmatrix} 16 \\ 126 \\ \overline{126} \end{pmatrix} \tag{72}$$

This matrix, obviously, has a real determinant. However, because of the mixing between the 16 and the $\overline{126}$ fermions, the submatrix which serves to diagonalize the light fermion sector is no longer real but complex [6].

Although, by construction, one has $ArgdetM = 0$ at tree level, at the loop level one will eventually generate a mass matrix phase. For the example above, it is clear that any 16-126 mass term will contribute to $ArgdetM$. Such a mass term can be induced, at one loop, by the Higgs exchange diagram indicated in Fig 2 The order of magnitude of the contribution of this diagram to the mass matrix is (here λ is the 4-Higgs coupling)

$$M_{loop} \sim M_1 \left(\frac{m M_2}{M_H^2}\right) \frac{\lambda}{16\pi^2} e^{i\phi'} \tag{73}$$

which leads to

$$ArgdetM \sim \frac{\lambda}{16\pi^2} \left(\frac{M_1}{M_H}\right)^2 (\phi + \phi') \tag{74}$$

It is clear that, if the contribution (74) is to be less than 10^{-9}, one requires a hierarchy between the masses of the superheavy fermions and the corresponding Higgs bosons. This hierarchy is technically natural, since one gains chiral symmetries in the limit as the light and heavy fermions decouple [40]. However, one must admit that, without a priori prejudices, one would not have expected (74) to be so small!

Let me briefly discuss the structure of the second class of soft CP models, in which the quark mass matrix is real at tree level, and CP is broken spontaneously at the GUT scale. These models [39] require that there be a filtering down of the CP phase to colored Higgs

[6] Actually, it is necessary to have more than one phase, in this particular model, for this to really happen. Also one should realize that this submatrix is not strictly unitary. However, the departures from unitarity are of order $(m/M)^2$ and therefore are totally negligible, if M is of order the GUT scale

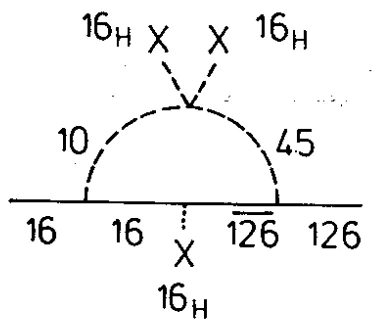


Figure 2: A one loop contribution to $ArgdetM$ in Barr's model

bosons, which can then couple to quarks [7]. Since the quark mass matrix is real, in these models, CP violation in the Kaon system is not of the Kobayashi Maskawa type. Thus these models must accomplish two things: they must provide an alternative explanation of low energy CP violation and they must be such that, at the loop level, the induced $ArgdetM$ is not too big.

Barr [43] has surveyed the classes of colored Higgs fields which may be useful for this type of model building. I shall not repeat his analysis here, but only use the "best" two of his examples to illustrate some of their general features - and also some of their drawbacks. The first example makes use of a color 6 of Higgs [39] [44] and leads to a superweak model of CP violation. This model is ruled out by the non zero measurement of ϵ' [36]. The second example has leptoquark Higgses and, in principle, can have a non zero ϵ'. However, how big ϵ' can be is severely restricted by rare Kaon decay bounds [45].

It is clear that a sextet colored scalar χ, which is either a triplet or a singlet under $SU(2)$, can provide a non zero value for ϵ, via the diagram of Fig 3. This requires that there be either a phase in the χ propagator of in the χqq vertex. Obviously this exchange provides only a $\Delta S = 2$ CP violation. Thus these models are of a superweak nature and, hence are now in trouble experimentally. Apart from this deadly fact, these models do have all the requirements to solve the strong CP problem. The CP phase connected with χ can originate spontaneously at the GUT scale. Furthermore, the induced value for $ArgdetM$ at the loop level, via χ exchange, is likely to be less than 10^{-9}. For instance, χ can be part of some 15 dimensional Higgs of $SU(5)$ and the propagator phase in Fig 3 can arise from a complex phase in the χ mass matrix, coming from the expectation of an $SU(15)$ singlet Higgs in the coupling $15^A 15^B 1$ [8] [39] [46]. Because χ only couples to $q_R q_R$, its first contribution to ArgdetM will occur at the two loop level and thus is probably sufficiently

[7] There is an alternative possibility which makes use instead of exotic light fermions [43], but I shall not discuss it further here

[8] One needs more than one 15 to get a propagator phase

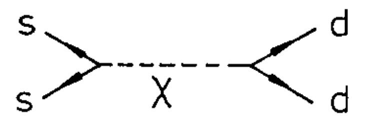


Figure 3: CP violation via exchange of a diquark Higgs. χ couples only to $q_L q_L$ or $q_R q_R$

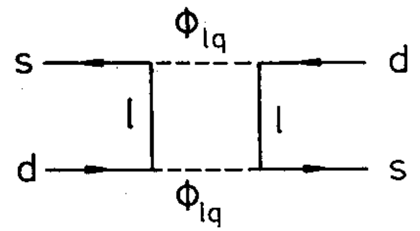
suppressed [39].

This same kind of scenario can also be implemented with leptoquarks. Here, however, there are some important differences. ϵ will now not be given by direct leptoquark exchange, but by the box diagram of Fig 4a. The leptoquark Penguin diagram of Fig 4b will then provide a contribution to ϵ'. So these models are not purely superweak. However, because the decay $K \to \mu e$, for example, can be induced by these same leptoquarks, there are strong bounds on the size of the allowed couplings. Models where the low energy violation is due to scalar leptoquarks have been investigated in detail by Hall and Randall [45], who worried in particular about the constraints coming from Kaon physics. To avoid problems with proton decay only leptoquarks which are $SU(2)$ doublets are considered. These scalars will only couple to quarks of a given helicity (either q_L or q_R) and hence their exchange will generate a non zero $ArgdetM$ [43] only beyond one loop. Indeed, as Hall and Randall [45] show, the first non vanishing contribution arises at the three loop level. These scalars can also be embedded in a GUT Higgs representation. Therefore the associated low energy phase can be imagined as being due to spontaneous CP breaking at the GUT scale.

If the ϵ parameter in Kaon physics is due to scalar leptoquark exchange, the strength of the imaginary part of the Yukawa coupling Γ_{lq} of the leptoquarks ϕ_{lq} to the quarks and leptons, as well as the mass M_ϕ of the lightest ϕ_{lq}, are constrained. The analysis of Hall and Randall [45] gives

$$\frac{Im(\Gamma_{lq}\Gamma^{\dagger}_{lq})^2_{sd}}{M^2_\phi} \simeq 5 \times 10^{-12} GeV^{-2} \qquad (75)$$

The measured value for $|\epsilon'/\epsilon|$ provides a new constraint on these parameters, since the Yukawa couplings enter quadratically and not quartically in the Penguin diagram of Fig 4b. Although Hall and Randall [45] estimated the contribution of the leptoquark Penguin diagram, since no value for ϵ'/ϵ was known at the time of their writing, they obviously did not use it as a constraint on Γ_{lq} and M_ϕ. An actual value for ϵ'/ϵ, however, severly



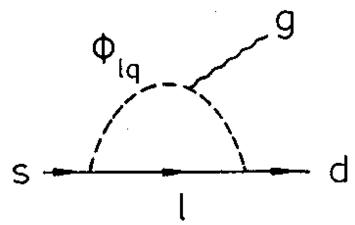


Figure 4: Leptoquark diagrams contributing to ϵ (a) and to ϵ' (b)

limits these quantities. Using again the results of [45] and the value of ϵ'/ϵ measured by the NA31 collaboration at CERN [36], one has

$$\frac{Im(\Gamma_{lq}\Gamma^{\dagger}_{lq})_{sd}}{M^2_{\phi}} \simeq 7 \times 10^{-8} GeV^{-2} \qquad (76)$$

Taking these equations at face value implies that, in order of magnitude,

$$Im(\Gamma_{lq}\Gamma^{\dagger}_{lq})_{sd} \simeq 10^{-4} \quad M_{\phi} \sim 30 GeV \qquad (77)$$

These numbers are too big to avoid the existing bounds on rare Kaon decay processes and $e-\mu$ conversion in Nuclei. Typically, as can be seen from figure 6 of [45], to avoid getting too large a rate for the process $K_L \rightarrow \mu^+\mu^-$ one must have that $M_{\phi} \geq 10^3 GeV$. Furthermore, the bound from $\mu \rightarrow e$ conversion is even stronger. One might be able to tolerate a leptoquark in the hundred Gev range [45] [47], but not one as light as that required by the new ϵ'/ϵ measurment. Of course, since the actual structure of the leptoquark Yukawa coupling matrix is unknown, it is always possible to plead special circumstances and avoid the above conclusion. In this case, however, one is to expect spectacular fireworks from HERA, where these leptoquarks will be copiously produced as s-channel resonances [48].

My personal conclusions on the soft CP option for solving the strong CP problem are not very sanguine - but it may be that I am prejudiced! What is wrong here, I believe, is not the general idea, but its implementation in terms of practical models. Asking not to have a large phase in the determinant of the quark mass matrix and to have both $\Delta S = 2$ and $\Delta S = 1$ CP violation are requirement that are not easily reconcilable. Furthermore, supposing that the soft CP breaking really occurs at the GUT scale, necessitates quite a large superstructure for which we have no evidence at all [9]. It is, of course, possible that CP violation at low energies occurs by some other interaction, which at the same time avoids the strong CP problem. However, the probability of this being the case will continue to decrease if also other CP violating phenomena - like those involving B mesons - are seen to fit the general Kobayashi Maskawa pattern.

[9]This is not quite true, since in these kinds of models there is a more direct relation between the Universe baryon asymmetry and CP violation in the Kaon system. However, in practice this connection is not without its own difficulties [49]

4 The Chiral Solution to the Strong CP Problem

In this Section, I want to discuss various theoretical, phenomenological and experimental issues connected with the chiral solution to the strong CP problem. We saw earlier that, by imposing a global chiral symmetry ($U(1)_{PQ}$) on the standard model Lagrangian [31], the strong CP problem disappeared, since the CP violating $\bar{\Theta}$-term gets replaced by the dynamical interaction of the axion field with $F\tilde{F}$ (c.f. Eq(65)). Obviously, however, for the chiral solution to be really a solution to the strong CP problem, axions must exist! Unfortunately, we have up to now no experimental indications for axions. In fact as I will discuss in more detail below, if the scale of the $U(1)_{PQ}$ breaking, v_{PQ}, is of order of the weak symmetry breaking scale.

$$v = (\sqrt{2}G_F)^{-\frac{1}{2}} \simeq 250 GeV \qquad (78)$$

one can establish experimentally that axions do not exist.

Since v_{PQ} characterizes the strength of the axions interactions, it is clear that as v_{PQ} becomes sufficiently larger than v, axions become progressively "invisible" to ordinary experiments. Remarkably, however, such invisible axions can play important astrophysical and cosmological roles. These non particle physics considerations turn out to restrict the allowed values of v_{PQ} to a rather narrow range. If v_{PQ} is above the weak scale, axions are so light that they can play an important role in the energy loss of stars. Indeed, soon after axions were proposed, it was realized [50] that unless v_{PQ} was really very big ($v_{PQ} \geq 10^8 GeV >> v$) the presence of axions would totally distort normal stellar evolution. The scale v_{PQ}, however, cannot be arbitrarily large [10], since cosmology provides an upper bound for it. This bound arises because the energy density stored in coherent axion oscillations, which is proportional to v_{PQ}, does not dissipate quickly with temperature. Hence, for v_{PQ} sufficiently big ($v_{PQ} \sim 10^{12} GeV$), the axion stored energy today would exceed the Universe's critical energy density [51].

