Discrete and Global Symmetries in Particle Physics

R. D. Peccei

Department of Physics and Astronomy, UCLA, Los Angeles, CA 90095-1547

Abstract. I begin these lectures by examining the transformation properties of quantum fields under the discrete symmetries of Parity, P, Charge Conjugation, C, and Time Reversal, T. With these results in hand, I then show how the structure of the Standard Model helps explain the conservation/violation of these symmetries in various sectors of the theory. This discussion is also used to give a qualitative proof of the CPT Theorem, and some of the stringent tests of this theorem in the neutral Kaon sector are reviewed. In the second part of these lectures, global symmetries are examined. Here, after the distinction between Wigner-Weyl and Nambu-Goldstone realizations of these symmetries is explained, a discussion is given of the various, approximate or real, global symmetries of the Standard Model. Particular attention is paid to the role that chiral anomalies play in altering the classical symmetry patterns of the Standard Model. To understand the differences between anomaly effects in QCD and those in the electroweak theory, a discussion of the nature of the vacuum structure of gauge theories is presented. This naturally raises the issue of the strong CP problem, and I present a brief discussion of the chiral solution to this problem and of its ramifications for astrophysics and cosmology. I also touch briefly on possible constraints on, and prospects for, having real Nambu-Goldstone bosons in nature, concentrating specifically on the simplest example of Majorons. I end these lectures by discussing the compatibility of having global symmetry in the presence of gravitational interactions. Although these interactions, in general, produces small corrections, they can alter significantly the Nambu-Goldstone sector of theories.

1 Discrete Space-Time Symmetries

Lorentz transformations

$$x^\mu \to x'^\mu = \Lambda^\mu_{\ \nu} x^\nu \tag{1}$$

preserve the invariance of the space-time interval

$$x_\mu x^\mu = \mathbf{r}^2 - c^2 t^2 = \mathbf{r'}^2 - c^2 {t'}^2 = x'_\mu x'^\mu \ . \tag{2}$$

This constrains the matrices $\Lambda^\mu_{\ \nu}$ to obey

$$\eta_{\mu\nu} = \Lambda^\lambda_{\ \mu} \eta_{\lambda\kappa} \Lambda^\kappa_{\ \nu} \ , \tag{3}$$

where the matrix tensor $\eta_{\mu\nu}$ is the diagonal matrix

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \ . \tag{4}$$

The pseudo-orthogonality of the Λ matrices detailed in Eq. (3)

$$\eta = \Lambda^T \eta \Lambda \qquad (5)$$

allows the classification of Lorentz transformations depending on whether

$$\det \Lambda = \begin{cases} +1 \\ -1 \end{cases} ; \quad \Lambda^0_0 = \pm\sqrt{1 + \sum_{i=1}^{3}(\Lambda^i{}_0)^2} = \begin{cases} \geq +1 \\ \leq -1 \end{cases} . \qquad (6)$$

As a result, the Lorentz group splits into four distinct pieces

$$\begin{aligned}
L^\uparrow_+ &: \ \det \Lambda = +1; \ \Lambda^0_0 \geq 1 \\
L^\uparrow_- &: \ \det \Lambda = -1; \ \Lambda^0_0 \geq 1 \\
L^\downarrow_+ &: \ \det \Lambda = +1; \ \Lambda^0_0 \leq -1 \\
L^\downarrow_- &: \ \det \Lambda = -1; \ \Lambda^0_0 \leq -1 .
\end{aligned} \qquad (7)$$

The transformation matrices Λ in $L^\uparrow_+$ by themselves form a sub-group of the Lorentz group: the proper orthochronous Lorentz group. All other transformations in the Lorentz group can be obtained from Λ in $L^\uparrow_+$ by using two discrete transformations, P and T, characterized by the matrices:

$$P^\mu{}_\nu = \begin{bmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix} ; T^\mu{}_\nu = \begin{bmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{bmatrix} \qquad (8)$$

corresponding to space inversion (Parity) and time reversal. It is clear that if $\Lambda \in L^\uparrow_+$, then $P\Lambda \in L^\uparrow_-$; $PT\Lambda \in L^\downarrow_+$; and $T\Lambda \in L^\downarrow_-$. Remarkably, nature is invariant only under the proper orthochronous Lorentz transformations. Parity is violated in the weak interactions, something which was first suggested by Lee and Yang (Lee and Yang 1956) in 1956 and soon thereafter observed experimentally (Wu et al 1957). The detection of the decay of K^0_L into pions by Christenson, Cronin, Fitch and Turlay (Christenson et al 1964) in 1964 provided indirect evidence that also time reversal is not a good symmetry of nature.

One can understand why this is so on the basis of the Standard Model of electroweak and strong interactions and of the, so called, CPT theorem, established by Pauli, Schwinger, Lüders and Zumino (Pauli 1955). To appreciate these facts I will need to sketch how quantum fields behave under the discrete space-time transformations of P and T, as well as their behavior under charge conjugation (C) which physically corresponds to reversing the sign of all charges. I will begin with parity.

1.1   Parity

The Parity transformation properties of the electromagnetic fields follow directly from classical considerations.[1] The Lorentz force

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{9}$$

obviously changes sign under Parity, since $\mathbf{p} \to -\mathbf{p}$.[2] Hence, it follows that $\mathbf{E}$ is odd and $\mathbf{B}$ is even under Parity:

$$\mathbf{E}(\mathbf{x}, t) \xrightarrow{P} -\mathbf{E}(-\mathbf{x}, t) \; ; \;\; \mathbf{B}(\mathbf{x}, t) \xrightarrow{P} B(-\mathbf{x}, t) \; . \tag{10}$$

Formally, the transformation above is induced by a Unitary operator $U(P)$. This operator takes the vector potential $A^\mu(\mathbf{x}, t)$ into a transformed vector potential $A^\mu(-\mathbf{x}, t)$. In view of Eq. (10), it is easy to see that

$$U(P)A^\mu(\mathbf{x}, t)U(P)^{-1} = \eta(\mu)A^\mu(-\mathbf{x}, t) \; , \tag{11}$$

where the symbol $\eta(\mu)$ is a useful notational shorthand, with

$$\eta(\mu) = \begin{cases} -1 \; \mu \neq 0 \\ +1 \; \mu = 0 \; . \end{cases} \tag{12}$$

Spin-zero scalar, $S(\mathbf{x}, t)$, and pseudoscalar, $P(\mathbf{x}, t)$, fields under parity are, respectively, even and odd. That is,

$$\begin{aligned} U(P)S(\mathbf{x}, t)U(P)^{-1} &= S(-\mathbf{x}, t) \\ U(P)P(\mathbf{x}, t)U(P)^{-1} &= -P(-\mathbf{x}, t) \; . \end{aligned} \tag{13}$$

The behavior of spin-1/2 Dirac fields $\psi(\mathbf{x}, t)$ under Parity is slightly more complex. However, this behavior can be straightforwardly deduced from the requirement that the Dirac equation be invariant under this operation. One finds that

$$U(P)\psi(\mathbf{x}, t)U(P)^{-1} = \eta_P \gamma^0 \psi(-\mathbf{x}, t) \; . \tag{14}$$

Here η_P is a phase factor of unit magnitude ($|\eta_P|^2 = 1$). Because one is always interested in fermion-antifermion bilinears, the phase factor η_P plays no role physically and one can set it to unity ($\eta_P \equiv 1$) without loss of generality.

Given Eq. (14), it is a straightforward exercise to deduce the Parity properties of fermion-antifermion bilinears.[3] Since

$$\gamma^0 \gamma^0 \gamma^0 = \gamma^0 \; ; \gamma^0 \gamma^i \gamma^0 = -\gamma^i \; ; \gamma^0 \gamma_5 \gamma^0 = -\gamma_5 \; , \tag{15}$$

---

[1] Henceforth, I shall use natural units where $c = \hbar = 1$.

[2] Since Parity reverses the sign of space coordinates $\mathbf{r} \to -\mathbf{r}$, the velocity also changes sign, $\mathbf{v} \to -\mathbf{v}$.

[3] In my conventions $\{\gamma^\mu, \gamma^\nu\} = -2\eta^{\mu\nu}$, $\gamma^{0\dagger} = \gamma^0$ but $\gamma^{i\dagger} = -\gamma^i$, and $\gamma_5 = i\gamma^0 \gamma^1 \gamma^2 \gamma^3$.

one easily deduces that

$$U(P)\bar{\psi}(\mathbf{x},t)\psi(\mathbf{x},t)U(P)^{-1} = \bar{\psi}(-\mathbf{x},t)\psi(-\mathbf{x},t) \quad \text{(Scalar)}$$
$$U(P)\bar{\psi}(\mathbf{x},t)i\gamma_5\psi(\mathbf{x},t)U(P)^{-1} = -\bar{\psi}(-\mathbf{x},t)i\gamma_5\psi(-\mathbf{x},t) \quad \text{(Pseudoscalar)}$$
$$U(P)\bar{\psi}(\mathbf{x},t)\gamma^\mu\psi(\mathbf{x},t)U(P)^{-1} = \eta(\mu)\bar{\psi}(-\mathbf{x},t)\gamma^\mu\psi(-\mathbf{x},t) \quad \text{(Vector)}$$
$$U(P)\bar{\psi}(\mathbf{x},t)\gamma^\mu\gamma_5\psi(\mathbf{x},t)U(P)^{-1} = -\eta(\mu)\bar{\psi}(-\mathbf{x},t)\gamma^\mu\gamma_5\psi(-\mathbf{x},t) \quad \text{(Pseudovector)}$$

$$(16)$$

From the above, one sees immediately that the electromagnetic interaction is parity invariant:

$$W_{\text{int}}^{\text{em}} = \int d^4x\, e A^\mu(x)\bar{\psi}(x)\gamma_\mu\psi(x) \xrightarrow{P} W_{\text{int}}^{\text{em}} . \tag{17}$$

On the other hand, because Parity transforms fields of a given **chirality** into each other[4]

$$\psi_{\text{L}}(\mathbf{x},t) \xrightarrow{P} \gamma^0\psi_{\text{R}}(-\mathbf{x},t); \ \ \psi_{\text{R}}(\mathbf{x},t) \xrightarrow{P} \gamma^0\psi_{\text{L}}(-\mathbf{x},t) , \tag{18}$$

it is obvious that the chirally asymmetric weak interactions will violate parity. Thus, this sector of the Standard Model is Parity violating. The strong interactions, however, are invariant under Parity. These interactions are governed by Quantum Chromodynamics and in QCD both the left-handed and right-handed quarks are triplets under the $SU(3)$ gauge group:

$$q_{\text{L}} \sim 3 \ ; \ \ q_{\text{R}} \sim 3 . \tag{19}$$

Note the difference here with respect to the weak interactions. Under the weak $SU(2)$ group of the $SU(2)\times U(1)$ theory, the left-handed fields ψ_{L} of both quarks and leptons are doublets, while the right-handed fields ψ_{R} are singlets

$$\psi_{\text{L}} \sim 2 \ ; \ \ \psi_{\text{R}} \sim 1 . \tag{20}$$

This is the root cause for the violation of Parity in the weak interactions.