The above qualitative considerations are independent of the precise couplings of the axion. However, to rule out all "visible" axion models and to establish the allowed range of v_{PQ}, where "invisible" axion models can still be countenanced, requires that one be a little bit more precise with the axion interactions than we have been in Eq(65). It turns out that there is some mild model dependence of the axion couplings to fermions, which eventually also emerges in the coupling of the axion to the gauge fields (e.g. the parameter ξ in Eq (62)). I want to, briefly, indicate what is the reason for this. This discussion will have the added benefit that it will allow the derivation of a mass formula for the axion and demonstrate the way in which the phase dependence of the vacuum expectation values adjust, so as to solve the $U(1)_A$ problem.

[10]Although $v_{PQ} \geq M_{Planck}$ also would not make much sense physically

4.1 Standard and Variant Visible Axion Models

It is useful to develop the relevant formalism in the simplest axion model and then detail the changes necessary for other situations. The most direct way to impose a $U(1)_{PQ}$ symmetry in the standard model [31] is to replace the ordinary Higgs doublet Φ and its charge conjugate $\tilde{\Phi} = i\tau_2\Phi^*$ by two separate doublets Φ_1 and Φ_2. These fields have the same hypercharge as Φ and $\tilde{\Phi}$ but, since they are independent fields, they can now carry another $U(1)$ charge. The excitation associated with this $U(1)_{PQ}$ charge is the axion. Since this is all we are really interested in, it is convenient to isolate this overall phase field in Φ_1 and Φ_2 and drop all other excitations. In so doing, however, we must make sure not to mix pieces of the axion with the hypercharge phase field, which eventually gets eaten by the Z^0. If the fields Φ_i have vacuum expectation values $\frac{1}{\sqrt{2}}v_i$, it is to see that one should write [52]

$$\Phi_1 = \frac{1}{\sqrt{2}}v_1 e^{\frac{iax}{v}}\begin{bmatrix}1\\0\end{bmatrix} \qquad \Phi_2 = \frac{1}{\sqrt{2}}v_2 e^{\frac{ia}{xv}}\begin{bmatrix}0\\1\end{bmatrix} \tag{79}$$

where

$$x = \frac{v_2}{v_1}; \qquad v = \sqrt{v_1^2 + v_2^2} \tag{80}$$

and a is the physical axion field [11]. In this simple model, $v_{PQ} = v$, so that Eq(61) becomes, for a finite $U(1)_{PQ}$ transformation:

$$a \to a + \alpha v \tag{81}$$

Having fixed the $U(1)_{PQ}$ transformation of the Higgs fields, the transformation laws for the fermions are now determined by demanding that the Yukawa interactions be also $U(1)_{PQ}$ invariant. A particularly convenient definition for the PQ symmetry transformation laws for the fermions [53] is to assume that all left handed doublet fields are PQ singlets. Then the PQ charges of the right handed fermions are fixed directly by specifying to which Higgs field they couple. In the model discussed in [31], all charge $\frac{2}{3}$ right handed quark fields couple to Φ_1 and all charge $-\frac{1}{3}$ right handed quark fields and right handed leptons couple to Φ_2. Thus their $U(1)_{PQ}$ transformations are simply

$$\begin{aligned} u_{Ri} &\to e^{-i\alpha x}u_{Ri} \\ d_{Ri} &\to e^{-\frac{i\alpha}{x}}d_{Ri} \\ l_{Ri} &\to e^{-\frac{i\alpha}{x}}l_{Ri} \end{aligned} \tag{82}$$

It is, however, possible to imagine changing these assignements by replacing in some Yukawa interactions, say, $\Phi_1 \to \tilde{\Phi}_2$, since these fields have the same hypercharge. These variant models [54], obviously, have different PQ assignements than those given in Eq(82), and so can give rise to different detailed predictions for axions. I will return to this point shortly.

From Eqs (81) and (82) the current associated with the $U(1)_{PQ}$ transformation is simply

$$\begin{aligned} J^\mu_{PQ} &= -v\partial^\mu a + x\sum_i^{N_g}\bar{u}_{Ri}\gamma^\mu u_{Ri} \\ &+ \frac{1}{x}\sum_i^{N_g}\bar{d}_{Ri}\gamma^\mu d_{Ri} + \frac{1}{x}\sum_i^{N_g}\bar{l}_{Ri}\gamma^\mu l_{Ri} \end{aligned} \tag{83}$$

[11] That is, a is what we denoted as a_{phys} in Eq(64). We have dropped the subscript here, so as not to encumber the notation unnecesarily. Note that, eventhough we are assuming that $\bar{\Theta}$ has already been set to zero, if we need to detail the $\bar{\Theta}$ dependence of quantities, it suffices to look at their a dependence

This current has a chiral color anomaly [3] which is proportional to the number of fermion families $N_g = \frac{1}{2}N_f$ and depends on x:

$$\partial_\mu J^\mu_{PQ} = [N_g(x + \frac{1}{x})]\frac{g^2}{32\pi^2}F^{\mu\nu}_a\tilde{F}_{a\mu\nu} \tag{84}$$

The quantity in the square brackets above is clearly model dependent, since it depends on the way we assumed the Higgs fields Φ_1 and Φ_2 couple to the quarks. For any other assignement, however, the relevant changes to Eq (84) are obvious. In particular, it is easy to show that if we always couple all right handed charge $-\frac{1}{3}$ quarks to Φ_2, thereby avoiding a severe flavor changing neutral current problem [54], then the factor N_g in Eq (84), in an arbitrary variant model, is replaced by N_u, the number of right handed charge $\frac{2}{3}$ quarks coupled to Φ_1 [53].

It follows, from the transformation of the axion field under $U(1)_{PQ}$, that the parameter ξ in the effective Lagrangian of Eq(62) is just the coefficient in the square brackets in Eq(84). More formally [53], consider making a local transformation on the right handed fermion fields, to remove the axion field from the Yukawa interactions:

$$\begin{aligned} u_{Ri} &\to e^{-\frac{iax}{v}}u_{Ri} \\ d_{Ri} &\to e^{-\frac{ia}{vx}}d_{Ri} \\ l_{Ri} &\to e^{-\frac{ia}{vx}}l_{Ri} \end{aligned} \tag{85}$$

This transformation will generate derivative interaction of the axion from the fermion kinetic energy terms. This is the term $\mathcal{L}(\partial_\mu a, \psi)$ of Eq(62), which for the model at hand reads:

$$\begin{aligned} \mathcal{L}(\partial_\mu a, \psi) &= \frac{1}{v}\partial_\mu a[x\sum_i^{N_g}\bar{u}_{Ri}\gamma^\mu u_{Ri} \\ &+ \frac{1}{x}\sum_i^{N_g}\bar{d}_{Ri}\gamma^\mu d_{Ri} + \frac{1}{x}\sum_i^{N_g}\bar{l}_{Ri}\gamma^\mu l_{Ri}] \end{aligned} \tag{86}$$

However, in making the chiral transformation (85) one generates also anomalous non derivative interaction of the axion to the $SU(3)$ and $U(1)$ fields, through loops containing the right handed fermion fields. One finds that [53]

$$\mathcal{L}_{anomaly} = \frac{a}{v}\left\{[N_g(x + \frac{1}{x})]\frac{g^2}{32\pi^2}F^{\mu\nu}_a\tilde{F}_{a\mu\nu} + [N_g(\frac{4}{3}x + \frac{1}{3x} + \frac{1}{x})]\frac{e^2}{16\pi^2}B^{\mu\nu}\tilde{B}_{\mu\nu}\right\} \tag{87}$$

In the above $B^{\mu\nu}$ is the field strength of the $U(1)$ weak gauge field, coupled to the right handed fermions

$$B^{\mu\nu} = F^{\mu\nu}_{em} - \tan\Theta_W F^{\mu\nu}_Z \tag{88}$$

For clarity, in Eq(87), I have displayed explicitly the contribution of the charge $\frac{2}{3}$, charge$-\frac{1}{3}$ and charge -1 fields. Different assignments of the PQ transformation for these fields [12],

[12] i.e. different choices of Yukawa interactions

obviously modify the strength of the interactions in Eq(87). For instance, if one had coupled all right handed leptons to Φ_1, then the last contribution in Eq(87) instead of being x^{-1} would be $-x$. So the axion couplings to photons and gluons are (somewhat) model dependent.

The presence of the gluon anomaly, makes the extraction of the interactions of axions with light hadrons not so straightforward [13]. One approach to this problem, pioneered by Bardeen and Tye [52], is to construct an anomaly free current out of J^μ_{PQ}, which has a divergence which vanishes as the light quark masses vanish. This current, because it has a soft divergence, can then be used for current algebra calculations of axion properties, including the axion mass. I shall not pursue this method here in detail. However, I want to give the form of the modified Bardeen-Tye current, since it will emerge in a different guise, in the discussion that follows. Consider, for simplicity, that only the u and d quarks are light. Then the Bardeen-Tye current [52] reads:

$$\tilde{J}^\mu_{PQ} = J^\mu_{PQ} - \frac{1}{2}N_g(x+\frac{1}{x})[\frac{m_d}{(m_u+m_d)}\bar{u}\gamma^\mu\gamma_5 u + \frac{m_u}{(m_u+m_d)}\bar{d}\gamma^\mu\gamma_5 d] \quad (89)$$

Clearly $\tilde{J}^\mu_{PQ}$ has no anomaly, by construction. Furthermore, it is not hard to see that the divergence of this current indeed vanishes both as $m_u \to 0$ and as $m_d \to 0$. I note also, for future reference, that $\tilde{J}^\mu_{PQ}$ contains in it both isoscalar and isovector axial currents. Defining

$$A^\mu_0 = \frac{1}{2}[\bar{u}\gamma^\mu\gamma_5 u + \bar{d}\gamma^\mu\gamma_5 d]; \qquad A^\mu_3 = \frac{1}{2}[\bar{u}\gamma^\mu\gamma_5 u - \bar{d}\gamma^\mu\gamma_5 d], \quad (90)$$

the axial piece of the Bardeen-Tye current for the light quarks reads:

$$(\tilde{J}^\mu_{PQ})_{axial} = \lambda_0 A^\mu_0 + \lambda_3 A^\mu_3 \quad (91)$$

where

$$\begin{aligned}\lambda_0 &= \frac{1}{2}[1-N_g](x+\frac{1}{x}) \\ \lambda_3 &= \frac{1}{2}[(x-\frac{1}{x}) - N_g(x+\frac{1}{x})\frac{(m_d-m_u)}{(m_d+m_u)}]\end{aligned} \quad (92)$$

The above mixing parameters λ_0 and λ_3 are those for the model of [31]. Again, if one changes the Higgs assignments, also the λ_i will vary [14].

Rather than using current algebra methods, it is perhaps more illustrative to try to derive the interactions of axions with light hadrons by using an effective Lagrangian technique [55]. I will concentrate here only in the interaction of axions with light mesons. For this purpose, it suffices to augment the usual meson chiral effective Lagrangian [56] by appropriate axions contributions, which respect all the symmetries of the problem at hand. If one again considers only the u and d quarks as light, the chiral Lagrangian will contain three separate pieces [53]. The first piece is just the usual $U(2) \times U(2)$ invariant

[13]The interactions of axions with heavy quarks Q can simply be computed from Eq(86) or, equivalently, from the relevant Yukawa couplings. The strength of these interactions are of order $\frac{m_Q}{v}$. Thus, as long as $m_Q >> \Lambda_{QCD}$, the anomaly will only give small corrections to the overall couplings

[14]For general formulas, for variant models, see [53]

chiral Lagrangian, describing the strong interactions of the pions and η, plus the axion kinetic energy

$$\mathcal{L}_{chiral} = -\frac{1}{4}F_\pi^2 Tr\partial_\mu U^\dagger\partial^\mu U - \frac{1}{2}\partial_\mu a\partial^\mu a \quad (93)$$

Here the chiral field U is given by

$$U = exp[i\frac{\vec{\tau}\cdot\vec{\pi}+\eta}{F_\pi}] \quad (94)$$

The second contribution in the effective Lagrangian mimics the effect of the Yukawa interactions, which give masses to the quarks as a result of the $SU(2)\times U(1)$ breakdown. At the meson level, the presence of quark mass terms breaks the $U(2)\times U(2)$ symmetry explicitly. However, since the Yukawa interactions preserve the PQ symmetry, one must ensure that the mass breaking term one writes down for the mesons also preserves this symmetry. It is easy to see that the desired term is

$$\mathcal{L}_{mass\ breaking} = -\frac{c_0}{2}Tr[UAM + M^\dagger A^\dagger U^\dagger] \quad (95)$$

where

$$M = \begin{bmatrix} m_u & 0 \\ 0 & m_d \end{bmatrix} \quad (96)$$

and

$$A = \begin{bmatrix} e^{-\frac{ixa}{v}} & 0 \\ 0 & e^{-\frac{ia}{xv}} \end{bmatrix} \quad (97)$$

Note that the PQ invariance is preserved by Eq(95), since the shift of Eq(81) is compensated by the (right handed) $U(2)$ transformation

$$U \to U\begin{bmatrix} e^{ix\alpha} & 0 \\ 0 & e^{\frac{i\alpha}{x}} \end{bmatrix} \quad (98)$$

The last piece of the effective Lagrangian contains terms which reflect the anomaly structure of the quark theory. Since we are considering explicitly the axion interactions with the light mesons (Eq(95)), we need only incorporate the heavy quark contributions of Eq(87). Thus, keeping only the gluon piece:

$$\mathcal{L}_{a\ anomaly} = \frac{a}{v}[(N_g-1)(x+\frac{1}{x})]\frac{g^2}{32\pi^2}F^{\mu\nu}_a\tilde{F}_{a\mu\nu} \quad (99)$$

In addition, the effective Lagrangian must also reflect the strong anomaly associated with the $U(1)_A$ current. Since the η meson shifts under these transformation by F_π, the relevant anomaly piece is [53]

$$\mathcal{L}_{\eta\ anomaly} = \frac{\eta}{F_\pi}[2]\frac{g^2}{32\pi^2}F^{\mu\nu}_a\tilde{F}_{a\mu\nu} \quad (100)$$

The effect of the strong anomaly above, when the gauge fields are integrated out, is supposed to give a mass term for the η, thereby solving the $U(1)_A$ problem [15]. For the case at

[15]Formally [14], one adds a "kinetic energy" term for $F\tilde{F}$ and eliminates this quantity from the theory, by using the equations of motion.

hand, since also the axion field couples to $F\tilde{F}$, the combined effect of Eqs(99) and (100) is to produce an effective mass term for a combination of a and η:

$$\mathcal{L}^{eff}_{anomaly} = -\frac{1}{2}m_0^2[\eta + \frac{F_\pi}{v}\{\frac{(N_g-1)}{2}(x+\frac{1}{x})\}a]^2 \quad (101)$$

The parameter m_0 is, to a good approximation, the mass for the physical η and so $m_0^2 >> m_\pi^2$. Whence, it follows that the linear combination of fields occuring in Eq(101) essentially decouples from the theory.

The presence of the mass breaking term, Eq(95), and of the anomaly term, Eq(101), gives mass for all the excitations: π, η and a. In the absence of the anomaly term, the presence of the PQ symmetry allows some linear combination of π^0, η and a to remain massless. In the absence of the mass breaking term, the π and the orthogonal combination to the field appearing in Eq(101) remain massless. When both terms are included, there are no more massless fields. In the charged pion sector the mass term

$$\mathcal{L}_{charged} = -\frac{(m_u+m_d)}{F_\pi^2}c_0\pi^+\pi^- \quad (102)$$

identifies

$$c_0 = \frac{F_\pi^2 m_\pi^2}{(m_u+m_d)} \quad (103)$$

In the neutral boson sector, the calculation of the mass eigenstates and physical fields is facilitated by the fact that $m_0^2 >> m_\pi^2$. Using Eq(103), one has

$$\begin{aligned}\mathcal{L}_{neutral} &= -\frac{1}{2}m_\pi^2\{\frac{m_u}{(m_u+m_d)}[\pi^0+\eta-\frac{axF_\pi}{v}]^2 + \frac{m_d}{(m_u+m_d)}[-\pi^0+\eta-\frac{aF_\pi}{xv}]^2\} \\ &\quad -\frac{1}{2}m_0^2[\eta+\frac{F_\pi}{v}\{\frac{(N_g-1)}{2}(x+\frac{1}{x})\}a]^2 \end{aligned} \quad (104)$$

For $m_0^2 >> m_\pi^2$ one readily finds that the axion mass is given by [52] [55]:

$$m_a^2 = \frac{m_\pi^2 F_\pi^2}{v^2}N_g^2(x+\frac{1}{x})^2\frac{m_u m_d}{(m_u+m_d)^2} \quad (105)$$

and that the field a contains a small admixture of the physical π^0 and η fields:

$$a \simeq a_{phys} - \xi_{a\pi}\pi^0_{phys} - \xi_{a\eta}\eta_{phys} \quad (106)$$

where

$$\xi_{a\pi} = \lambda_3\frac{F_\pi}{v}; \quad \xi_{a\eta} = \lambda_0\frac{F_\pi}{v} \quad (107)$$

with the λ_i being the mixing parameters defined in Eq(92).

Several remarks are in order:

1. The formula (105) for the axion mass vanishes both as m_u and/or $m_d \to 0$. This is as expected, since in this case one has more symmetry. The precise structure in Eq(105) applies to the model of [31]. However, the modification for variant models is rather trivial: one just replaces N_g by the number of families which feel the PQ

symmetry. Because of the ratio $\frac{F_\pi}{v}$, the axion's mass is very much smaller than that of the pion. Using [27]

$$\frac{(m_d - m_u)}{(m_d + m_u)} \simeq 0.26 \quad (108)$$

one has, numerically,

$$m_a \simeq 25N_g(x+\frac{1}{x})KeV \quad (109)$$

2. The identifications (106) and (107) clarify the structure of the Bardeen Tye current [52]. Including the axion piece, the axial part of this current is

$$\tilde{J}^\mu_{PQ} = -v\partial^\mu a + \lambda_3 A_3^\mu + \lambda_0 A_0^\mu = -v\partial^\mu a - F_\pi\lambda_3\partial^\mu\pi^0 - F_\pi\lambda_0\partial_\mu\eta, \quad (110)$$

where the second line is the form taken by the axial currents in the effective Lagrangian. Therefore, in view of Eqs (106) and (107), the Bardeen Tye current is seen to contain only the physical axion contribution:

$$\tilde{J}^\mu_{PQ} = -v\partial^\mu a_{phys} \quad (111)$$

3. The effective Lagrangian constructed serves to exemplify how the $\bar{\Theta}$ dependence needed to solve the $U(1)_A$ problem obtains and how the dynamical cancellation of $\bar{\Theta}$ works. For the vacuum configuration, U becomes a pure phase matrix. Corresponding to the phase choice made at the quark level in Eq(46), one has [16]:

$$U = \begin{bmatrix} e^{i[\phi_u(\bar\Theta)+\frac{x<a>}{v}]} & 0 \\ 0 & e^{i[\phi_d(\bar\Theta)+\frac{<a>}{xv}]} \end{bmatrix} \quad (112)$$

Using the above and the identification (103), the vacuum energy density from Eq(95) is seen to be

$$\mathcal{H}^{vac}_{mass} = -m_\pi^2F_\pi^2[\frac{m_u}{(m_u+m_d)}cos\phi_u(\bar\Theta) + \frac{m_d}{(m_u+m_d)}cos\phi_d(\bar\Theta)] \quad (113)$$

If we consider the case of one generation ($N_g = 1$), to match our discussion in Sec 3, then the anomaly contribution contains only the phase information associated with the η field. Since, from Eq(112), one identifies:

$$\frac{2<\eta>}{F_\pi} = \phi_u(\bar\Theta) + \phi_d(\bar\Theta) + (x+\frac{1}{x})\frac{<a>}{v} \quad (114)$$

the anomaly contribution to the vacuum energy is

$$\mathcal{H}^{vac}_{anomaly} = \frac{1}{4}m_0^2F_\pi^2[\phi_u(\bar\Theta) + \phi_d(\bar\Theta) + (x+\frac{1}{x})\frac{<a>}{v}]^2 \quad (115)$$

Minimization of the total vacuum energy with respect to $<a>$, yields the dynamical PQ condition (63) [31] with $\phi_u(\Theta) + \phi_d(\Theta) = \bar\Theta$. Minimization of the total vacuum energy with respect to $\bar\Theta$ reproduces Eq(53), which implies the phase dependence of Eq(54), necessary to solve the $U(1)_A$ problem [9] [14] [15].

[16] I have kept here the contribution proportional to the axion vacuum expectation value $<a>$, coming at the quark level from the Yukawa couplings, since I want to demonstrate the PQ mechanism

One sees from Eq(109) that unless x (or x^{-1}) is large, the axion mass will be below $2m_e$. In this case, the axion is very long lived, since it can only decay into two photons. The Lagrangian giving the coupling of the axion to the electromagnetic field is, essentially, that of Eq(87). However, because the unphysical axion field a mixes slightly with the π^0 and η, there are small modifications, which must be included in a more accurate treatment. These corrections can be easily computed, using the effective Lagrangian given above. For this purpose, I note that one can use Eq(87) to detail the contributions of the heavy quarks and leptons to the $a\gamma\gamma$ coupling:

$$\mathcal{L}^{Q,l}_{a\gamma\gamma} = \frac{a_{phys}}{v}\{(N_g-1)[\frac{4x}{3}+\frac{1}{3x}] + N_g[\frac{1}{x}]\}\frac{e^2}{16\pi^2}F^{\mu\nu}_{em}\tilde{F}^{em}_{\mu\nu} \quad (116)$$

In addition to this interaction, the effective coupling of the π^0 and the η to two photons, due to the electromagnetic anomaly [3]

$$\mathcal{L}_{em\ anomaly} = (\frac{\pi^0}{F_\pi}+\frac{5}{3}\frac{\eta}{F_\pi})\frac{e^2}{16\pi^2}F^{\mu\nu}_{em}\tilde{F}^{em}_{\mu\nu} \quad (117)$$

induces a further $a\gamma\gamma$ coupling, through the mixing (107):

$$\pi^0 \simeq \pi^0_{phys} + \xi_{a\pi}a_{phys} \qquad \eta \simeq \eta_{phys} + \xi_{a\eta}a_{phys} \quad (118)$$

One finds

$$\mathcal{L}^{mixing}_{a\gamma\gamma} = \frac{a_{phys}}{v}[\lambda_3 + \frac{5}{3}\lambda_0]\frac{e^2}{16\pi^2}F^{\mu\nu}_{em}\tilde{F}^{em}_{\mu\nu} \quad (119)$$

The quantity in the curly bracket in Eq(116), as well as λ_3 and λ_0, are model dependent. However, for any individual model, they can be easily ascertained [17]. For the original model of [31], one obtains for the total $a\gamma\gamma$ coupling:

$$\mathcal{L}_{a\gamma\gamma} = \frac{a_{phys}}{v}[N_g(x+\frac{1}{x})]\{\frac{4}{3} - \frac{4m_d+m_u}{3(m_d+m_u)}\}\frac{e^2}{16\pi^2}F^{\mu\nu}_{em}\tilde{F}^{em}_{\mu\nu} \quad (120)$$

In view of Eq(105), one can eliminate, in the above, the dependence on N_g and x in favour of the axion's mass:

$$\mathcal{L}_{a\gamma\gamma} = \frac{a_{phys}}{F_\pi}(\frac{m_a}{m_\pi})K_{a\gamma\gamma}\frac{e^2}{16\pi^2}F^{\mu\nu}_{em}\tilde{F}^{em}_{\mu\nu} \quad (121)$$

where, for the model of [31],

$$K_{a\gamma\gamma} = \sqrt{\frac{m_u}{m_d}} \simeq 0.75 \quad (122)$$

One sees from (117) and (121) that the axion lifetime into two photons is simply related to the π^0 lifetime. Since these decays are p-wave, one has:

$$\tau(a\to 2\gamma) = K^{-2}_{a\gamma\gamma}(\frac{m_\pi}{m_a})^5\tau(\pi^0\to 2\gamma) \simeq 0.38K^{-2}_{a\gamma\gamma}(\frac{100KeV}{m_a})^5 sec \quad (123)$$

[17] The literature is, unfortunately, full of mistakes concerning the $a\gamma\gamma$ couplings. Some of these are purely transcription errors [57], but in other instances there are real misconceptions. The correct answer, for the original PQ model, is given by Bardeen and Tye [52], and this is generalized to other models by Kaplan and by Srednicki [58]

Clearly, even an 1 MeV axion is quite long lived!