1.2 Charge Conjugation

As I alluded to earlier, the process of charge conjugation is connected physically with the reversal of the sign of all electric charges. For the electromagnetic field, therefore, the charge conjugation transformation C brings the vector potential $A^\mu(x)$ into minus itself

$$U(C)A^\mu(x)U(C)^{-1} = -A^\mu(x) . \tag{21}$$

For Dirac fields, since charge conjugation should transform particles into antiparticles, this operation essentially corresponds to Hermitian conjugation. That is, one has

$$U(C)\psi(x)U(C)^{-1} = \eta_c C\psi^\dagger(x) . \tag{22}$$

---

[4] Here $\psi_{\text{L}}(x) = \frac{1}{2}(1-\gamma_5)\psi(x); \ \psi_{\text{R}}(x) = \frac{1}{2}(1+\gamma_5)\psi_{\text{R}}(x).$

Here η_c is again a phase factor of unit magnitude and, without loss of generality, one can take $\eta_c \equiv 1$. The form of the matrix C can be deduced from the requirement that the transformation (22) should leave the Dirac equation invariant. For this to be the case necessitates that

$$C\gamma_\mu^* C^{-1} = -\gamma_\mu \ . \tag{23}$$

The particular form of C one obtains depends on the form of the γ-matrices used. In the Majorana representation, where the γ-matrices are purely imaginary [Majorana: $\gamma_\mu^* = -\gamma_\mu$] then $C = 1$. On the other hand, in the Dirac representation [Dirac: $\gamma^0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$; $\gamma^i = \begin{bmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{bmatrix}$], then $C = \gamma_2$. Because of the simplicity of C in the Majorana representation, in what follows we shall make use of this representation when dealing with charge conjugation.

Using Eq. (22), it is straightforward to compute the C-conjugation properties of fermion antifermion bilinears. Let me do this explicitly for the scalar density $\bar{\psi}\psi$ and then quote the results for the other bilinears. One has

$$\begin{aligned}
U(C)\bar{\psi}(x)\psi(x)U(C)^{-1} &= U(C)\psi_\alpha^\dagger(x)(\gamma^0)_{\alpha\beta}\psi_\beta(x)U(C)^{-1} \\
&= \psi_\alpha(x)(\gamma^0)_{\alpha\beta}\psi_\beta^\dagger(x) \\
&= -\psi_\beta^\dagger(x)(\gamma^0)_{\alpha\beta}\psi_\alpha(x) \\
&= -\psi_\beta^\dagger(x)(\gamma^{0T})_{\beta\alpha}\psi_\alpha(x) \\
&= +\bar{\psi}(x)\psi(x) \ .
\end{aligned} \tag{24}$$

The second line above is the result of using Eq. (22), taking $C = 1$ assuming one is working in the Majorana representation. The third line above follows because fermion fields anticommute (apart from an irrelevant infinite piece which can be subtracted away). Finally, the last line follows since in the Majorana representation γ^0 is an antisymmetric matrix ($\gamma^{0T} = -\gamma^0$).

The full set of results for the behavior of fermion-antifermion bilinears under C is displayed below:

$$\begin{aligned}
U(C) \ \bar{\psi}(x)\psi(x)U(C)^{-1} &= \bar{\psi}(x)\psi(x) \quad \text{(Scalar)} \\
U(C) \ \bar{\psi}(x)i\gamma_5\psi(x)U(C)^{-1} &= \bar{\psi}(x)i\gamma_5\psi(x) \quad \text{(Pseudoscalar)} \\
U(C) \ \bar{\psi}(x)\gamma^\mu\psi(x)U(C)^{-1} &= -\bar{\psi}(x)\gamma^\mu\psi(x) \quad \text{(Vector)} \\
U(C) \ \bar{\psi}(x)\gamma^\mu\gamma_5\psi(x)U(C)^{-1} &= \bar{\psi}(x)\gamma^\mu\gamma^5\psi(x) \quad \text{(Pseudovector)} \ .
\end{aligned} \tag{25}$$

These results lead to some immediate consequences. For instance, it follows that electromagnetic interactions are C-invariant. Using Eqs. (21) and (25) it follows that

$$W_{\text{int}}^{\text{em}} = \int d^4x e A^\mu(x)\bar{\psi}(x)\gamma_\mu\psi(x) \xrightarrow{C} W_{\text{int}}^{\text{em}} \ , \tag{26}$$

since both A^μ and the electromagnetic current $\bar{\psi}\gamma^\mu\psi$ change sign under C.

The strong interactions are also invariant under charge conjugation. This takes a small discussion, but it is also easy to see. The principal point to note is that the $SU(3)$ currents of QCD do not have the same simple transformation properties as the electromagnetic current, because they involve the non-trivial $SU(3)$ matrices λ_a. Effectively these matrices get transposed in the bilinears, if one makes a charge conjugation transformation. That is, one has

$$U(C)\bar{q}\gamma^\mu \frac{\lambda_a}{2} q U(C)^{-1} = -\bar{q}\gamma^\mu \left(\frac{\lambda_a}{2}\right)^T q \ . \tag{27}$$

Because $\lambda_1, \lambda_3, \lambda_4, \lambda_6$, and λ_8 are symmetric, while λ_2, λ_5, and λ_7 are antisymmetric, it follows that

$$J_a^\mu \rightarrow -\eta(a) J_a^\mu \ , \tag{28}$$

where

$$\eta(a) = \begin{cases} +1 \text{ for } a = 1, 3, 4, 6 \text{ and } 8 \\ -1 \text{ for } a = 2, 5 \text{ and } 7 \end{cases} \tag{29}$$

To guarantee invariance of the quark gluon interaction terms

$$W_{\text{int}} = \int d^4 x g_3 A_a^\mu J_{\mu a} \tag{30}$$

under charge conjugation it is necessary to assume that the charge conjugation properties of the gluon fields themselves vary according to which component one is dealing with. Namely, for invariance of Eq. (30) under C one needs

$$U(C)A_a^\mu(x)U(C)^{-1} = -\eta(a)A_a^\mu(x) \ . \tag{31}$$

It is easy to check that the above transformation property is precisely what is needed to have the nonlinear gluon field strengths have well defined C-properties. Recall that

$$G_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial^\nu A_a^\mu + g f_{abc} A_b^\mu A_c^\nu \ . \tag{32}$$

Now, for $SU(3)$, the only non-vanishing structure constants f_{abc} are (Slansky 1981)

$$f_{abc} \neq 0 \text{ for } abc = \{123, 147, 156, 246, 257, 345, 367, 458, 678\} \ . \tag{33}$$

One sees that $f_{abc} \neq 0$ only for cases in which there is an odd number of indices which themselves are odd (i.e. the indices: 2,5, and 7). This assures that, indeed, $G_a^{\mu\nu}$ transforms in the same way as A_a^μ does under C:

$$U(C)G_a^{\mu\nu}(x)U(C)^{-1} = -\eta(a)G_a^{\mu\nu}(x) \ . \tag{34}$$

This last property then insures that

$$W^{\text{QCD}} = \int d^4 x \left[-\bar{q}\left(\gamma^\mu \frac{1}{i}D_\mu + m_q\right)q - \frac{1}{4}G_a^{\mu\nu}G_{a\mu\nu}\right] \xrightarrow{C} W^{\text{QCD}} \ . \tag{35}$$

The situation is different for the weak interactions since these involve both vector and pseudovector interactions. Let us focus, for example, on the $SU(2)$ current for leptons of the first generation

$$J_i^\mu = (\bar{\nu}_e\ \bar{e})_{\rm L} \gamma^\mu \frac{\tau_i}{2} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_{\rm L} = \frac{1}{4}(\bar{\nu}_e\ \bar{e})\gamma^\mu(1-\gamma_5)\tau_i \begin{pmatrix} \nu_e \\ e \end{pmatrix} \ . \tag{36}$$

This current transforms differently in its vector and pseudovector pieces as well as in its 1, 3 and 2 components:

$$U(C)J_{1,3}^\mu U(C)^{-1} = -\frac{1}{4}(\bar{\nu}_e\ \bar{e})\gamma^\mu(1+\gamma_5)\tau_{1,3} \begin{pmatrix} \nu_e \\ e \end{pmatrix}$$

$$U(C)J_2^\mu U(C)^{-1} = +\frac{1}{4}(\bar{\nu}_e\ \bar{e})\gamma^\mu(1+\gamma_5)\tau_2 \begin{pmatrix} \nu_e \\ e \end{pmatrix} \ . \tag{37}$$

The difference in behavior in the 1,3 and 2 components is absorbed by postulating the following C-transformation properties for the W_i^μ fields.[5]

$$U(C)W_i^\mu(x)U(C)^{-1} = -\eta(i)W_i^\mu(x) \ , \tag{38}$$

with

$$\eta(i) = \begin{cases} +1\ i = 1, 3 \\ -1\ i = 2 \end{cases} \tag{39}$$

Note that these properties are what one might expect since they imply that

$$W_\pm^\mu = \frac{1}{\sqrt{2}}(W_1^\mu \mp iW_2^\mu) \xrightarrow{C} -W_\mp^\mu \ . \tag{40}$$

However, even so, the simultaneous presence of vector and pseudovector pieces in the currents which enter the weak interactions force one to conclude that

$$W_{\text{weak interactions}} \xrightarrow{C} W_{\text{weak interactions}} \ , \tag{41}$$

as is observed experimentally.

1.3   Time Reversal

Classically, T-invariance corresponds to the fact that the equations of motion describing a particle going from A to B along some path also allow, as a permitted motion, the time reversed motion. That is, a motion where the particle follows the same path, but is now going from B to A. Clearly, in this time reversed motion all momenta are reflected, but the coordinates remain the same. So, classically, under a T-transformation

$$\mathbf{p} \xrightarrow{T} -\mathbf{p} \ ; \quad \mathbf{F} = \frac{d\mathbf{p}}{dt} \xrightarrow{T} \mathbf{F} \ . \tag{42}$$

---

[5] These transformation properties guarantee that $F_i^{\mu\nu}$ and W_i^μ have the same C-properties.