If $m_a > 2m_e$, however, the axion decays rapidly into electrons. Using Eq(86), one readily computes that

$$\tau(a\to e^+e^-) = \frac{8\pi v^2}{m_e^2[m_a^2-4m_e^2]^{1/2}}[x^2] \quad (124)$$

The factor of x^2, in the above, gets replaced by x^{-2} if one coupled the electrons to $\tilde{\Phi}_1$ rather than Φ_2 [18]. Numerically, one finds

$$\tau(a\to e^+e^-) \simeq 4\times 10^{-9}(\frac{1MeV}{m_a})[1-\frac{4m_e^2}{m_a^2}]^{-1/2}[x^2]sec \quad (125)$$

4.2 The Demise of Visible Axion Models

Unfortunately, a multitude of experiments carried out in the last ten years have conclusively ruled out both the long lived ($m_a < 2m_e$) and short lived ($m_a > 2m_e$) varieties of "visible" axion models. I will not summarize here all the experimental evidence against axions [60], but only concentrate on a few experiments which serve to deal a telling blow to visible axion models. The strong bound obtained at KEK on the process $K^+\to\pi^+ +$ Nothing [61]

$$B(K^+\to\pi^+ + \ Nothing) \leq 3.8\times 10^{-8} \quad (126)$$

with "Nothing" interpreted as a long lived axion which escapes detection, is perhaps the most damaging evidence against this excitation. Since this is a nonleptonic decay, the theoretical estimates for the branching fraction [62] are rather uncertain. However, they lie sufficiently above the bound (126) to be compelling. For example, using an effective Lagrangian to describe the $\Delta S = 1$ non leptonic K decays, and extending the model Lagrangian for the axion interactions here described to the case of 3 light flavors, Bardeen, Yanagida and I [53] find

$$A(K^+\to\pi^+ + a) \simeq \sqrt{2}A(K^0\to\pi^+ + \pi^-)\lambda_0(\frac{F_\pi}{v}) \quad (127)$$

This implies

$$B(K^+\to\pi^+ + a) \simeq 3\times 10^{-5}\lambda_0^2 = 3\times 10^{-5}(x+1/x)^2 \quad (128)$$

where the second line is the result for a 3 family version of the model [31]. There is some uncertainty in our calculation since Eq(127) relates the axion decay amplitude to a $\Delta I = 1/2$ enhanced amplitude, but our model cannot obtain this dynamical enhancement. However, we do not believe we are wrong by more than three orders of magnitude [19]!

[18] One can also contemplate models [59] where the leptons couple to a third Higgs doublet with no PQ charge. Then there is no tree level axion-lepton coupling and one is left with a, so called, hadronic axion [58]

[19] The only way to avoid the Kaon bound (126), for a long lived axion, is to construct a variant model where λ_0 vanishes. Such models exist [For example the variant models of [54] where only u_R couples to Φ_1, so that $N_u = 1$] but they can be ruled out by nuclear deexcitation experiments [63] involving isovector transitions, which are proportional to λ_3 [64]. [It is even possible to construct a variant model where both λ_0 and λ_3 vanish [65], if $x\simeq\sqrt{\frac{5}{3}}$. However, since there is only one active PQ family in this model, the mass of the resulting axion is so low, $m_a \simeq 60KeV$, that stellar evolution bounds [50] appear to effectively rule this possibility out]

A different set of bounds on axions arises from quarkonium decay. The Wilczek process [66] $Q\bar{Q} \to a\gamma$ measures the coupling of axions to heavy quarks. The ratio of this rate to the quarkonium rate into μ - pairs is independent of the details of the quarkonium wavefunction, although it has a rather large QCD correction [67]. One finds, including the order α_s QCD correction,

$$R_Q = \frac{\Gamma(Q\bar{Q} \to a\gamma)}{\Gamma(Q\bar{Q} \to \mu^+\mu^-)} = \frac{G_F m_Q^2}{\sqrt{2}\pi\alpha}[1 - \frac{[\frac{\pi^2}{2} + 8ln2]}{3\pi}\alpha_s]Z_Q^2 \qquad (129)$$

where Z_Q is related to the strenght of the axion coupling to the heavy quarks. In the model of [31], $Z_Q = x$ for all charge 2/3 heavy quarks and $Z_Q = x^{-1}$ for all charge $-1/3$ heavy quarks. Thus one cannot suppress both the axion decays of charmonium and bottomonium, at the same time. However, in variant models [54], it is possible to have $Z_Q = x^{-1}$ for both Ψ and Υ decays, so that if x is large both decays are suppressed.

Numerically, taking $m_b = 4.9GeV$, $m_c = 1.5GeV$, along with the measured value for the μ - pair branching ratios, and evaluating the large QCD correction at face value (using $\alpha_s = 0.17 \pm 0.03$ and $\alpha_s = 0.24 \pm 0.04$, respectively, at the Υ and the Ψ [68]), one finds

$$\begin{aligned} B(\Upsilon \to a\gamma) &= (2.0 \pm 0.7) \times 10^{-4} Z_b^2 \\ B(\Psi \to a\gamma) &= (3.7 \pm 0.8) \times 10^{-5} Z_c^2 \end{aligned} \qquad (130)$$

Experimentally axions have been searched for, in both these processes, but have not been found, leading to the bounds [69]

$$\begin{aligned} B(\Upsilon \to a\gamma) &< 3 \times 10^{-4} \\ B(\Psi \to a\gamma) &< 1.4 \times 10^{-5} \end{aligned} \qquad (131)$$

One sees that these bounds force Z_b and Z_c to be less or of the order of unity. Thus, for the standard axion model [31], only very light axions ($m_a \sim 150KeV$) are allowed by the quarkonia bounds. If one wants to contemplate short lived axion models, where $m_a > 2m_e$ and therefore x (or x^{-1}) is large, the quarkonia bounds can only be avoided by choosing $Z_b = Z_c = x^{-1}$ (or $Z_b = Z_c = x$) and one is, therefore, naturally lead to variant axion models [54].

Although the quarkonia bounds are ineffective for variant axion models, with $m_a > 2m_e$, these models have been ruled out by a combination of experiments which bound the mixing parameters λ_3 and λ_0 independently. A very strong bound on λ_3 is obtained from an elegant experiment at SIN which measured the process $\pi^+ \to e^+e^-e^+\nu_e$ with a branching ratio [70]

$$B(\pi^+ \to e^+e^-e^+\nu_e) = (3.4 \pm 0.5) \times 10^{-9} \qquad (132)$$

From an analysis of the e^+e^- invariant mass distribution, one can set a bound on the decay mode $\pi^+ \to ae^+\nu_e$, for axion masses in the MeV range and sufficiently short lifetimes ($\tau_a < 10^{-11} sec$). The typical bounds one finds are of the order of [71]

$$B(\pi^+ \to ae^+\nu_e) = (1 - 2) \times 10^{-10} \qquad (133)$$

The rate for the $\pi^+ \to ae^+\nu_e$ decay can be calculated theoretically in terms of the π^0-axion mixing parameter $\xi_{a\pi} = \lambda_3 \frac{F_\pi}{v}$ and one finds [72]

$$\Gamma(\pi^+ \to ae^+\nu_e) = \frac{G_F^2 m_\pi^5}{384\pi^3}(\xi_{\pi a})^2 \qquad (134)$$

The above implies a branching ratio

$$B(\pi^+ \to ae^+\nu_e) = 3 \times 10^{-9}(\lambda_3)^2 \qquad (135)$$

Thus the SIN bound requires λ_3 to be small: $\lambda_3 \leq 0.18 - 0.26$. Unfortunately, it turns out that, except for a very special case, λ_3 is very large in variant axion models, which have $m_a > 2m_e$ [53]. Typically, one has

$$|\lambda_3| \sim [\frac{m_a}{100KeV}] \qquad (136)$$

So the non observation of an axion signal in the $\pi^+ \to e^+e^-e^+\nu_e$ decay mode, essentially single handedly eliminates short lived "visible" axion models.

This conclusion is further strengthened by experiments which have searched for evidence of short lived axions in isoscalar nuclear transitions. The ratio of the rates of axion to photon nuclear deexcitation is quite reliably calculated [64]. Furthermore, for the case at hand, the axion decays are proportional to the isoscalar mixing parameter $\xi_{a\eta} = \lambda_0(\frac{F_\pi}{v})$. A particularly interesting example is afforded by the $M1$ decay of the 3.58 MeV $2^+(T = 0)$ state of ^{10}B to be $3^+(T = 0)$ ground state. Theoretically, one predicts a rate ratio [53]

$$\frac{\Gamma_a}{\Gamma_\gamma} = [1.15 \times 10^{-3}](\frac{k_a}{k})^3\lambda_0^2 \qquad (137)$$

where k_a and k are the momenta of the axion and photon, respectively. Experimentally, one has the bound

$$\frac{\Gamma_a}{\Gamma_\gamma} \leq [3.8 \times 10^{-3}](\frac{k_a}{k})^3 \qquad (138)$$

Whence, it follows that for axions in the kinematically allowed region, λ_0 cannot be too much bigger than unity.

This bound, per se, is not terribly meaningful, since there are variant models where λ_0 vanishes identically [54]. However, there is a model independent prediction for the difference between λ_0 and λ_3 in variant axion models, if $m_a > 2m_e$, [53]

$$(\lambda_0 - \lambda_3)^2 \simeq [\frac{m_a v}{m_\pi F_\pi}]^2 \frac{m_u}{m_d} \geq (15)^2 \qquad (139)$$

One sees, therefore, that it is not possible to have both λ_0 and λ_3 small. Hence the Boron bound (138), in conjunction with the pion decay bound (133), serves to completely eliminate short lived "visible" axion models.

The above considerations suggest that there is really no window for axions to exist, independently of whether they are short or long lived, <u>if</u> the scale of the breakdown of the PQ symmetry is the weak scale. Thus, if the solution to the strong CP puzzle is to be found by using an additional chiral symmetry, this symmetry must be broken at a large scale and the axion is of the invisible type. I want to conclude this Section, therefore, by detailing certain properties and various physical constraints of these kind of axion models.

4.3 Invisible Axion Models

All invisible axion models make use of some complex scalar field σ which:

- carries a PQcharge and possesses a large vacuum expectation value $<\sigma>=\frac{v_{PQ}}{\sqrt{2}}>>v$

- is an $SU(2)\times U(1)$ singlet

Both these conditions are obviously necessary, since one wants to split the scales of the PQ breaking from that of the electroweak breaking. Where invisible axion models differ is in the transformation properties of ordinary quarks (and leptons) under the extra chiral symmetry. Broadly speaking, two options have been considered:

1. The known quark and leptons do not feel the PQ symmetry, but there exists some (presumably very heavy) new quarks which carry a PQ charge. Prototype models of this type were first discussed by Kim [74] and by Shifman, Vainshtein and Zakharov [75] ($KSVZ$ axions)

2. Ordinary quarks and leptons carry a PQ charge, so that one necessitates again two Higgs doublets in the theory. Since the fermions in the theory do not couple directly to σ, they feel the PQ breaking only through the Higgs potential. Invisible axion models of this sort were first suggested by Dine, Fischler and Srednicki [76] and by Zhitnitskii [77] (DFSZ axions)

Both the $KSVZ$ and $DFSZ$ axion models solve the strong CP problem, since the vacuum expectation value of the axion field cancels the CP violating $\bar{\Theta}$ term in the Lagrangian [cf Eq (63)]. The resulting axions, however, have somewhat different couplings to the ordinary fermions and to photons. These couplings can be derived by following steps analogous to those used earlier for the standard axion [31]. So, in what follows, I will not give too many details [20].

4.3.1 The KSVZ Axion

The model [74] [75] introduces a heavy quark X which carries a PQ charge and therefore can couple to σ

$$\mathcal{L}=-h\bar{X}_L\sigma X_R-h^*\bar{X}_R\sigma^\dagger X_L \tag{140}$$

It is again convenient to adopt a definition of the PQ transformation in which only X_R carries a PQ charge and to concentrate on the axion content of σ, by freezing the radial component of this field:

$$\sigma=\frac{v_{PQ}}{\sqrt{2}}e^{i\frac{a}{v_{PQ}}} \tag{141}$$

Rotating the axion field away from (140) by a local X_R transformation, the anomalous interactions of the axion for this model are easily seen to be [cf Eq (87)]

$$\mathcal{L}_{anomaly}=\frac{a}{v_{PQ}}\left\{\frac{g^2}{32\pi^2}F_a^{\mu\nu}\tilde{F}_{a\mu\nu}+3[q_X^2]\frac{e^2}{16\pi^2}F_{em}^{\mu\nu}\tilde{F}_{\mu\nu}^{em}\right\} \tag{142}$$

[20]For a more extensive pedagogical discussion, see for example [58]. One should be aware that there exist a plethora of conventions in the literature so that, for example, the meaning of v_{PQ} is not universal. (e.g. $(v_{PQ})_{Srednicki}\equiv f_a=2v_{PQ}$)

where q_X is the electromagnetic charge of the heavy quark X, in units of e. The mixing of the axion with light hadrons can now be handled analogously to what was done previously. One finds, in this way, that the $KSVZ$ axion mass is

$$m_a^{KSVZ}=\frac{m_\pi F_\pi}{v_{PQ}}\frac{\sqrt{m_d m_u}}{(m_u+m_d)} \tag{143}$$

and that the mixing parameters with the π^0 and the η and the $\gamma-\gamma$ coupling of the axion are

$$\xi_{a\pi}^{KSVZ} = -\frac{1}{2}[\frac{m_d-m_u}{m_d+m_u}](\frac{F_\pi}{v_{PQ}}) \tag{144}$$

$$\xi_{a\eta}^{KSVZ} = -\frac{1}{2}(\frac{F_\pi}{v_{PQ}}) \tag{145}$$

$$C_{a\gamma\gamma}^{KSVZ} = \{3q_X^2-\frac{4m_d+m_u}{3(m_d+m_u)}\} \tag{146}$$

Note that the $KSVZ$ axion is a hadronic axion, in that it has no tree level coupling to electrons. A small axion-electron coupling, however, can be generated at the one loop level [78], using the axion coupling to photons. The coupling to nucleons follows immediately from the (chiral covariant) couplings of the π^0 and η to nucleons, via the mixing parameters of Eq (144) and of Eq (145) [53] [58] [63].

4.3.2 The $DFSZ$ axion

In this model [76] [77] the quarks and leptons have the same PQ symmetry properties as in the standard axion model [31] [21]. The two Higgs fields Φ_1 and Φ_2 necessary to allow the implementation of a PQ symmetry at the quark level are coupled to the Higgs field σ via a quartic term in the Higgs potential

$$\mathcal{L}_{quart}=\kappa[\Phi_1^T C(\sigma^\dagger)^2\Phi_2]+h.c. \tag{147}$$

which fixes the PQ properties of σ relative to those of the Higgs fields and fermions [22]. In the limit that $v_{PQ}>>v$, the axion content of the three Higgs fields is easily seen to be

$$\begin{aligned}\sigma &= \frac{v_{PQ}}{\sqrt{2}}e^{i\frac{a}{v_{PQ}}}\\ \Phi_1 &= \frac{v_1}{\sqrt{2}}e^{i\frac{2v_2^2}{v^2v_{PQ}}a}\begin{bmatrix}1\\0\end{bmatrix}\equiv\frac{v_1}{\sqrt{2}}e^{i\frac{X_1}{v_{PQ}}a}\begin{bmatrix}1\\0\end{bmatrix}\\ \Phi_2 &= \frac{v_2}{\sqrt{2}}e^{i\frac{2v_1^2}{v^2v_{PQ}}a}\begin{bmatrix}0\\1\end{bmatrix}\equiv\frac{v_2}{\sqrt{2}}e^{i\frac{X_2}{v_{PQ}}a}\begin{bmatrix}0\\1\end{bmatrix}\end{aligned} \tag{148}$$

That is, the hypercharge zero excitation a of Eq (79) only contains a fraction $\alpha=\frac{2v_1v_2}{vv_{PQ}}$ of the invisible axion field, thereby guaranteeing that (147) is invariant under the PQ shift:

$$a\to a+\alpha v_{PQ} \tag{149}$$

[21]One can, of course, also contemplate models where the quarks and leptons are given variant PQ assignments

[22]One can consider instead of (147) a trilinear coupling of Φ_1,Φ_2 and σ. This model [79] has the nice feature that the invisible axion is essentially a Majoron [80] and there is a relation between the PQ breaking scale and neutrino masses.

The parameters X_1 and X_2, defined in (148), play identical roles to x and $\frac{1}{x}$, respectively, in the standard axion model [31], while v_{PQ} replaces v. Hence, the relevant properties of the $DFSZ$ axion are immediate to detail. Since $X_1 + X_2 = 2$, one has

$$m_a^{DFSZ} = \frac{m_\pi F_\pi}{v_{PQ}} \frac{\sqrt{m_u m_d}}{(m_u + m_d)} [2N_g] \equiv \frac{m_\pi F_\pi}{\bar{v}_{PQ}} \frac{\sqrt{m_u m_d}}{(m_u + m_d)} \tag{150}$$

In the above I have defined a modified PQ-breaking scale $\bar{v}_{PQ} \equiv \frac{v_{PQ}}{2N_G}$ for the DFSZ axion, so as to make the mass formula for both the KSVZ and the DFSZ axions be the same. As we shall see soon, $\bar{v}_{PQ}$ turns out to be, anyway, the more relevant parameter to consider. For the mixing parameters and the $\gamma\gamma$ coupling one finds

$$\xi_{a\pi}^{DFSZ} = \frac{1}{2}[\frac{X_1 - X_2}{2N_g} - \frac{(m_d - m_u)}{(m_d + m_u)}](\frac{F_\pi}{\bar{v}_{PQ}}) \tag{151}$$

$$\xi_{a\eta}^{DFSZ} = [\frac{1 - N_g}{2N_g}](\frac{F_\pi}{\bar{v}_{PQ}}) \tag{152}$$

$$C_{a\gamma\gamma}^{DFSZ} = \{\frac{4}{3} - \frac{(4m_d + m_u)}{3(m_d + m_u)}\} \simeq 0.37 \tag{153}$$

In the above we have made use of Eq (108). Note also that the $\gamma\gamma$ coupling has the analogous meaning to that used for the KSFZ axion. That is, it is the coefficient of $\frac{a}{\bar{v}_{PQ}} \frac{e^2}{16\pi^2} F_{em}^{\mu\nu} \tilde{F}_{\mu\nu}^{em}$ in the Lagrangian. The $DFSZ$ axion has a direct derivative coupling to the electrons [cf Eq (86)]. Using the electron's equation of motion this yields the effective term

$$\mathcal{L}_{ae^+e^-}^{DFSZ} = -iX_2[\frac{m_e}{v_{PQ}}]\bar{e}\gamma_5 ea = -i\frac{X_2}{2N_g}[\frac{m_e}{\bar{v}_{PQ}}]\bar{e}\gamma_5 ea \tag{154}$$

4.3.3 Astrophysical and Cosmological Constraints of Invisible Axion Models

I have already noted earlier that invisible axion models are subjected to astrophysical and cosmological constraints. The original astrophysical bounds of [50] have been considerably sharpened in recent years by a number of detailed studies[81]. Similarly, also the cosmological bounds [51] on invisible axions have been reexamined and some of the uncertainties inherent in deriving these bounds have been better understood. I want to enter briefly into these matters here since:

- There has been a trend in recent times for the invisible axion window to begin to close (Although theoretical considerations which might help to open this window have also been suggested! [82])
- A number of experiments to detect invisible axions have been proposed [83] (and some have actually been carried out [84]), whose feasibility clearly relies on the range where invisible axions might exist

The astrophysical bounds one obtains for invisible axions depend somewhat on the type of axion one is considering. More precisely, the limits arising from stellar evolution are sensitive to the axion-electron coupling and hence are different for $KSVZ$ and $DFSZ$ axions. However, limits on axion emission from neutron star cores [85] or the very recent limits obtained from the observation of a neutrino pulse from the Supernova 1987a [86],

do not make a sharp distinction between $KSVZ$ and $DFSZ$ axions, since they depend on the mixing parameters $\xi_{a\eta}$ and $\xi_{a\pi}$, which are comparable for both models.

The strongest limits on axions, derived from considerations of the energy loss in stars due to axion emission, make use mostly of plasma processes involving electrons (Compton axion production $[\gamma + e \to a + e]$ and axion bremsstrahlung $[e + Z \to e + Z + a]$). The axion emission rate for the Primakoff process $[\gamma + e, Z \to e, Z + a]$ is substantially reduced, due to Debye Hückel screening of the electric charges [87]. Clearly, therefore, limits obtained for hadronic axions, where only the axion coupling to two photons is effective, are considerably weaker than those for $DFSZ$ axions.

Raffelt [87] derived a bound on $KSVZ$ axions by requiring that the axion emission rate, due to the Primakoff process, in the sun (the axion luminosity) does not exceed the surface photon luminosity. In terms of the PQ breaking scale this bound reads

$$v_{PQ} \geq |C_{a\gamma\gamma}^{KSVZ}|\ 9.7 \times 10^5\ GeV \tag{155}$$

Raffelt's bound has been confirmed by a more accurate and extensive analysis [88], in which the effect of axions has been incorporated directly into the stellar evolution code. Furthermore, by considering the effect of hadronic axions on the lifetime of helium burning red giants, and asking that there should not be a severe conflict between the calculated and observed lifetimes, Dearborn and Raffelt [88] were able to improve the bound on hadronic axions by more than an order of magnitude, obtaining

$$v_{PQ} \geq |C_{a\gamma\gamma}^{KSVZ}|\ 2.3 \times 10^7\ GeV \tag{156}$$

Since

$$m_a^{KSZV} = 6.2 \times 10^6\ [\frac{1}{v_{PQ}(GeV)}]\ eV, \tag{157}$$

the Dearborn-Raffelt bound implies

$$m_a^{KSZV} \leq \frac{0.27}{|3q_X^2 - 0.96|}\ eV \tag{158}$$

where we have used Eq (108) again, in the above.

For $DFSZ$ axions the strongest bound, arising from stellar evolution considerations, comes from the suppression of helium ignition in low mass red giants, due to the presence of axions [89]. If the axion electron coupling is too strong, Compton emission cools the helium in the star and insulates the core from conductive heating. The star then evolves up the giant branch, but its evolution time is significantly affected. Asking that there should not be more than a factor of two reduction in the evolution time for helium burning giants, due to the presence of axions, or, what is equivalent, that the axion energy loss should not exceed the stellar energy generation rate, requires [23]

$$g_{aee} \leq 1.4 \times 10^{-13} \tag{159}$$

Comparable bounds arise also from studying axion emission from white dwarfs [90] and from the bremsstrahlung of axions off electrons in the crust of neutron stars [85].