Quantum mechanically, the interchange of initial and final states is implemented by having the operator $U(T)$, corresponding to time reversal, be an **anti-unitary** operator (Wigner 1932), with

$$U(T) = V(T)K .\qquad (43)$$

In the above, $V(T)$ is a unitary operator while K complex conjugates any c-number quantity it acts on. The operation of complex conjugation as part of $U(T)$ is what renders this operator anti-unitary. The need for complex conjugation, in connection with time reversal, is already seen at the level of the Schrödinger equation. From

$$i\frac{\partial}{\partial t}\psi(\mathbf{x}, t) = H\psi(\mathbf{x}, t)\qquad (44)$$

one deduces that $\psi^*(\mathbf{x}, -t)$ obeys the equation

$$i\frac{\partial}{\partial t}\psi^*(\mathbf{x}, -t) = H^*\psi^*(\mathbf{x}, -t) .\qquad (45)$$

So, provided that the Hamiltonian is real ($H^* = H$), then one sees that $\psi^*(\mathbf{x}, -t)$ is also a solution of the Schrödinger equation. Therefore, in quantum mechanics, complex conjugation of the wave function (along with the reality of the Hamiltonian) accompanies the reversal in the direction of time.

The association of complex conjugation with time reversal effectively interchanges incoming and outgoing states (Low 1967)

$$\langle U(T)\phi | U(T)\psi \rangle = \langle \psi | \phi \rangle .\qquad (46)$$

Thus, if T is a good symmetry of the theory, one relates processes to their time reversed process (e.g. the decay $A \to BC$ to the formation of A from the coalescence of B and C, $BC \to A$). More precisely, if time reversal is a good symmetry, then one relates the S-matrix element S_{fi} to that for $S_{\tilde{i}\tilde{f}}$, where the states, $\tilde{i}, \tilde{f}$ have all the momentum directions $\{\mathbf{p}\}$ reversed in comparison to the states i, f. That is

$$S_{fi} = {}_{\text{out}}\langle f | i \rangle_{\text{in}} = {}_{\text{in}}\langle U(T)i | U(T)f \rangle_{\text{out}} = {}_{\text{out}}\langle \tilde{i} | \tilde{f} \rangle_{\text{in}} = S_{\tilde{i}\tilde{f}} .\qquad (47)$$

The next to last step above is only valid if time reversal is a good symmetry of the theory, since in this case it follows that

$$U(T)|f\rangle_{\text{out}} = |\tilde{f}\rangle_{\text{in}} ; \quad U(T)|i\rangle_{\text{in}} = |\tilde{i}\rangle_{\text{out}} .\qquad (48)$$

I should add a comment here about the issue of the reality of the Hamiltonian needed for time reversal to hold at the Schrödinger equation level. This is not quite the case when spin is involved and is the reason for the possible additional operator $V(T)$ in the definition of $U(T)$ in Eq. (43). More correctly, in general, what is needed is that

$$V(T)H^*V(T)^{-1} = H .\qquad (49)$$

When there is no spin $V(T)$ is just the unit matrix, but with spin its presence allows for T-invariance. The simplest example of this is provided by the ordinary spin-orbit interaction of atomic physics

$$H_{\text{s-o}} = \lambda \boldsymbol{\sigma} \cdot \mathbf{L} \ , \tag{50}$$

with λ some real constant. Since $\mathbf{L} = \mathbf{r} \times \frac{1}{i}\boldsymbol{\nabla}$, it follows that

$$H_{\text{s-o}}^* = \lambda \boldsymbol{\sigma}^* \cdot \mathbf{L}^* = -\lambda \boldsymbol{\sigma}^* \cdot \mathbf{L} \ , \tag{51}$$

which is not the same as Eq. (50) because $\sigma_2^* = -\sigma_2$ but $\sigma_{1,3}^* = \sigma_{1,3}$. However, since $\sigma_2 \boldsymbol{\sigma}^* \sigma_2 = -\boldsymbol{\sigma}$, using $V(T) = \sigma_2$ guarantees that

$$V(T)H_{\text{s-o}}^* V(T)^{-1} = H_{\text{s-o}} \ , \tag{52}$$

reflecting physically that, indeed, time reversal not only changes $\mathbf{L} \to -\mathbf{L}$, but also, effectively, $\boldsymbol{\sigma} \to -\boldsymbol{\sigma}$.

In field theory, it is again straightforward to deduce what is the effect of a time-reversal transformation on the electromagnetic fields by focusing on what happens classically. Since the Lorentz force is invariant under T

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \xrightarrow{T} \mathbf{F} \ , \tag{53}$$

it follows that $\mathbf{E}$ is even and $\mathbf{B}$ is odd under time-reversal. In terms of the vector potential, therefore, one has

$$U(T)A^\mu(\mathbf{x}, t)U(T)^{-1} = \eta(\mu)A^\mu(\mathbf{x}, -t) \ . \tag{54}$$

For spin-1/2 fields one can deduce the transformation properties of $\psi(\mathbf{x}, t)$ under T-transformations by again asking that the action of $U(T)$ on $\psi(\mathbf{x}, t)$ produce another solution of the Dirac equation. Writing

$$U(T)\psi(\mathbf{x}, t)U(T)^{-1} = \eta_T T \psi(\mathbf{x}, -t) \ , \tag{55}$$

with η_T a phase of unit magnitude (which we shall take, without loss of generality, to be unity, $\eta_T \equiv 1$), and remembering that $U(T)$ complex conjugates all c-numbers, one finds that for invariance of the Dirac equation the matrix T must obey

$$\begin{aligned} T\gamma^{0*}T^{-1} &= \gamma^0 \\ T\gamma^{i*}T^{-1} &= -\gamma^i \ . \end{aligned} \tag{56}$$

As was the case for the charge conjugation matrix C, the form of the matrix T also depends on which representation of the γ-matrices one uses. In the convenient Majorana representation, where $\gamma^{\mu*} = -\gamma^\mu$, one finds that

$$T = \gamma^0 \gamma_5 \ . \tag{57}$$

Armed with Eqs. (55) and (57), a simple calculation then produces the following transformation properties for the familiar fermion-antifermion bilinears:[6]

$$U(T)\,\bar{\psi}(\mathbf{x},t)\psi(\mathbf{x},t)U(T)^{-1} = \bar{\psi}(\mathbf{x},-t)\psi(\mathbf{x},-t) \quad \text{(Scalar)}$$

$$U(T)\,\bar{\psi}(\mathbf{x},t)i\gamma_5\psi(\mathbf{x},t)U(T)^{-1} = -\bar{\psi}(\mathbf{x},-t)i\gamma_5\psi(\mathbf{x},-t) \quad \text{(Pseudoscalar)}$$

$$U(T)\,\bar{\psi}(\mathbf{x},t)\gamma^\mu\psi(\mathbf{x},t)U(T)^{-1} = \eta(\mu)\bar{\psi}(\mathbf{x},-t)\gamma^\mu\psi(\mathbf{x},-t) \quad \text{(Vector)}$$

$$U(T)\,\bar{\psi}(\mathbf{x},t)\gamma^\mu\gamma_5\psi(\mathbf{x},t)U(T)^{-1} = \eta(\mu)\bar{\psi}(\mathbf{x},-t)\gamma^\mu\gamma_5\psi(\mathbf{x},-t) \quad \text{(Pseudovector)}$$

$$\tag{58}$$

It is obvious from the above and Eq. (54), as well from the reality of the electromagnetic coupling constant e, that the electromagnetic interactions are T-invariant

$$W_{\text{int}}^{\text{em}} = \int d^4x\, e A^\mu(x)\bar{\psi}(x)\gamma_\mu\psi(x) \xrightarrow{T} W_{\text{int}}^{\text{em}} . \tag{59}$$

It is easy to check also that the gauge interactions in both QCD and the $SU(2) \times U(1)$ electroweak theory are also T-invariant, provided one properly defines how the gauge fields transform. Since for $SU(3)$ only λ_2, λ_5 and λ_7 are imaginary, and for $SU(2)$ only σ_2 is imaginary, it is easy to check that the desired T-transformation properties are:[7]

$$U(T)A_a^\mu(\mathbf{x},t)U(T)^{-1} = \eta(\mu)\eta(a)A_a^\mu(\mathbf{x},-t) \quad (SU(3))$$

$$U(T)W_i^\mu(\mathbf{x},t)U(T)^{-1} = \eta(\mu)\eta(i)W_i^\mu(\mathbf{x},-t) \quad (SU(2))$$

$$U(T)Y^\mu(\mathbf{x},t)U(T)^{-1} = \eta(\mu)Y^\mu(\mathbf{x},-t) \quad (U(1)) . \tag{60}$$

Note that in contrast to C, T-transformations affect vector and pseudovector currents in the same way. Thus, using (58) and (60), it follows immediately that

$$W_{\text{gauge interactions}}^{\text{SM}} \xrightarrow{T} W_{\text{gauge interactions}}^{\text{SM}} . \tag{61}$$

The Standard Model can have, however, T-violating interactions in the electroweak sector involving the scalar Higgs field. The couplings of the Higgs field, in contrast to the gauge couplings, do not need to be real. These complex couplings then provide the possibility of having T-violating interactions. I examine this point in the simplest case where one has only one complex Higgs doublet

$$\Phi = \begin{pmatrix} \phi^0 \\ \phi^- \end{pmatrix} \tag{62}$$

in the theory. The scalar Higgs self-interactions, which trigger the breakdown of $SU(2) \times U(1)$, only involve real coefficients since one must require the Higgs potential to be Hermitian. That is

$$V = \lambda \left(\Phi^\dagger \Phi - \frac{v^2}{2} \right)^2 = V^\dagger \tag{63}$$

---

[6] In deducing Eq. (58), care must be taken to remember that $U(T)$ complex conjugates c-numbers.

[7] Of course, the gauge coupling constants, just like e, are real.

implies that both λ and v are real parameters.