[23] This is the bound given in the first reference in [89]

Turning Eq(159) into a lower bound on $\tilde{v}_{PQ}$, or an upper bound on the axion mass, again involves an unknown parameter. Since [cf. Eq(154)] for the $DFSZ$ axion

$$g_{aee} = \frac{X_2}{2N_g}[\frac{m_e}{\tilde{v}_{PQ}}] = \frac{2v_1^2}{2N_g(v_1^2+v_2^2)}[\frac{m_e}{\tilde{v}_{PQ}}], \tag{160}$$

the bound (159) translates, for the case of three generations, into

$$\tilde{v}_{PQ} \geq [\frac{2v_1^2}{v_1^2+v_2^2}]\ 6\times 10^8\ GeV \tag{161}$$

or, equivalently,

$$m_a^{DFSZ} \leq [\frac{v_1^2+v_2^2}{2v_1^2}]\ 10^{-2}\ eV \tag{162}$$

One sees that, apart from model dependent constants, typically the lower bound on the PQ breaking scale for the $DFSZ$ axion is about a factor of thirty above that for the $KSVZ$ axion.

The observation of a neutrino pulse from the Supernova 1987a [91], of roughly the expected characteristics [92], provides further (stronger) constraints on both kinds of invisible axions [86]. Axion emission during the core collapse would substantially affect the observed neutrino pulse, if the axion luminosity was of the order of the neutrino luminosity ($\sim 10^{53}\ ergs\ sec^{-1}$). Obviously, if v_{PQ} is sufficiently large one can keep the axion luminosity below $10^{53}\ ergs\ sec^{-1}$. However, this also can occur if v_{PQ} is sufficiently small. In this latter case, although the collapsing star produces many more axions, these axions never escape since the absorption cross section for axions also grows and eventually the axions are thermalized in the hot core. Thus from the Supernova 1987a one can deduce a "forbidden range" for v_{PQ}.

The dominant process for axion emission in the formation of the proto neutron star is axion bremsstrahlung from nucleon-nucleon interactions. There is some confusion in the literature regarding the magnitude of the cross section for this process [93], which is connected with whether one shold use a pseudoscalar or a pseudovector axion coupling to nucleons. As our preceding discussion should have made clear, the correct procedure is to compute the nucleon nucleon bremsstrahlung of pions and η's in a chiral Lagrangian and then mix the axion in, via the parameters $\xi_{a\pi}$ and $\xi_{a\eta}$. Unfortunately, as far as I can see, this is not precisely what has been done to date [24]. At any rate, I will quote below bounds which make use of Iwamoto's bremsstrahlung calculation [85] (corrected by some appropriate statistical factors [86]), since these bounds are the more conservative ones.

The restrictions from SN 1987a which one obtains are essentially bounds on the strength of the axion coupling to nucleons, which is connected to the parameters $\xi_{a\pi}$ and $\xi_{a\eta}$. Thus the Supernova bounds are valid both for $KSVZ$ and $DFSZ$ axions. In what follows I will use, for definitiveness, the analysis of Turner [86], since he also discusses what happens in the case of strong coupling in some detail. Turner gives a bound $[g \lesssim 1.4\times 10^{-10}]$ on the axion-nucleon coupling squared $g^2 \equiv 3(g_{ann}^2 + g_{app}^2)$, which translates to the constraint [25]:

$$[\xi_{a\pi}^2 + \frac{9}{25}\xi_{a\eta}^2] \leq 1.7\times 10^{-23} \tag{163}$$

[24] For example, in a chiral Lagrangian there are $\pi\pi\bar{N}N$ contact interactions [94] and these appear not to have been included in the calculations of [93]

[25] Note that $g^2 = 6g_{\pi NN}^2[\xi_{a\pi}^2 + \frac{9}{25}\xi_{a\eta}^2]$

where the factor of $\frac{9}{25}$ is an estimate of the ratio of the isoscalar to the isovector nucleon axial form factors. For the $KSVZ$ model this implies, using Eqs (144) and (145):

$$v_{PQ} \geq 7.3\times 10^9\ GeV \tag{164}$$
$$m_a^{KSVZ} \leq 8.4\times 10^{-4}\ eV \tag{165}$$

while for the DSVZ model, using Eqs (151) and (152), one obtains, for 3 generations:

$$\tilde{v}_{PQ} \geq [1.44 + \frac{(X_1 - X_2 - 1.56)^2}{2}]\ 3.7\times 10^9\ GeV \tag{166}$$
$$m_a^{DFSZ} \leq \frac{1.7\times 10^{-3}}{[1.44 + \frac{X_1 - X_2 - 1.56)^2}{2}]}\ eV \tag{167}$$

The above limits are considerably stronger than those arising from stellar evolution. However, as I have indicated above, for small enough v_{PQ} the axions interact so strongly that they are trapped in the collapsing star and cannot contribute to the cooling which would distort the neutrino pulse. Trapping occurs when the axion mean free path is less than or of the size of the core. According to Turner's calculations [86], this obtains if $g \geq 3\times 10^{-9}$. However, even in the trapped regime, the thermalized axions will radiate away energy. Only when the temperature of this "axiosphere" is low enough ($T_a < 8 - 10\ MeV$) will the axion luminosity be less than $10^{53} ergs\ sec^{-1}$. This constrains $g \geq 4\times 10^{-7}$, which corresponds to the limit

$$\xi_{a\pi}^2 + \frac{9}{25}\xi_{a\eta}^2 \geq 1.3\times 10^{-16} \tag{168}$$

For the $DFSZ$ axion the corresponding v_{PQ} is in a region $[\tilde{v}_{PQ} \leq [1.44 + \frac{(X_1 - X_2 - 1.56)^2}{2}]\ 1.3\times 10^6\ GeV]$ which is already well excluded by the stellar bound (161). However, for the $KSVZ$ axion the upper bound on v_{PQ} implied by (168) is

$$v_{PQ} \lesssim 2.6\times 10^6\ GeV \quad (KSVZ) \tag{169}$$

which is not very far removed from the lower bound (156) [26]. Thus it is possible that a small (Turner) window near $v_{PQ} \simeq 10^6 - 10^7\ GeV$ $[m_a \simeq 1\ eV]$ remains open for hadronic axions.

The existence of a window near $m_a \simeq 1\ eV$, where hadronic axions may be allowed by astrophysical considerations is of particular importance because, precisely for this range of values, one can conceive of an experiment for their detection. Axions originating from the sun, with typical KeV energies, can be converted in the presence of a strong magnetic field into X-rays, as a result of the nontrivial axion-photon-photon coupling. The original suggestion of a possible axion helioscope for detecting this sort of invisible axions is due to Sikivie [83]. The most advanced proposal I know of, to realize this in practice, is due to Van Bibber et al. [95], who want to make use of the magnet from the FNAL 15 foot bubble chamber for this purpose. The flux of hadronic axions from the sun has been calculated by Raffelt [96] to be

$$F_a = [\frac{C_{a\gamma\gamma}^{KSVZ}}{v_{PQ}(GeV)}]^2\ 1.9\times 10^{26}\ cm^{-2}sec^{-1}$$
$$= 3.6\times 10^{11}\ cm^{-2}sec^{-1} \tag{170}$$

[26] One can imagine models where $C_{a\gamma\gamma}^{KSVZ}$ is of $0(10^{-1})$. Furthermore, the theoretical uncertainties in estimating the upper bound (169) could easily approach an order of magnitude

where the second line is the value corresponding to the Dearborn-Raffelt bound (156). The maximum probability of axion to photon conversion in a homogenous magnetic field B, of linear extent L, can be simply shown to be [83]

$$P(a \to \gamma) = \left|\frac{\alpha BL}{2\pi}\frac{C^{KSVZ}_{a\gamma\gamma}}{v_{PQ}}\right|^2 \quad (171)$$

which, using again (156), for $B \simeq 3\ Tesla$ and $L \simeq 4m$, gives counting rates of the order of $10^{-2} sec^{-1}$. Van Bibber et al. [95] argue that this rate may be actually realizeable in practice if one tunes the density of a low Z gas in the detector, so that the effective mass of the produced X-ray photon matches that of the impinging axion. Obviously, results from such a proposed experiment (even if null!) would be of considerable interest.

If the Turner window is nonexistent, the present astrophysical bounds on invisible axions appear to force the PQ breaking scale near $10^{10} GeV$ [cf. Eqs (164) and (166)]. This value is sufficiently close to the usually quoted cosmological upper bound [51], $v_{PQ} \leq 10^{12} GeV$, that one should peruse this bound also more closely. It is relatively easy to understand the origin of the cosmological bound and to derive an order of magnitude estimate for the limiting value for v_{PQ}. It is, however, considerably harder to try to really pin down a, hard and fast, number for v_{PQ}.

The cosmological bound on v_{PQ} requires paying attention to the evolution of the axion expectation value, through the history of the Universe. This expectation value [which was set to zero in Eqs (141) and (148), as a matter of convention] is zero in the early Universe. However, the expectation value of the invisible axion field then assumes some arbitrary value, as the Universe cools down through the temperature range, $T \sim v_{PQ}$, of the PQ phase transition [27]. This value, in general, is not the value to which $< a >$ will be driven to dynamically by the chiral color anomaly [cf. Eq (63), with $\bar{\Theta} = 0$ in the convention adopted here.] Hence, as the Universe cools below temperatures of order $T \sim \Lambda_{QCD}$, where the effects of the color anomaly become significant, $< a >$ must begin to adjust and it will oscillate towards its final dynamical value [zero here]. The bound on v_{PQ} arises by requiring that the energy density stored in the axion field oscillations today should not be larger than the energy density which closes the Universe.

In the approximation that the only important part of the axion potential is its mass term, the equation of motion for $< a >$, in the expanding Universe, is

$$\frac{d^2 < a >}{dt^2} + 3H(t)\frac{d < a >}{dt} + m_a^2(t) < a > = 0 \quad (172)$$

where $H(t)$ is the Hubble constant: $H = \frac{d}{dt}\ln R$, with $R(t)$ being the Robertson Walker scale factor. The Hubble constant, in fact, varies with time and temperature as $H \sim \frac{1}{t} \sim T^2$. Well below temperatures of order Λ_{QCD}, $m_a(t)$ takes the value I previously detailed for the invisible axion models. Much above Λ_{QCD}, $m_a(t)$ vanishes. Clearly if $m_a(t) << H(t)$ the value of $< a >$ will not change, while if $m_a(t) >> H(t)$, $< a >$ will oscillate. This oscillation will be approximately sinusoidal if the time variation of $m_a(t)$ is not strong

[27] One can assume that $< a(x) >$ is essentially space independent for $T << v_{PQ}$, since the Fourier components of non zero k are rapidly redshifted away, once the corresponding wavelength has entered the horizon [51]

$[\frac{dm_a}{dt} << m_a^2]$. In this adiabatic regime, it is easy to check that an approximate solution of (172) is given by

$$< a(t) > \simeq \frac{A}{[m_a(t)R^3(t)]^{\frac{1}{2}}} \cos\, m_a(t)t \quad (173)$$

with A a constant. The energy density stored in the axion field is thus

$$\rho_a(t) = \frac{1}{2}[m_a^2 < a >^2 + (\frac{d < a >}{dt})^2] \simeq \frac{1}{2}\frac{m_a(t)A^2}{R^3(t)} \quad (174)$$

One can calculate with this formula the axion energy density today, in terms of the density at the time when $< a >$ began to oscillate. This instant is conventionally taken [52] as the time when

$$3H(t_{osc}) = m_a(t_{osc}) \quad (175)$$

and the axion energy density is

$$\rho_a(t_{osc}) = \frac{1}{2}m_a^2(t_{osc})(Cv_{PQ})^2 \quad (176)$$

with C some number of $0(1)$, related to the original axion vacuum alignment: $< a > = Cv_{PQ}$ [28]. It follows then immediately that

$$\begin{aligned} \rho_a &= \rho_a(t_{osc})[\frac{m_a}{m_a(t_{osc})}][\frac{R(t_{osc})}{R}]^3 \\ &= \frac{1}{2}C^2 v_{PQ}^2 m_a m_a(t_{osc})[\frac{R(t_{osc})}{R}]^3 \end{aligned} \quad (177)$$

The time t_{osc}, or equivalently the temperature T_{osc}, where (175) holds is closely connected to Λ_{QCD}. Thus one can immediately gain an order of magnitude estimate for ρ_a by replacing

$$m_a(t_{osc}) \simeq H(T_{QCD}) \simeq \frac{\Lambda_{QCD}^2}{M_P} \quad (178)$$

with M_P being the Planck mass, and using $m_a \sim \frac{\Lambda_{QCD}^2}{v_{PQ}}$. Then one secures

$$\rho_a \simeq v_{PQ}T^3(\frac{\Lambda_{QCD}}{M_P}) \quad (179)$$

where $T \simeq 3K$ is the present temperature. Demanding that $\rho_a \leq \rho_{crit} \simeq 10^{-46}\ GeV^4$ implies the rough bound

$$v_{PQ} \leq 10^{12} GeV \quad (180)$$

which, amazingly enough, is the usual quoted bound!

To do better one must:

[28] It is useful to use a convention for the PQ breaking scale so that the parameter ξ in Eq (65), multiplying the color anomaly, is unity. This obtains if one uses v_{PQ} for this scale for KSVZ axions and $\tilde{v}_{PQ}$ for DFSZ axions, respectively. In this case, the expectation value $< a >$, divided by the PQ scale, has always periodicity 2π. Thus, with the understanding that v_{PQ} in Eq (176) means either v_{PQ} or $\tilde{v}_{PQ}$, depending on the type of axion, C is really of $O(1)$

1. Calculate the time t_{osc} of Eq (175). This requires a knowledge of the temperature dependence of the axion mass.

2. Calculate the cosmic scale ratio $[\frac{R(t_{osc})}{R}]$. An uncertainty here is the number of active species at t_{osc}, since this number influences the evolution of the Hubble constant with temperature.

3. Try to estimate the constant C, which typifies the value of $< a >$ at the time of the PQ phase transition.

4. Try to estimate the errors made in dropping nonlinear terms in the axion potential and the range of validity of the adiabatic approximation, leading to Eq (173).

Most of the above points can be reasonable addressed, albeit with a certain degree of uncertainty [51] [97]. For instance, for the temperature dependence of the axion mass one can use the dilute instanton gas calculation of Gross Pisarski and Yaffe [98]. Knowing how m_a varies with T allows one to secure a value for T_{osc}, or equivalently t_{osc}. Then most of the uncertainty in the ratio $\frac{R(t_{osc})}{R}$ is related to whether or not there has been entropy increase since t_{osc}. Similarly, if one assumes that all values of $\frac{<a>}{v_{PQ}}$ are allowed in the relevant periodicity domain, then C should be given by the root mean square value $C = \frac{\pi}{\sqrt{3}}$ [97]. Furthermore, one can also incorporate the effect of non linearities in the axion potential and check the self consistency of the adiabatic approximation [97].

The most recent detailed analysis of these problems is due to Turner [97] and I quote his result, including (some) of the uncertainties,

$$\begin{aligned} \bar{v}_{PQ} &\leq (1.6-7.8)\times 10^{12}[\gamma]^{0.85}[C]^{-1.7}\ GeV \quad (DFSZ) \\ v_{PQ} &\leq (1.6-7.8)\times 10^{12}[\gamma]^{0.85}[C]^{-1.7}\ GeV \quad (KSVZ) \end{aligned} \tag{181}$$

Here $\gamma \geq 1$ would account for entropy production, after axion production. Note that the result is the same for both axion models, if one uses the appropiate PQ scale. If one assumes C is given by the root mean square value $C = \frac{\pi}{\sqrt{3}}$, the bound on v_{PQ} is close to $10^{12} GeV$. Thus, it appears that there is still some window for invisible axions. Although the astrophysical bounds have crept considerably up with time, the cosmological bound has, at least, not decreased [29]!

If axions indeed serve to close the Universe, it is again possible to try to search for their traces experimentally [83]. Axions in the galactic halo can be converted in an external magnetic field into photons. Since the relic axions move with the non relativistic virial velocity, the converted photons will have a very sharp frequency distribution centered around m_a. Thus one can try to see the signal of these relic axions by means of a variable frequency resonant cavity. A preliminary (but very challenging!) experiment of this sort has already been carried out [101], exploring the mass range 4.5 to $5\mu eV$. Although the experiment is about two orders of magnitude above the sensitivity level necessary to detect an axion photon conversion, the result obtained provides a non trivial direct upper bound

[29] R. Davis [99] argues that the presence of cosmic axion strings, present in a non inflationary scenario, may be an additional source of cosmological axions. He obtains, as a consequence, a bound on v_{PQ} which is about two order of magnitude lower than Eq (181). This calculation, however, has been questioned by Harari and Sikivie [100], who obtain only a lowering of the v_{PQ} bound by a factor of 2 or so.

on the axion-photon-photon coupling, for axions in this mass range. Assuming a relic density of axions near the earth of $< \rho_a >= 300 MeV/cm^3$ [102], the bound obtained in [101] for $C_{a\gamma\gamma}$ is

$$C_{a\gamma\gamma} \leq 1.6 \tag{182}$$

This is to be compared to Eqs (146) and (153).

Although a window for invisible axions exists, and experiments are valiantly probing the edges of this range, no one has any real idea of why the scale of the PQ breaking should lie in this range. It would be much more pleasing if one could remove v_{PQ} to, at least, the GUT scale, where perhaps one could relate it to some other physical phenomena [30]. Unfortunately, to substantially alter the upper bound (181) on the scale v_{PQ} requires that the treatment of the QCD phase transition be substantially altered. Some early work along these lines was done by Unruh and Wald [104], who suggested several mechanisms during the QCD phase transition to pump less energy into the axion field. This topic has reemerged anew recently [82], and it is claimed that with sufficient supercooling it might be possible to increase the v_{PQ} bound by a couple of orders of magnitude. That is something, but probably not enough to tie the PQ scale naturally to phenomena near the Planck scale.

These considerations notwithstanding, one should not despair. There may well be physical reasons for having an intermediate scale below the Planck mass. Furthermore, there are compelling astrophysical and cosmological reasons for wanting some dark matter in the Universe, and invisible axions are among the most "sensible" candidates for this dark component. Finally, scenarios for galaxy formation, based on cold dark matter (which could very well be axionic), remain of considerable interest. So invisible axions - the grandchilds of the strong CP problem we began discussing - appear to be well and to have acquired a life of their own. One should not, therefore, try to snuff them out, solely on pure theoretical prejudices.

5 Envoi

The strong CP problem is no closer to a solution today than it was twelve years ago, when it was first identified. This is frustrating, but perhaps not unexpected, since we also do not have a very good understanding of CP violation. If one had a CP conserving world, one would never worry about $\bar{\Theta}$. It would be just set to zero automatically, as a symmetry requirement. But CP is not conserved and we have to either resort to tricks (the soft CP solution) or to extra dynamical assumptions (the chiral solution) to get rid of $\bar{\Theta}$. Neither scheme really works in practice, but the latter approach opens up a plethora of interesting physical phenomena, connected with the appearance of axions.

Axions, if nothing else, have become a common theme that cuts horizontally across a broad spectrum of problems, in cosmology, astrophysics, particle physics, nuclear physics and atomic physics, involving both experimental and theoretical issues. However, the bottom line on axions must remain: are they there? This we do not know yet, but the hunting

[30] In any superymmetric theory which has a dilatational symmetry, the superpartner of the dilaton is always an axion. Hence, axions, and a concomitant PQ symmetry, emerge very naturally in superstring theories. Unfortunately, the natural scale for v_{PQ} in these theories is much closer to M_P than $10^{12} GeV$! [103]

grounds are becoming increasingly sparse. It is my belief that the searches for axions in the cosmologically and astrophysically allowed region should be pursued with vigor, even though the realistic chances of finding a signal are slim. At the same time, more theoretical effort should be invested in the strong CP problem itself. It is very well possible that there are other solutions to this problem, besides the ones that I have discussed, which are just waiting to be discovered. On this hopeful note, let me close.

Acknowledgements

The list of people who have influenced my thinking concerning the strong CP problem is long. I remain grateful to Steve Weinberg for having asked the right question and to Helen Quinn for a most fruitful and pleasant collaboration. Bill Donnelly helped me understand nuclei and Mel Schwarz and Stu Freedman how experiments really work. I owe the little I know of anomalies to Bill Bardeen, who in addition contributed crucially to my understanding of the low energy aspects of axions. I owe many insights into the collective aspects of axions to Tsotumo Yanagida and to Frank Wilczek. Georg Raffelt, Pierre Sikivie and Mike Turner were my guide to the heavens. Jihn Kim, on more than one occasion, straightened me out with his encyclopeadic knowledge. To all go my heartfelt thanks. Last, but not least, I would like to thank Cecilia Jarlskog for asking me to write this chapter for her book and for her patience in waiting for it to materialize.

References

[1] S. Weinberg, Phys. Rev. D11 (1975) 3583

[2] G. t'Hooft, Phys. Rev. Lett. 37 (1976) 8; Phys. Rev. D14 (1976) 3432

[3] S.L. Adler, Phys. Rev. 177 (1969) 2426; J.S. Bell and R. Jackiw, Nuovo Cim. 60 (1969) 47; W.A. Bardeen, Phys.Rev. 184 (1969) 1848

[4] C.G. Callan, R. Dashen and D. Gross. Phys. Lett. 63B (1976) 334

[5] R. Jackiw and C. Rebbi, Phys. Rev. Lett. 37 (1976) 172

[6] For a nice pedagogical discussion, see for example R.J. Crewther, Acta Phys. Austriaca Suppl. 19 (1978) 47

[7] W.A. Bardeen, Nucl. Phys. B75 (1974) 246

[8] For a masterful exposition see S. Coleman in Proc 1977 Erice Summer School, **The Ways of Subnuclear Physics**, ed. A. Zichichi (Plenum press 1979)

[9] The $U(1)_A$ problem has been reviewed in depth recently by G.A. Christos, Phys. Rept. 116 (1984) 751. A more condensed discussion is contained in G. t'Hooft Phys. Rept. 142 (1986) 357, in which some of the more controversial points of Christos' review are addressed

[10] A.A. Belavin, A.M. Polyakov, A.S. Schwartz and Yu.S. Tyupkin, Phys. Lett. 59B (1975) 85

[11] M.F. Atiyah and I.M. Singer, Ann. Math. 87 (1968) 484, 546

[12] This point is discussed in detail by R. Crewther, Riv Nuovo Cimento 2 (1979) 63 and in the review of G. Christos, Phys. Rept. 116 (1984) 751

[13] R. Dashen, Phys. Rev.D3 (1971) 1879

[14] P. Di Vecchia and G. Veneziano, Nucl. Phys. B171 (1980) 253

[15] R. Crewther, Riv Nuovo Cimento 2 (1979) 63

[16] This is where the controversy regarding the solution of the $U(1)_A$ problem is centered in the treatment of 't Hooft and that of Christos [9] - which follows the point of view of Crewther [17]

[17] R. Crewther, Phys. Lett. 70B (1977) 349

[18] E. Witten, Nucl. Phys. B156 (1979) 269

[19] I.S. Altarev et al, Phys. Lett. B136 (1984) 327

[20] V. Baluni, Phys. Rev. D19 (1979) 2227

[21] R. Crewther, P. Di Vecchia, G. Veneziano and E. Witten, Phys. Lett. 88B (1979) 123; 91B (1980) 487 (E)

[22] These bounds are reviewed by H.Y. Cheng, Phys. Rept. 158 (1988) 1

[23] M. Morgan and G. Miller, Phys. Lett. 179B (1986) 379

[24] E.P. Shabalin, Sov. J. Nucl. Phys. 36 (1982) 575

[25] J. Ellis and M.K. Gaillard, Nucl. Phys. B150 (1979) 141

[26] S. Khlebnikov and M. Shaposhnikov, Phys. Lett. 203B (1988) 121

[27] J. Gasser and H. Leutwyler, Phys. Repts. 87 (1982) 77

[28] G. t'Hooft, Nucl. Phys. B61 (1973) 455

[29] D.B. Kaplan and A.V. Manohar, Phys. Rev. Lett. 56 (1986) 2004

[30] K. Choi, C.W. Kim and W.K. Sze, John Hopkins preprint JHU-TIPAC 8804 March 1988