The Yukawa interactions of Φ with the quark fields, however, can have complex coefficients.[8] With i, j being family indices, one can write, in general, these interactions as

$$\mathcal{L}_{\text{Yukawa}} = -\Gamma_{ij}^{u}(\bar{u}, \bar{d})_{\text{L}i}\Phi u_{\text{R}j} - \Gamma_{ij}^{d}(\bar{u}, \bar{d})_{\text{L}i}\tilde{\Phi}d_{\text{R}j} + \text{h.c.} \ . \tag{64}$$

Here $\tilde{\Phi} = i\sigma_2\Phi^*$ and the coefficient matrices Γ_{ij}^{u} Γ_{ij}^{d} are arbitrary complex matrices. After the electroweak interactions are spontaneously broken ($SU(2) \times U(1) \to U(1)_{\text{em}}$), effectively all that remains of the doublet field Φ is one scalar excitation—the Higgs boson H—and the vacuum expectation value v:

$$\Phi \to \frac{1}{\sqrt{2}} \begin{pmatrix} v + H \\ 0 \end{pmatrix} \tag{65}$$

Thus the Yukawa interactions (64) generate mass terms for the charge 2/3 and charge -1/3 quarks

$$M_{ij}^{u,d} = \frac{1}{\sqrt{2}}\Gamma_{ij}^{u,d}v \ . \tag{66}$$

As is well known, these mass matrices can be diagonalized by a bi-unitary transformation

$$(U_{\text{L}}^{u,d})^\dagger M^{u,d}U_{\text{R}}^{u,d} = \mathcal{M}^{u,d} \ . \tag{67}$$

The diagonal matrices $\mathcal{M}^{u,d}$ have real eigenvalues m_i, corresponding to the physical quark masses. Further, the bi-unitary transformations on the quark fields diagonalizes the Yukawa coupling matrices, since M and Γ are linearly related. Whence, all that remains of the Yukawa sector after these transformations is the simple interaction

$$\mathcal{L}_{\text{Yukawa}}^{\text{eff}} = - \sum_i m_i \bar{q}_i(x)q_i(x) \left[1 + \frac{H(x)}{v}\right] \ . \tag{68}$$

Provided $H(\mathbf{x}, t)$ has the canonical T-transformation one expects for a scalar field,

$$U(T)H(\mathbf{x}, t)U(T)^{-1} = H(\mathbf{x}, -t) \ . \tag{69}$$

Eq. (68) is a T-conserving interaction also. Nevertheless, the complex nature of the original Yukawa couplings does end up by producing some T-violating interactions.

It is easy to understand this last point. The bi-unitary transformations performed on the quarks to diagonalize the quark mass matrices alter the form of the charged current weak interactions. Before these transformations, these interactions had the form

$$\mathcal{L}^{\text{cc}} = \frac{e}{2\sqrt{2}\ \sin\theta_W} \left[W_+^\mu J_{-\mu}^0 + W_-^\mu J_{+\mu}^0\right] \ , \tag{70}$$

---

[8] I concentrate here only on the quark sector, because if one does not introduce right-handed neutrinos in the theory—so that neutrinos are effectively massless—then all the phases in the Yukawa couplings in the lepton sector can be rotated away.

with

$$J^0_{-\mu} = (\bar{u}_1, \bar{u}_2, \bar{u}_3)\gamma_\mu(1 - \gamma_5)\,\mathbf{1}\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \tag{71}$$

and

$$J^0_{+\mu} = (J^0_{-\mu})^\dagger \; . \tag{72}$$

Clearly, this interaction is T-invariant. However, after the bi-unitary transformation on the quark fields to diagonalize M [Eq. (67)], the charged current $J^0_{-\mu}$ is altered to

$$J_{-\mu} = (\bar{u}, \bar{c}, \bar{t})\gamma_\mu(1 - \gamma_5)\mathbf{V}_{\mathrm{CKM}}\begin{pmatrix} d \\ s \\ b \end{pmatrix} \; , \tag{73}$$

where the Cabibbo-Kobayashi-Maskawa quark mixing matrix (Cabibbo 1963, and Kobayashi and Maskawa 1973)

$$\mathbf{V}_{\mathrm{CKM}} = U^{u\dagger}_{\mathrm{L}} U^d_{\mathrm{L}} \tag{74}$$

is a unitary matrix, since U^u_{L} and U^d_{L} are. Because, in general, $\mathbf{V}_{\mathrm{CKM}}$ is complex, its presence in the currents J^μ_- (and J^μ_+) can lead to T-violation.

For three families of quarks and leptons, as we apparently have, it is not difficult to show that the matrix $\mathbf{V}_{\mathrm{CKM}}$ has only one physical phase, δ. All the other phases can be rotated away through further harmless redefinitions of the quark fields. If $\delta \neq 0$, then the charged current weak interactions are not T-invariant

$$\mathcal{L}^{\mathrm{cc}}(\mathbf{x}, t) = \frac{e}{2\sqrt{2}\,\sin\theta_{\mathrm{W}}}\left[W^\mu_+ J_{-\mu} + W^\mu_- J_{+\mu} \right] \xrightarrow{T}\!\!\!\!\!/\;\; \mathcal{L}^{\mathrm{cc}}(\mathbf{x}, -t) \tag{75}$$

and the standard model can give rise to observable manifestations of T-violation. We return to this point in more detail in the next subsection, after we discuss the CPT theorem.

1.4   The CPT Theorem

If nature is described by a local Lorentz invariant field theory, where there is the usual connection between spin and statistics, then one can prove a deep theorem, now known as the CPT Theorem (Pauli 1955). Namely, in these circumstances, one can show that the action of the theory is **always invariant** under the combined application of a C-, a P-, and a T-transformation. That is

$$W \xrightarrow{CPT} W \; . \tag{76}$$

I will not attempt here to establish the CPT theorem with rigor. The interested reader can turn, for example, to the erudite manuscript of Streater and Wightman (Streater and Wightman 1964) for this. Rather, I want to show why and

how the CPT Theorem works, based on the preceding discussion of the C, P, and T transformation properties of quantum fields.

To get started, let us look at the effect of a CPT transformation on the electromagnetic interactions. Using Eqs. (11), (16), (21), (25), (54), and (58), one has

$$
\begin{aligned}
A^\mu(\mathbf{x},t) &\xrightarrow{CPT} [-1][\eta(\mu)][\eta(\mu)]A^\mu(-\mathbf{x},-t) = -A^\mu(-\mathbf{x},-t) \\
J^\mu_{\text{em}}(x,t) &= \bar\psi(\mathbf{x},t)\gamma^\mu\psi(\mathbf{x},t) \\
&\xrightarrow{CPT} [-1][\eta(\mu)][\eta(\mu)]J^\mu_{\text{em}}(-\mathbf{x},-t) = -J^\mu_{\text{em}}(-\mathbf{x},-t) \ .
\end{aligned}
\tag{77}
$$

Obviously, therefore, under a CPT transformation

$$
W^{\text{em}}_{\text{int}} = \int d^4x\, eA^\mu(x)J^{\text{em}}_\mu(x) \xrightarrow{CPT} W^{\text{em}}_{\text{int}} \ .
\tag{78}
$$

This, however, is a trivial case, since $W^{\text{em}}_{\text{int}}$ was **separately** invariant under C-, P-, and T-transformations!

CPT invariance, if it is a general property, must hold also when there is violation of the individual symmetries. A more significant test is provided by the electroweak theory. There, for example, both C and P are violated in the neutral current interactions, while T and CPT are conserved. Let us check this. The action for the neutral current interactions is given by

$$
W^{\text{NC}}_{\text{int}} = \frac{e}{2\cos\theta_W\sin\theta_W}\int d^4x\, Z_\mu J^\mu_{\text{NC}} \ .
\tag{79}
$$

The neutral current

$$
J^\mu_{\text{NC}} = 2[J^\mu_3 - \sin^2\theta_W J^\mu_{\text{em}}] = V^\mu + A^\mu
\tag{80}
$$

contains both vector and pseudovector pieces, since these latter components are present in the $SU(2)$ current J^μ_3. Parity and Charge Conjugation are violated in Eq. (79) because the vector and pseudovector currents transform in opposite ways under each of these transformations. That is, one has, under Parity

$$
\begin{aligned}
Z^\mu(\mathbf{x},t) &\xrightarrow{P} \eta(\mu)Z^\mu(-\mathbf{x},t) \ ; \quad V^\mu(\mathbf{x},t) \xrightarrow{P} \eta(\mu)V^\mu(-\mathbf{x},t); \\
A^\mu(\mathbf{x},t) &\xrightarrow{P} -\eta(\mu)A^\mu(-\mathbf{x},t)
\end{aligned}
\tag{81}
$$

while, under Charge Conjugation,

$$
Z^\mu(\mathbf{x},t) \xrightarrow{C} -Z^\mu(\mathbf{x},t); \ V^\mu(\mathbf{x},t) \xrightarrow{C} -V^\mu(\mathbf{x},t); A^\mu(\mathbf{x},t) \xrightarrow{C} A^\mu(\mathbf{x},t)
\tag{82}
$$

On the other hand, T is conserved by Eq. (79), since under time reversal

$$
\begin{aligned}
Z^\mu(\mathbf{x},t) &\xrightarrow{T} \eta(\mu)Z^\mu(\mathbf{x},-t); \ V^\mu(\mathbf{x},t) \xrightarrow{T} \eta(\mu)V^\mu(\mathbf{x},-t); \\
A^\mu(\mathbf{x},t) &\xrightarrow{T} \eta(\mu)\mathbf{A}(\mathbf{x},-t) \ .
\end{aligned}
\tag{83}
$$

Using the above three equations, it is easy to see that the neutral current interactions conserve CPT. One has

$$Z^\mu(x,t) \xrightarrow{CPT} -Z^\mu(-\mathbf{x},-t) \ ; \ V^\mu(\mathbf{x},t) \xrightarrow{CPT} -V^\mu(-\mathbf{x},-t);$$
$$A^\mu(\mathbf{x},t) \xrightarrow{CPT} -A^\mu(-\mathbf{x},-t) \ . \tag{84}$$

From the above, it is also clear that CP and T are **equivalent** transformations for the neutral current action

$$W_{\text{int}}^{\text{NC}} \xrightarrow{CP} W_{\text{int}}^{\text{NC}} \xrightarrow{T} W_{\text{int}}^{\text{NC}} \ . \tag{85}$$

The equivalence between a T-transformation and a CP-transformation also holds when both of these potential symmetries are violated. Hence, even in this case, the combined CPT-transformation is indeed an invariance of the action. This is the essence of the CPT Theorem. To appreciate this point let me examine, specifically, the T-violating charged current interaction between the u and b quarks, typified by the complex CKM matrix element V_{ub}.[9] One has

$$W_{ub}^{\text{cc}} = \frac{e}{2\sqrt{2} \ \sin\theta_W} \int d^4x \left\{ V_{ub} W_+^\mu \bar{u}\gamma_\mu(1-\gamma_5)b + V_{ub}^* W_-^\mu \bar{b}\gamma_\mu(1-\gamma_5)u \right\} \ , \tag{86}$$

where

$$W_\pm^\mu = \frac{1}{\sqrt{2}}(W_1^\mu \mp iW_2^\mu) \ . \tag{87}$$

Because under T

$$W_1^\mu(\mathbf{x},t) \xrightarrow{T} \eta(\mu)W_1^\mu(\mathbf{x},-t); \ W_2^\mu(\mathbf{x},t) \xrightarrow{T} -\eta(\mu)W_2^\mu(\mathbf{x},-t) \tag{88}$$

and remembering the i factor in Eq. (87), it follows that

$$W_\pm^\mu(\mathbf{x},t) \xrightarrow{T} \eta(\mu)W_\pm^\mu(\mathbf{x},-t) \ . \tag{89}$$