[31] R.D. Peccei and H.R. Quinn, Phys. Rev. Lett. 38 (1977) 1440; Phys. Rev. D16 (1977) 1791

[32] S. Weinberg, Phys. Rev. Lett. 40 (1978) 223; F. Wilczek, Phys. Rev. Lett. 40 (1978) 271

[33] T.D. Lee, Phys. Rev. D8 (1973) 1326

[34] P. Sikivie, Phys. Lett. 65B (1976) 41

[35] There is an extensive discussion of the soft CP alternative in both the recent reviews of J. Kim, Phys. Rept. 150 (1987) 1 and of H.Y. Cheng, Phys. Rept. 158 (1988) 1 of the strong CP problem

[36] H. Burkhardt et al, Phys. Lett. 206B (1988) 169

[37] Representative examples include: M.A.B. Beg and H.S. Tsao, Phys. Rev. Lett. 41 (1978) 278; H. Georgi, Hadronic Jour. 1 (1978) 155; R.N. Mohapatra and G. Senjanovic, Phys. Lett. 126B (1978) 283; G. Segré and A. Weldon, Phys. Rev. Lett. 42 (1979) 1191; S. Barr and P. Langacker, Phys. Rev. Lett. 42 (1979) 1654

[38] Y.B. Zeldovich, I.B. Kobzarev and L. Okun, Sov. Phys. JETP 40 (1975) 1

[39] S. Barr and A. Zee, Phys. Rev. Lett. 55 (1985) 2253

[40] A. Nelson, Phys. Lett. 136B (1983) 387

[41] S. Barr, Phys. Rev. D30 (1984) 1805

[42] A. Nelson, Phys. Lett. 143B (1984) 165

[43] S. Barr, Phys. Rev. D34 (1986) 1567

[44] A. Masiero, R.N. Mohapatra and R.D. Peccei, Nucl. Phys. B192 (1981) 66

[45] L.S. Hall and L.J. Randall, Nucl. Phys. B274 (1986) 157

[46] J.F. Nieves, Phys. Lett. 164B (1985) 85

[47] Similar bounds have been obtained also by W. Buchmüller and D. Wyler, Nucl. Phys. B268 (1986) 621

[48] See for example, W. Buchmüller, R. Rückl and D. Wyler, Phys. Lett. 191B (1987) 442

[49] For a discussion, see for example the review of H.Y. Cheng, Phys. Rept. 158 (1988) 1

[50] D. Dicus, E.W. Kolb, V.I. Teplitz and R.W. Wagoner, Phys. Rev. D18 (1978) 1829; K. Sato, Prog. Theor. Phys. 60 (1978) 1942; M. Vsyotsskii, Y. Zeldovich, M. Khlopov and V. Chechetkin, JETP Lett. 27 (1978) 502

[51] J. Preskill, M. Wise and F. Wilczek, Phys. Lett. 120B (1983) 127; L.F. Abbott and P. Sikivie, Phys. Lett. 120B (1983) 133; M. Dine and W. Fischler, Phys. Lett. 120B (1983) 137

[52] W.A. Bardeen and S.H.H. Tye, Phys. Lett 74B (1978) 580

[53] W.A. Bardeen, R.D. Peccei and T. Yanagida, Nucl. Phys. B279 (1987) 401

[54] R.D. Peccei, T.T. Wu and T. Yanagida, Phys. Lett. 172B (1986) 435; L.M. Krauss and F. Wilczek, Phys. Lett. 173B (1986) 189

[55] S. Weinberg, Phys. Rev. Lett. 40 (1978) 223; J. Kandaswamy, P. Salomonson and J. Schechter, Phys. Rev. D17 (1978) 3051; H. Georgi, D.B. Kaplan and L. Randall, Phys. Lett. 169B (1986) 73; W.A. Bardeen, R.D. Peccei and T. Yanagida, Nucl. Phys. B279 (1987) 401

[56] For a discussion of effective Lagrangians, see for example, H. Georgi. **Weak Interactions and Modern Particle Theory** (Benjamin 1984)

[57] In the review of Kim, Phys. Rept. 150 (1987) 1, the formula which details the $a\gamma\gamma$ coupling has a piece missing. It should read $c_{a\gamma\gamma} = \bar{c}_{a\gamma\gamma} - \frac{2}{3}\frac{4+Z}{1+Z}[\frac{x+x^{-1}}{2}]$, where the factor in the square bracket was omitted. I am grateful to J. Kim for making me aware of this correction.

[58] D.B. Kaplan, Nucl. Phys. B260 (1985) 215; M. Srednicki, Nucl. Phys. B260 (1985) 689; see also, H. Georgi, D.B. Kaplan and L. Randall Phys. Lett. 169B (1986) 73

[59] A.A. Anselm and N.G. Uraltsev, Phys. Lett. B114 (1982) 39

[60] This matter is discussed in more detail, for example, in the review of Kim, Phys. Rept. 150 (1987) 1. See also, R.D. Peccei in **Neutrino 81** R.J. Cence et al., ed. (Univ. of Hawai, 1981) Vol I [long lived axion bounds] and R.D. Peccei in Proceedings of the XXIII Conference of High Energy Physics, S. Loken ed. (World Scientific, 1987) [short lived axion bounds]

[61] Y. Asano et al., Phys. Lett. 107B (1981) 159

[62] T. Goldman and C.M. Hoffmann, Phys. Rev. Lett. 40 (1978) 220; S. Weinberg, Phys. Rev. Lett. 40 (1978) 223; J. Kandaswamy, P. Salomonson and J. Schechter, Phys. Lett. 74B (1978) 337; J.M. Frere, M.B. Gavela and J.A.M. Vermaseren, Phys. Lett. 103B (1981) 159; I. Antoniadis and T.N. Truong, Phys. Lett. 109B (1982) 67; J. Kim, Phys. Rept. 150 (1987) 1; W.A. Bardeen, R.D. Peccei and T. Yanagida, Nucl. Phys. B279 (1987) 401

[63] For a review of these bounds see A. Zehnder, Phys. Lett. 104B (1981) 494

[64] T.W. Donnelly, S.J. Freedman, R.S. Lytel, R.D. Peccei and M. Schwartz, Phys. Rev. D18 (1978) 1607

[65] L.M. Krauss and D. J. Nash, Phys. Lett. 202B (1988) 560

[66] F. Wilczek, Phys. Rev. Lett. 39 (1977) 1304

[67] P. Nason, Phys. Lett. 175B (1986) 223

[68] I use the value $\alpha_s = 0.13 \pm 0.02$ at $q^2 = (35\ GeV)^2$, obtained from a recent global analysis by A. Ali and F. Barreiro, Desy preprint DESY 88-075

[69] S. Yamada, Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energy, ed. by D. G. Cassel and D. L. Kreinick (Cornell 1983)

[70] S. Egli et al, Phys. Lett. 175B (1986) 97

[71] R. Eichler et al, Phys. Lett. 175B (1986) 101

[72] M. Suzuki, Phys. Lett. 175B (1986) 364; L.M. Krauss and M. Wise, Phys. Lett. 176B (1986) 483; W.A. Bardeen, R.D. Peccei and T. Yanagida, Nucl. Phys. B279 (1987) 401

[73] A.L. Hallin et al., Phys. Rev. Lett. 57 (1986) 2105; see also, F.W.N. de Boer, Phys. Lett. 180B (1986) 4

[74] J. Kim, Phys. Rev. Lett. 43 (1979) 103

[75] M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, Nucl. Phys. B166 (1980) 493

[76] M. Dine, W. Fischler and M. Srednicki, Phys. Lett. 104B (1981) 199

[77] A.P. Zhitnitskii, Sov. J. Nucl. Phys. 31 (1980) 260

[78] M. Srednicki, Nucl. Phys. B260 (1985) 689

[79] P. Langacker, R.D. Peccei and T. Yanagida, Mod. Phys. Lett. A1 (1986) 541

[80] Y. Chikashige, R.N. Mohapatra and R.D. Peccei, Phys. Lett. 98B (1981) 265

[81] M. Fukugita, S. Watamura and M. Yoshimura, Phys. Rev. Lett. 48 (1982) 1522; Phys. Rev. D26 (1982) 1840; D.S. Dearborn, D.N. Schramm and G. Steigman, Phys. Rev. Lett. 56 (1986) 26; G. Raffelt, Phys. Lett. 166B (1986) 402; Phys. Rev. D33 (1986) 897; G. Raffelt and D.S. Dearborn, Phys. Rev. D36 (1987) 2211; J.A. Frieman, S. Dimopoulos and M.S. Turner, Phys. Rev. D36 (1987) 2201

[82] M.B. Voloshin, ITEP preprint ITEP-88-044; N. Dowrick and N.A. McDougall, Oxford preprint 1988

[83] P. Sikivie, Phys. Rev. Lett. 51 (1983) 1415; 52 (1984) 695 (E); Phys. Rev. D32 (1985) 1988; L. Krauss, J. Moody, F. Wilczek and D. Morris, Phys. Rev. Lett. 55 (1985) 1797 ; L. Maiani, R. Petronzio and E. Zavattini, Phys. Lett. 175B (1986) 359; K. Van Bibber et al., Phys. Rev. Lett. 59 (1987) 759; K. Van Bibber et al., Letter of Intent to Fermilab, March 1988

[84] S. de Panfilis et al., Phys. Rev. Lett. 59 (1987) 839; F. Avignone et al., Phys. Rev. D35 (1987) 2752

[85] N. Iwamoto, Phys. Rev. Lett. 53 (1984) 1198

[86] G. Raffelt and D. Seckel, Phys. Rev. Lett. 60 (1988) 1793; M. Turner, Phys. Rev. Lett. 60 (1988) 1797; R. Mayle, J. Wilson, J. Ellis, K. Olive, D.N. Schramm and G. Steigman, Phys. Lett. 203B (1988) 203

[87] G. Raffelt, Phys. Rev. D33 (1986) 897

[88] G. Raffelt and D.S. Dearborn, Phys. Rev. D36 (1987) 2211. For a somewhat less technical discussion see, J. Frieman, S. Dimopoulos and M.S. Turner, Phys. Rev. D36 (1987) 2201

[89] D.S. Dearborn, D.N. Schramm and G. Steigman, Phys. Rev. Lett. 56 (1986) 26; J. Frieman, S. Dimopoulos and M.S. Turner, Phys. Rev. D36 (1987) 2201

[90] G. Raffelt, Phys. Lett. 166B (1986) 402

[91] K. Hirata et al., Phys. Rev. Lett. 58 (1987) 1490; R.M. Bionta et al., Phys. Rev. Lett. 58 (1987) 1494

[92] See for example, D.N. Schramm, Comments Nucl. Part. Phys. 17 (1987) 2391

[93] N. Iwamoto, Phys. Rev. Lett. 53 (1984) 1198; A. Pantziris and K. Kang, Phys. Rev. D33 (1986) 3509. See also [86]

[94] See for example, R.D. Peccei, Phys. Rev. 176 (1968) 1812

[95] K. Van Bibber et al., Letter of Intent to Fermilab, March 1988. I am grateful to G. Raffelt for communicating this document to me.

[96] G, Raffelt, Phys. Rev. D; to be published

[97] J. Kim, Phys. Rept. 150 (1987) 1; M.S. Turner, Phys. Rev. D33 (1986) 819

[98] D. Gross, R. Pisarski and L. Yaffe, Rev. Mod. Phys. 53 (1981) 43

[99] R. Davis, Phys. Lett. 180B (1986) 225

[100] D. Harari and P. Sikivie, Phys. Lett. 195B (1987) 361

[101] S. de Panfilis et al., Phys. Rev. Lett 59 (1987) 839

[102] M.S. Turner, Phys. Rev. D33 (1986)

[103] For a discussion of the axion scale problem in superstring theories, see for example, J. Kim, in Proceedings of the XXIII International Conference in Energy Physics, Berkeley 1986, ed. by S. Loken (World Scientific 1987)

[104] W.G. Unruh and R.M. Wald, Phys. Rev. D32 (1985) 831