On the other hand, under T, the $u-b$ currents behave as

$$\bar{u}(\mathbf{x},t)\gamma_\mu(1-\gamma_5)b(\mathbf{x},t) \xrightarrow{T} \eta(\mu)\bar{u}(\mathbf{x},-t)\gamma_\mu(1-\gamma_5)b(\mathbf{x},-t)$$
$$\bar{b}(\mathbf{x},t)\gamma_\mu(1-\gamma_5)u(\mathbf{x},t) \xrightarrow{T} \eta(\mu)\bar{b}(\mathbf{x},-t)\gamma_\mu(1-\gamma_5)u(\mathbf{x},-t) \ . \tag{90}$$

Hence, one sees, indeed, that the action W_{ub}^{cc} is not T-invariant

$$W_{ub}^{\text{cc}} \xrightarrow{T} \tilde{W}_{ub}^{\text{cc}} = \frac{e}{2\sqrt{2}\sin\theta_W} \int d^4x \left\{ V_{ub}^* W_+^\mu \bar{u}\gamma_\mu(1-\gamma_5)b + V_{ub} W_-^\mu \bar{b}\gamma_\mu(1-\gamma_5)u \right\} .$$
$$\tag{91}$$

---

[9] One can pick phase conventions where V_{ub} is real. In this case, however, other pieces in the charged current Lagrangian give rise to T-violation. The final result for physically measured parameters must be phase-convention independent. I focus here on the V_{ub} term for definitiveness, since in the standard convention for the CKM matrix (Cabibbo 1963, and Kobayashi and Maskawa 1973) V_{ub} is complex and its phase is precisely $-\delta$.

The behavior of the various ingredients in W^{cc}_{ub} under CP is individually different than it is under T. For instance, one has

$$W^\mu_1(\mathbf{x}, t) \xrightarrow{CP} -\eta(\mu)W^\mu_1(-\mathbf{x}, t); \ W^\mu_2(\mathbf{x}, t) \xrightarrow{CP} \eta(\mu)W^\mu_2(-\mathbf{x}, t) \ . \qquad (92)$$

Hence, since one also does not complex conjugate the i in $W^\mu_\pm$ in this case, one has

$$W^\mu_\pm(\mathbf{x}, t) \xrightarrow{CP} -\eta(\mu)W_\mp(-\mathbf{x}, t) \ . \qquad (93)$$

Similarly, one finds, that under CP the $u - b$ currents transform as

$$\bar{u}(\mathbf{x}, t)\gamma_\mu(1 - \gamma_5)b(\mathbf{x}, t) \xrightarrow{CP} -\eta(\mu)\bar{b}(-\mathbf{x}, t)\gamma_\mu(1 - \gamma_5)u(-\mathbf{x}, t)$$

$$\bar{b}(\mathbf{x}, t)\gamma_\mu(1 - \gamma_5)u(\mathbf{x}, t) \xrightarrow{CP} -\eta(\mu)\bar{u}(-\mathbf{x}, t)\gamma_\mu(1 - \gamma_5)b(-\mathbf{x}, t) \ . \qquad (94)$$

The net effect, however, on W^{cc}_{ub} is the same as that of a T-transformation. One finds

$$W^{\text{cc}}_{ub} \xrightarrow{CP} \tilde{W}^{\text{cc}}_{ub} = \frac{e}{2\sqrt{2}\,\sin\theta_W}\int d^4x \left\{V_{ub}W^\mu_-\bar{b}\gamma_\mu(1 - \gamma_5)u + V^*_{ub}W^\mu_+\bar{u}\gamma_\mu(1 - \gamma_5)b\right\} \ . \qquad (95)$$

One can extract from this example the underlying reason why the CPT theorem holds. It results really from a combination of the needed Hermiticity of the Lagrangian and the complementary role that T and CP play on the operators and c-numbers that enter in the Lagrangian. Hermiticity means that a given term in the Lagrangian, containing some operator $O(x)$ and some c-number a, has the form

$$\mathcal{L}(x) = aO(x) + a^*O^\dagger(x) \ . \qquad (96)$$

Under T, the operator is unchanged (except for replacing t by $-t$), but the c-number is complex conjugated

$$O(\mathbf{x}, t) \xrightarrow{T} O(\mathbf{x}, -t) \ ; \ a \xrightarrow{T} a^* \ . \qquad (97)$$

Under CP, on the other hand, the operator O gets essentially replaced by its Hermitian adjoint, but the c-number a stays the same:

$$O(\mathbf{x}, t) \xrightarrow{CP} O^\dagger(-\mathbf{x}, t) \ ; \ a \xrightarrow{CP} a \ . \qquad (98)$$

Combining the operations of T and CP changes, effectively, the first term in Eq. (96) into the second term and *vice versa*

$$\mathcal{L} = aO(x) + a^*O^\dagger(x) \xrightarrow{CPT} \mathcal{L}(-x) = a^*O^\dagger(-x) + aO(-x) \qquad (99)$$

leaving the action invariant

$$W = \int d^4x\mathcal{L}(x) \xrightarrow{CPT} W \ . \qquad (100)$$

1.5 CP and CPT Tests in the Neutral Kaon Complex

The $K^0 \sim \bar{s}d$ and $\bar{K}^0 \sim s\bar{d}$ states provide an excellent laboratory to test CP and CPT. These states are unstable, decaying into particles with no strangeness through a first-order weak process. In addition, second order weak processes, giving rise to the transition $\bar{s}d \leftrightarrow s\bar{d}$, allow the K^0 to mix with the $\bar{K}^0$. The quantum mechanical evolution of this two-state system leads to the physical eigenstates K_L^0 and K_S^0, characterized by their, respective, long and short life-times.

The physical eigenstates K_L^0 and K_S^0 are obtained by diagonalizing the 2×2 effective Hamiltonian

$$H_{\text{eff}} = M - \frac{i}{2}\Gamma . \tag{101}$$

Here M and Γ are Hermitian matrices describing the mass mixing and decay properties of the neutral Kaon complex. If CPT is a good symmetry of nature, then the diagonal matrix elements of M and Γ are equal, since this symmetry changes effectively K^0 into $\bar{K}^0$.

$$M_{11} = M_{22} ; \quad \Gamma_{11} = \Gamma_{22} \ \text{[CPT Conservation]} . \tag{102}$$

CP conservation, on the other hand, guarantees the reality of the mass and decay matrices. It provides therefore a constraint on the off-diagonal matrix elements of M and Γ. Namely:

$$M_{12} = M_{12}^* ; \quad \Gamma_{12} = \Gamma_{12}^* \ \text{[CP Conservation]} . \tag{103}$$

If one does not impose the above constraints of CPT and CP conservation on M and Γ, the eigenstates of the Schrödinger equation

$$H_{\text{eff}} \begin{pmatrix} |K^0\rangle \\ |\bar{K}^0\rangle \end{pmatrix} = i\frac{\partial}{\partial t} \begin{pmatrix} |K^0\rangle \\ |\bar{K}^0\rangle \end{pmatrix} \tag{104}$$

are linear superpositions of the $|K^0\rangle$ and $|\bar{K}^0\rangle$ states, involving parameters δ_K and ϵ_K which reflect the breaking of these symmetries. The physical $|K_L^0\rangle$ and $|K_S^0\rangle$ eigenstates have the standard time evolution

$$|K_{L,S}(t)\rangle = \exp[-im_{L,S}t] \exp\left[-\frac{1}{2}\Gamma_{L,S}t\right] |K_{L,S}(0)\rangle , \tag{105}$$

characterized by the mass and width of these particles. The states $|K_{L,S}(0)\rangle$ involve the following superposition of the $|K^0\rangle$ and $|\bar{K}^0\rangle$ states:

$$|K_L(0)\rangle = \frac{1}{\sqrt{2}} \left\{ (1 + \epsilon_K + \delta_K)|K^0\rangle + (1 - \epsilon_K - \delta_K)|\bar{K}^0\rangle \right\}$$

$$|K_S(0)\rangle = \frac{1}{\sqrt{2}} \left\{ (1 + \epsilon_K - \delta_K)|K^0\rangle - (1 - \epsilon_K + \delta_K)|\bar{K}^0\rangle \right\} . \tag{106}$$

In the above

$$\epsilon_K = e^{i\phi_{SW}} \left[\frac{-\text{Im } M_{12} + \frac{i}{2} \text{ Im } \Gamma_{12}}{\sqrt{2}\ \Delta m} \right]$$

$$\delta_K = ie^{i\phi_{SW}} \left[\frac{(M_{11} - M_{22}) - \frac{i}{2}(\Gamma_{11} - \Gamma_{22})}{2\sqrt{2}\ \Delta m} \right] \ , \tag{107}$$

where

$$\phi_{SW} = \tan^{-1} \frac{2\Delta m}{\Gamma_S - \Gamma_L} \ ; \quad \Delta m = m_L - m_S \ . \tag{108}$$

Experimentally, one finds (Particle Data Group 1996)

$$\phi_{SW} = (43.49 \pm 0.08)^o \ ; \quad \Delta m = (3.491 \pm 0.009) \times 10^{-12} \text{ MeV} \ . \tag{109}$$

Note that $\epsilon_K = 0$, if CP is conserved and $\delta_K = 0$, if CPT is conserved. Only if both ϵ_K and δ_K vanish are the eigenstates $|K_L^0\rangle$ and $|K_S^0\rangle$ CP eigenstates. If both these symmetries hold then

$$CP|K_{L,S}^0\rangle = \mp|K_{L,S}^0\rangle \quad [\text{CP, CPT Conservation}] \ . \tag{110}$$

What is measured experimentally are the CP violating ratios of the amplitude of the K_L and K_S to go into two pions

$$\eta_{+-} = \frac{A(K_L \to \pi^+\pi^-)}{A(K_S \to \pi^+\pi^-)} = |\eta_{+-}|e^{i\phi_{+-}} = \epsilon + \epsilon'$$

$$\eta_{00} = \frac{A(K_L \to \pi^0\pi^0)}{A(K_S \to \pi^0\pi^0)} = |\eta_{00}|e^{i\phi_{00}} = \epsilon - 2\epsilon' \ . \tag{111}$$

Experimentally, one finds that $\eta_{+-} \simeq \eta_{00}$ (so $\epsilon \gg \epsilon'$), with (Particle Data Group 1996)

$$|\eta_{+-}| = (2.285 \pm 0.019) \times 10^{-3} \ ; \quad \phi_{+-} = (43.7 \pm 0.6)^o \ . \tag{112}$$

Neglecting the contribution of the widths compared to the masses, which is a very good approximation, one finds that the parameter ϵ above is simply (Buchanan et al 1992)

$$\epsilon \simeq \epsilon_K - \delta_K \simeq e^{i\phi_{SW}} \left[\frac{-\text{Im } M_{12}}{\sqrt{2}\ \Delta m} \right] + ie^{i\phi_{SW}} \left[\frac{M_{22} - M_{11}}{2\sqrt{2}\ \Delta m} \right] \ . \tag{113}$$

Note that the CPT violating contribution in the above is 90° **out of phase** from the CP violating contribution. Because $\phi_{SW} = (43.49 \pm 0.08)^o$ is consistent with $\phi_{+-} = (43.7 \pm 0.6)^o$, one deduces immediately that the non-zero value for η_{+-} observed is mostly a signal of CP-violation [Im $M_{12} \neq 0$] rather than of CPT violation [$M_{11} \neq M_{22}$].

If one neglects altogether the possibility that there is any CPT violation in the neutral Kaon decay amplitudes—something one would eventually need to check—then one can write approximately

$$M_{22} - M_{11} \simeq |\eta_{+-}|2\sqrt{2}\ \Delta m \tan(\phi_{+-} - \phi_{SW}) \ . \tag{114}$$

This equation, given the values of the experimental parameters involved, provides a spectacularly strong bound on CPT violation, because the $K_L - K_S$ mass difference Δm is so small. One finds, at the 90% CL,

$$\left| \frac{m_{\bar{K}^0} - m_{K^0}}{m_{K^0}} \right| < 9 \times 10^{-19} \;, \tag{115}$$

which is an incredibly stringent test of CPT.

Experiments at the just completed Frascati Phi Factory will be able to directly measure δ_K, without further assumptions, to an accuracy similar to the present accuracy for ϵ. This will be accomplished by studying the difference in relative time decay patterns of the doubly semileptonic decays of the $K_L K_S$ states produced in the Φ decay. If one studies the relative time dependence of the process $\Phi \to K_L K_S \to \pi^- e^+ \nu_e(t_1) \pi^+ e^- \bar{\nu}_e(t_2)$, then one can show that the pattern at large $\Delta t = t_1 - t_2$ is sensitive to Re δ_K, while the pattern at small Δt is sensitive to Im δ_K (Buchanan et al 1992).

2   Continuous Global Symmetries

In the Standard Model there are a variety of global symmetries, both exact and approximate. Some of these symmetries are manifest [Wigner-Weyl realized], while others are spontaneously broken [Nambu-Goldstone realized]. I wish here to examine these matters in some detail.

An important distinction exists for a continuous global symmetry depending on whether or not the vacuum state respects the symmetry. Let us denote the global symmetry group for the theory by G. This group, in general, will have generators g_i which obey an algebra

$$[g_i, g_j] = i c_{ijk} g_k \;, \tag{116}$$

where c_{ijk} are the structure constants for the group. If the generators g_i, for all i, annihilate the vacuum

$$g_i |0\rangle = 0 \;, \tag{117}$$

then the symmetry group is realized in a Wigner-Weyl way, with degenerate multiplets of states in the spectrum (Wigner 1952 and Weyl 1929). If, on the other hand, for some generators g_i

$$g_i |0\rangle \neq 0 \tag{118}$$

then the symmetry group G is spontaneously broken to a subgroup H ($G \to H$) and $n = \dim G/H$ massless scalars appear in the spectrum of the theory. This is the Nambu-Goldstone realization of the symmetry G and the massless scalars are known as Nambu-Goldstone bosons (Nambu 1980 and Goldstone 1981).

Physically, approximate global symmetries are easy to understand. These symmetries result from being able to neglect dynamically certain parameters in

the theory. A well known example is provided by Quantum Chromodynamics (QCD). The Lagrangian of QCD

$$\mathcal{L}_{\mathrm{QCD}} = - \sum_i \bar{q}_i \left(\gamma^\mu \frac{1}{i} D_\mu + m_i \right) q_i - \frac{1}{4} G_a^{\mu\nu} G_{a\mu\nu} \tag{119}$$

has an approximate global symmetry, connected to the fact that the lightest quark masses m_u and m_d are much smaller than the dynamical scale of the theory, Λ_{QCD}.[10] Neglecting the light quark masses, one sees that the QCD Lagrangian is invariant under a large global symmetry transformation

$$\mathcal{L}_{\mathrm{QCD}} \stackrel{U(n_f)_{\mathrm{L}} \times U(n_f)_{\mathrm{R}}}{\longrightarrow} \mathcal{L}_{\mathrm{QCD}} \ , \tag{120}$$

where n_f is the number of flavors whose masses are neglected. Under this group of transformations the n_f light quarks go into each other. For example, for $n_f = 2$, neglecting m_u and m_d in the QCD Lagrangian allows the symmetry transformation

$$\begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{L}} \to e^{ia_{i\mathrm{L}} T_i} \begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{L}} \ ; \quad \begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{R}} \to e^{ia_{i\mathrm{R}} T_i} \begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{R}} \ ; \tag{121}$$

where $T_i = (\tau_i, 1)$.

The global $U(2)_{\mathrm{L}} \times U(2)_{\mathrm{R}}$ approximate symmetry of QCD, arising from the fact that $m_u, m_d \ll \Lambda_{\mathrm{QCD}}$, is actually only a symmetry at the classical level. At the quantum level, there is an Adler-Bell-Jackiw (Adler, Bell and Jackiw 1969) anomaly in a $U(1)_{\mathrm{R-L}}$ subgroup of this symmetry and the real approximate global symmetry of QCD is reduced to

$$G = SU(2)_{\mathrm{R+L}} \times SU(2)_{\mathrm{R-L}} \times U(1)_{\mathrm{R+L}} \equiv SU(2)_V \times SU(2)_A \times U(1)_B \ . \tag{122}$$

Only $SU(2)_V$ and $U(1)_B$, however, are manifest symmetries of nature. The $SU(2)_A$ symmetry is spontaneously broken by the formation of u and d quark condensates, due to the QCD dynamics (see, for example, Donoghue et al 1992)

$$\langle \bar{u}u \rangle = \langle \bar{d}d \rangle \neq 0 \ . \tag{123}$$

The manifest $SU(2)_V$ symmetry, is the well-known isospin symmetry of the strong interactions (Heisenberg 1932), leading to the approximate nucleon $N = (p, n)$ and pion $\pi = (\pi^\pm, \pi^0)$ multiplets. $U(1)_B$ corresponds to baryon number and its existence as a good symmetry guarantees that nucleons and antinucleons have the same mass. The spontaneously broken $SU(2)_A$ symmetry leads to the appearance of three Nambu-Goldstone bosons, which are identified as the pions. Indeed, one can show that (see, for example, Peccei 1987)

$$m_\pi^2 \to 0 \ \ \text{as} \ \ m_{u,d} \to 0 \ . \tag{124}$$

---

[10] The strange quark mass $m_s \sim \Lambda_{\mathrm{QCD}}$ may also be neglected in some circumstances, leading to a larger $SU(3) \times SU(3)$ global symmetry.

Although $SU(2)_V \times SU(2)_A$ are only **approximate** symmetries of QCD, valid of we neglect m_u and m_d in the QCD Lagrangian, $U(1)_B$ is actually an exact global symmetry of the theory corresponding to the transformation

$$q_i \rightarrow \exp\left[\frac{i}{3}\alpha_B\right] q_i . \tag{125}$$

This transformation, since it affects all quarks equally, is also clearly a symmetry of the electroweak theory. Indeed, since all interactions always involve $q-\bar{q}$ pairs, it follows immediately that

$$\mathcal{L}_{\text{SM}} \xrightarrow{U(1)_B} \mathcal{L}_{\text{SM}} , \tag{126}$$

with the associated conserved current being given by

$$J_B^\mu = \frac{1}{3}\sum_i \bar{q}_i \gamma^\mu q_i . \tag{127}$$

Precisely the same argument can be made for leptons, since again all interactions in the Standard Model always involve a lepton-antilepton pair. Whence, one has

$$\mathcal{L}_{\text{SM}} \xrightarrow{U(1)_L} \mathcal{L}_{\text{SM}} , \tag{128}$$

with

$$J_{\text{L}}^\mu = \sum_i \bar{\ell}_i \gamma^\mu \ell_i \tag{129}$$

being the corresponding conserved current.

At the quantum level, however, it turns out that neither $U(1)_{\text{L}}$ or $U(1)_{\text{B}}$ are good symmetries, because of the chiral nature of the weak interactions. Because the left-handed fields under the $SU(2) \times U(1)$ Standard Model group behave differently than the right-handed fields, effectively in the electroweak theory both J_{B}^μ and J_{L}^μ feel corresponding ABJ anomalies ('t Hooft 1976a). As we shall see, the breaking of $U(1)_{\text{B}}$ and $U(1)_{\text{L}}$ by these anomalies is the same. Hence, in the electroweak theory, at the quantum level, there remains only one true global quantum symmetry, $U(1)_{\text{B}-\text{L}}$:

$$\mathcal{L}_{\text{SM}} \xrightarrow{U(1)_{\text{B}-\text{L}}} \mathcal{L}_{\text{SM}} . \tag{130}$$

We shall soon discuss these matters in some detail. However, before doing so, let me remark that the electroweak theory has actually a larger set of global symmetries if the neutrino masses vanish ($m_{\nu_i} = 0$).[11] In this case, each **individual** lepton number (L_e, L_μ and L_τ) is separately conserved at the classical level, while, say, $3L_e - B$, $3L_\mu - B$, $3L_\tau - B$ are conserved at the quantum level.

If one includes right-handed neutrinos in the standard model, so that $m_{\nu_i} \neq 0$, then one expects in general neutrino mixing, much as in the quark case.

---

[11] Theoretically, this is simply achieved by not including any right-handed neutrino fields ν_{R_i} in the theory.

One knows, however, experimentally that neutrino masses, if they exist at all are very light (Particle Data Group 1996)—typically with masses in the eV range. With such light neutrino masses, effectively the Standard Model produces extremely small lepton flavor violations. For instance, one knows experimentally that (Particle Data Group 1996)

$$BR(\mu \rightarrow e\gamma) < 5 \times 10^{-11} \ . \tag{131}$$

Such a transition can occur at the one-loop level in the SM, but its ratio is extremely suppressed due to the tiny neutrino masses (Pal and Wolfenstein 1982). Typically, one finds

$$BR(\mu \rightarrow e\gamma) \sim \frac{\alpha G_F \sin\theta_\nu (m_{\nu_i}^2 - m_{\nu_2}^2)}{M_W^2} \sim 10^{-24} \ . \tag{132}$$

Here θ_ν is a neutrino mixing angle and the numerical result corresponds to taking $\sin\theta_\nu \sim 10^{-1}$ and $\Delta m_\nu^2 \sim (\text{eV})^2$.

2.1   Chiral Anomalies

The existence of chiral anomalies (Adler, Bell and Jackiw 1969) has important consequences for the Standard Model. Anomalies, as we shall see, alter the classical global symmetry structure of the model. In addition, they bring into play the gauge field strength structure

$$F_a^{\mu\nu} \tilde{F}_{a\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} F_{a\alpha\beta} F_{a\mu\nu} \ . \tag{133}$$

This structure is C even, but is both P and T odd. Hence, it can provide additional sources of CP violation. In the Standard Model, it does so through the, so-called, $\bar{\theta}$-term effective interaction

$$\mathcal{L}_{\text{CP viol.}} = \bar{\theta} \frac{\alpha_3}{8\pi} G_a^{\mu\nu} \tilde{G}_{a\mu\nu} \ , \tag{134}$$

where $G_a^{\mu\nu}$ is the gluon field strength for QCD and α_3 is the corresponding (squared) coupling constant [$\alpha_3 = g_3^2/4\pi$].

For pedagogical reasons, it is important to sketch the *raison d'etre* for chiral anomalies. This is done best in the simple example provided by a theory which has a single fermion field ψ and a $U(1)_V \times U(1)_A$ global symmetry. In such a theory, at the classical (Lagrangian) level there are two conserved currents

$$J_V^\mu = \bar{\psi}\gamma^\mu\psi \ \ \text{with} \ \ \partial_\mu J_V^\mu = 0 \tag{135}$$

and

$$J_A^\mu = \bar{\psi}\gamma^\mu\gamma_5\psi \ \ \text{with} \ \ \partial_\mu J_A^\mu = 2m\bar{\psi}i\gamma_5\psi \xrightarrow{m\rightarrow 0} 0 \ . \tag{136}$$

That is, the chiral $U(1)_A$ symmetry obtains if the fermion ψ is massless. At the quantum level, however, it is not possible to preserve **both** the conservation laws
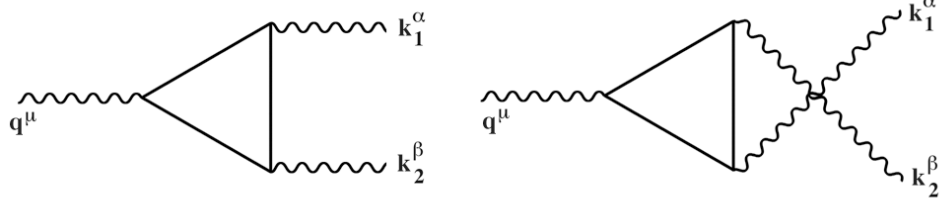
Fig. 1. Triangle graphs contributing to the AVV anomaly

for J_A^μ and J_V^μ. This is the origin of the chiral anomaly (Adler, Bell and Jackiw 1969).

More specifically, the source of the anomaly is the singular behavior of the triangle graph (shown in Fig. 1) involving one axial current J_A^μ and two vector currents J_V^μ. The individual graphs in Fig. 1 are each logarithmic divergent. However, their sum is finite. One can write the Green's function for two vector currents J_V^μ and an axial current as (Adler 1970)

$$T^{\mu\alpha\beta} = F(q^2, k_1^2, k_2^2) P^{\mu\alpha\beta}(k_1, k_2) . \tag{137}$$

The pseudotensor $P^{\mu\alpha\beta}(k_1, k_2)$ by Bose symmetry obeys

$$P^{\mu\alpha\beta}(k_1, k_2) = P^{\mu\beta\alpha}(k_2, k_1) . \tag{138}$$

Further, the conservation of the vector currents imposes the constraints

$$k_{1\alpha} P^{\mu\alpha\beta}(k_1, k_2) = k_{2\beta} P^{\mu\alpha\beta}(k_1, k_2) = 0 . \tag{139}$$

The above equations imply a unique structure for the pseudotensor $P^{\mu\alpha\beta}(k_1, k_2)$, namely

$$P^{\mu\alpha\beta}(k_1, k_2) = \epsilon^{\alpha\beta\rho\sigma} k_{1\rho} k_{2\sigma} q^\mu . \tag{140}$$

Because of the momentum factors in $P^{\mu\alpha\beta}(k_1, k_2)$, it follows that the invariant function $F(q^2, k_1^2, k_2^2)$ is indeed finite.

Given the above, imagine regularizing the triangle graphs in Fig. 1 via a Pauli-Villars regularization, to make each of the individual graphs finite (Adler 1970). Denoting the graphs in Fig. 1, respectively, by $t^{\mu\alpha\beta}(k_1, k_2)$ and $t^{\mu\beta\alpha}(k_2, k_1)$ this procedure yields for $T^{\mu\alpha\beta}$ the expression

$$\begin{aligned}
T^{\mu\alpha\beta} &= \epsilon^{\alpha\beta\rho\sigma} k_{1\rho} k_{2\sigma} q^\mu F(q^2, k_1^2, k_2^2) \\
&= \left[t^{\mu\alpha\beta}(k_1, k_2) \big|_m - t^{\mu\alpha\beta}(k_1, k_2) \big|_M \right] \\
&\quad + \left[t^{\mu\beta\alpha}(k_2, k_1) \big|_m - t^{\mu\beta\alpha}(k_2, k_1) \big|_M \right] .
\end{aligned} \tag{141}$$

Here M is the Pauli-Villars regularization mass. Taking the divergence of the above and setting the fermion mass $m \to 0$ yields the expression

$$q_\mu T^{\mu\alpha\beta} = -2iMP^{\alpha\beta}(M) \ . \tag{142}$$

Here the pseudoscalar structure $P^{\alpha\beta}(M)$ involves similar graphs to those in Fig. 1, except that the axial vertex is proportional to γ_5 and not $\gamma^\mu\gamma_5$.

Because the function $F(q^2, k_1^2, k_2^2)$ is finite, one knows that the Pauli-Villars regularization is really irrelevant and that one can therefore let $M \to \infty$. By straightforward calculation (Adler 1970) one finds that

$$\lim_{M\to\infty} -2iMP^{\alpha\beta}(M) = \frac{i}{2\pi^2}\epsilon^{\alpha\beta\rho\sigma}k_{1\rho}k_{2\sigma} \ . \tag{143}$$

Hence, one deduces the Adler-Bell-Jackiw anomalous divergence equation (Adler, Bell and Jackiw 1969)

$$q_\mu T^{\mu\alpha\beta} = \frac{i}{2\pi^2}\epsilon^{\alpha\beta\rho\sigma}k_{1\rho}k_{2\sigma} \ . \tag{144}$$

The anomalous Ward identity for $T^{\mu\alpha\beta}$ above can be interpreted in terms of an effective violation of the conservation equation for the axial current J_A^μ. Because the $U(1)_V$ gauge bosons- "photons"-couple to J_V^α and J_V^β, it is easy to show that Eq. (144) is equivalent to the anomalous divergence equation

$$\partial_\mu J_A^\mu = \frac{e^2}{8\pi^2}F_{\alpha\beta}\tilde{F}^{\alpha\beta} = \frac{\alpha}{2\pi}F_{\alpha\beta}\tilde{F}^{\alpha\beta} \ , \tag{145}$$

where e is the $U(1)_V$ coupling constant. The above is the famous Adler-Bell-Jackiw chiral anomaly (Adler, Bell and Jackiw 1969).

The above result, whose derivation we sketched for the $U(1)_V \times U(1)_A$ theory, can easily be generalized to the case where the fields in the current J_A^μ carry some non-Abelian charge. In this case the fermions in the anomalous triangle graphs carry some non-Abelian index and the graph, instead of simply involving e^2, now contains a factor of

$$g^2 \ \mathrm{Tr} \ \frac{\lambda_a}{2}\frac{\lambda_a}{2} = \frac{1}{2}g^2\delta_{ab} \ . \tag{146}$$

Here g is the coupling constant associated to the non-Abelian group and $\lambda_a/2$ is the appropriate generator matrix for the fermion fields, assuming they transform according to the fundamental representation of the non-Abelian group. It follows, therefore, that in the non-Abelian case the chiral anomaly (145) is replaced by

$$\partial_\mu J_A^\mu = \frac{g^2}{16\pi^2}F_a^{\alpha\beta}\tilde{F}_{a\alpha\beta} = \frac{\alpha_g^2}{4\pi}F_a^{\alpha\beta}\tilde{F}_{a\alpha\beta} \ , \tag{147}$$

where $F_a^{\alpha\beta}$ are the field strengths for the non-Abelian gauge bosons.

One can use the above results to analyze the Baryon (B) and Lepton (L) number currents in the Standard Model ('t Hooft 1976a). These currents, as

we mentioned earlier, are conserved at the Lagrangian level. Decomposing these currents into chiral components, one has

$$J_{\rm B}^{\mu} = \frac{1}{3} \sum_i \bar{q}_i \gamma^{\mu} q_i = \frac{1}{3} \sum_i (\bar{q}_{i\rm L} \gamma^{\mu} q_{i\rm L} + \bar{q}_{i\rm R} \gamma^{\mu} q_{i\rm R})$$

$$J_{\rm L}^{\mu} = \sum_i \bar{\ell}_i \gamma^{\mu} \ell_i = \sum_i (\bar{\ell}_{i\rm L} \gamma^{\mu} \ell_{i\rm L} + \bar{\ell}_{i\rm R} \gamma^{\mu} \ell_{i\rm R}) \ . \tag{148}$$

Because the quarks and leptons interact with the $SU(2) \times U(1)$ electroweak fields the divergence of $J_{\rm B}^{\mu}$ and $J_{\rm L}^{\mu}$ will not vanish, as a result of the chiral anomalies. A straightforward computation of the relevant triangle graphs gives

$$\partial_{\mu} J_{\rm B}^{\mu} = -\frac{\alpha_2}{8\pi} N_g W_i^{\mu\nu} \tilde{W}_{i\mu\nu} + \frac{\alpha_1}{8\pi} N_g \left(\frac{4}{9} + \frac{1}{9} - \frac{1}{18} \right) Y^{\alpha\beta} \tilde{Y}_{\alpha\beta} \tag{149}$$

and

$$\partial_{\mu} J_{\rm L}^{\mu} = -\frac{\alpha_2 N_g}{8\pi} W_i^{\mu\nu} \tilde{W}_{i\mu\nu} + \frac{\alpha_1}{8\pi} N_g \left(1 - \frac{1}{2} \right) Y^{\alpha\beta} \tilde{Y}_{\alpha\beta} \ . \tag{150}$$

In the above, N_g is the number of generations. The various numbers in front of the contributions involving the $U(1)$ gauge bosons contain the squares of the appropriate hypercharges, multiplied by the corresponding number of states [e.g. $u_{\rm R}$ contributes a factor of 4/9, while the doublet $(u,d)_{\rm L}$ contributes a factor of $2 \times 1/36$]. Note that for the Baryon number current and for the Lepton number current, not only the $SU(2)$ but also the $U(1)$ factors are the same [(4/9 + 1/9-1/18) = (1-1/2) = 1/2]. It follows therefore that, as advertized, the total fermion number B+L is broken at the quantum level, but B-L is conserved:

$$\partial_{\mu} J_{\rm B+L}^{\mu} = \frac{\alpha_1^2}{8\pi} N_g Y^{\alpha\beta} \tilde{Y}_{\alpha\beta} - \frac{\alpha_2^2}{4\pi} N_g W_i^{\alpha\beta} \tilde{W}_{i\alpha\beta}$$

$$\partial_{\mu} J_{\rm B-L}^{\mu} = 0 \ . \tag{151}$$

A similar situation obtains in QCD. In the limit as $m_u, m_d \to 0$, this theory has a global symmetry at the classical level of $SU(2)_V \times SU(2)_A \times U(1)_V \times U(1)_A$. However, the $U(1)_A$ current

$$J_5^{\mu} = \frac{1}{2} [\bar{u} \gamma^{\mu} \gamma_5 u + \bar{d} \gamma^{\mu} \gamma_5 d] \tag{152}$$

has a chiral anomaly, since the quarks carry color and interact with the gluons. Taking into account the contribution of both the u and d quarks in the triangle graph, one finds

$$\partial_{\mu} J_5^{\mu} = \frac{\alpha_3^2}{4\pi} G_a^{\alpha\beta} \tilde{G}_{a\alpha\beta} \ . \tag{153}$$

The violation of the (B+L)-current in the electroweak theory and of the $U(1)_A$ current in QCD, codified by Eqs. (151) and (153), have a similar aspect. Nevertheless, these quantum corrections are quite different physically in their import. As we shall see, the current J_5^{μ} is really badly broken by the above

quantum QCD effects. As a result, as we mentioned earlier, the classical $U(1)_A$ symmetry is never a good (approximate) symmetry of the strong interactions. In contrast, $J^\mu_{\text{B+L}}$ is extraordinarily weakly broken by the quantum corrections, except in the early Universe where temperature-dependent effects enhance these contributions. Thus, at zero temperature, the total fermion number (B+L) is essentially conserved.

Physically, these two results are what is needed. The formation of u and d-quark condensates

$$\langle \bar{u}u \rangle = \langle \bar{d}d \rangle \neq 0 \tag{154}$$

in QCD clearly breaks both the $SU(2)_A$ and $U(1)_A$ symmetries spontaneously. If $U(1)_A$ were really a symmetry, one would expect to have an associated Nambu-Goldstone boson—the η—with similar properties to the $SU(2)_A$ Nambu-Goldstone bosons—the π mesons. Although these states are supposed to be massless when the respective global symmetries are exact, both states should get similar masses once one includes quark mass terms for the u and d quarks (Weinberg 1975). However, experimentally, one finds $m^2_\eta \gg m^2_\pi$ and one concludes that $U(1)_A$ cannot really be a true symmetry of QCD. Thus the strong breaking of J^μ_A by the anomaly is a welcome result.

In contrast, for the electroweak theory it is important that the anomalous breaking of (B+L) should not physically lead to large effects, since one has very strong experimental bounds on baryon number violation. For instance the B-violating decay $p \to e^+\pi^0$ has a bound (Particle Data Group 1996)

$$\tau(p \to e^+\pi^0) > 5.5 \times 10^{32} \text{ years} . \tag{155}$$

To undersand why the anomaly contribution in Eq. (153) connected to the $U(1)_A$ current is important, while the anomaly contribution in Eq. (151) connected to the (B+L) current is irrelevant, requires an examination of the properties of the gauge theory vacuum. We turn to this next.

2.2   The Gauge Theory Vacuum

The resolution of the above issues came through a better understanding of the vacuum structure of gauge theories ('t Hooft 1976b and Polyakov 1977). The vacuum state is, by definition, a state where all fields vanish. For gauge fields, this needs to be slightly extended since these fields themselves are not physical. So, in the case of gauge fields, the vacuum state is one where either $A^\mu_a = 0$ or the gauge fields are a gauge transformation of $A^\mu_a = 0$. For our purposes it suffices to examine an $SU(2)$ gauge theory, since this example serves to exemplify what happens in a more general case.

It proves particularly convenient (Callan, Dashen and Gross 1976) to study the $SU(2)$ gauge theory in a temporal gauge where $A^0_a = 0$ $\{a = 1, 2, 3\}$. In this gauge the space components of the gauge fields are time-independent $A^i_a(\mathbf{r}, t) = A^i_a(\mathbf{r})$. Even so, there is still some residual gauge freedom. Defining a gauge matrix $A^i(\mathbf{r})$ by contracting the gauge fields with the Pauli matrices, $A^i(\mathbf{r}) =$

$\frac{\tau_a}{2}A_a^i(\mathbf{r})$, in the $A_a^0 = 0$ gauge one is left with the freedom to perform the following gauge transformations

$$A^i(\mathbf{r}) \rightarrow \Omega(\mathbf{r})A^i(\mathbf{r})\Omega(\mathbf{r})^{-1} + \frac{i}{g}\Omega(\mathbf{r})\nabla^i\Omega(\mathbf{r})^{-1} , \qquad (156)$$

where g is the gauge coupling for the $SU(2)$ theory. In view of the above, one concludes that in the $A_a^0 = 0$ gauge, pure gauge fields corresponding to the vacuum configuration are simply the set $\{0, \frac{i}{g}\Omega(\mathbf{r})\nabla^i\Omega(\mathbf{r})^{-1}\}$.

The behavior of $\Omega(\mathbf{r})$ as $\mathbf{r} \rightarrow \infty$ distinguishes classes of pure gauge fields. In particular, the requirement that (Callan, Dashen and Gross 1976)

$$\Omega(\mathbf{r}) \overset{\mathbf{r}\rightarrow\infty}{\longrightarrow} 1, \qquad (157)$$

provides a map of physical space $[S_3]$ onto the group space $[SU(2) \sim S_3]$. This $S_3 \rightarrow S_3$ map splits the matrices $\Omega(\mathbf{r})$ into different homotopy classes $\{\Omega_n(\mathbf{r})\}$, characterized by an integer n—the winding number—specifying how $\Omega(\mathbf{r})$ goes to unity at spatial infinity:

$$\Omega_n(\mathbf{r}) \overset{\mathbf{r}\rightarrow\infty}{\longrightarrow} e^{2\pi in} . \qquad (158)$$

Thus the set of pure gauge fields is $\{0, A_n^i(\mathbf{r})\}$, where

$$A_n^i(\mathbf{r}) = \frac{i}{g}\Omega_n(\mathbf{r})\nabla^i\Omega_n(\mathbf{r})^{-1} . \qquad (159)$$

The winding number n is just the Jacobian of the $S_3 \rightarrow S_3$ transformation (Crewther 1978) and one can show that

$$n = \frac{ig^3}{24\pi^2} \int d^3r \ \text{Tr} \ \epsilon_{ijk}A_n^i(\mathbf{r})A_n^j(\mathbf{r})A_n^k(\mathbf{r}) . \qquad (160)$$

Furthermore, one can construct the transformation matrix $\Omega_n(\mathbf{r})$ with winding number n by compounding n-times the transformation matrix of unit winding

$$\Omega_n(\mathbf{r}) = [\Omega_1(\mathbf{r})]^n . \qquad (161)$$

A representative $n = 1$ matrix, giving rise to a, so called, **large gauge transformation** is given by

$$\Omega_1(\mathbf{r}) = \frac{\mathbf{r}^2 - \lambda^2}{\mathbf{r}^2 + \lambda^2} + \frac{2i\lambda\boldsymbol{\tau} \cdot \mathbf{r}}{\mathbf{r}^2 + \lambda^2} , \qquad (162)$$

with λ an arbitrary scale parameter.

Using the above properties, it is clear that the n-vacuum state—corresponding to the pure gauge field configuration $A_n^i(\mathbf{r})$—is not fully gauge invariant. Indeed, a large gauge transformation can change the gauge field $A_n^i(\mathbf{r})$ into that of $A_{n+1}^i(\mathbf{r})$

$$A_{n+1}^i(\mathbf{r}) = \Omega_1(\mathbf{r})A_n^i(\mathbf{r})\Omega_1^{-1}(\mathbf{r}) + \frac{i}{g}\Omega_1(\mathbf{r})\nabla^i\Omega_1^{-1}(\mathbf{r}) , \qquad (163)$$

or

$$\Omega_1 |n\rangle = |n+1\rangle \ . \tag{164}$$

The correct vacuum state for a gauge theory must be gauge invariant. As such it must be a linear superposition of these n-vacuum states. This is the, so-called, θ-vacuum ('t Hooft 1976b and Polyakov 1977)

$$|\theta\rangle = \sum_n e^{-in\theta} |n\rangle \ . \tag{165}$$

Clearly, since

$$\Omega_1 |\theta\rangle = \sum_n e^{-in\theta} \Omega_1 |n\rangle = \sum_n e^{-in\theta} |n+1\rangle = e^{i\theta} |\theta\rangle \ , \tag{166}$$

the $|\theta\rangle$ vacuum is gauge invariant.

Using the θ-vacuum as the correct vacuum state for gauge theories, it is clear that the vacuum functional for these theories splits into distinct sectors (Callan, Dashen and Gross 1976). If $|\theta\rangle_{\pm}$ are the θ-vacuum states at $t = \pm\infty$, then the vacuum functional for a gauge theory takes the form

$$_+\langle\theta|\theta\rangle_- = \sum_{n,m} e^{im\theta} e^{-in\theta} \ _+\langle m|n\rangle_-$$

$$= \sum_{\nu} e^{i\nu\theta} \left[\sum_n \ _+\langle n+\nu|n\rangle_- \right] \ . \tag{167}$$

That is, the vacuum functional sums over vacuum to vacuum amplitudes in which the winding number at $t = \pm\infty$ differ by ν, weighing each by a factor $e^{i\nu\theta}$. We anticipate here that the superposition of amplitudes with **different** phases $e^{i\nu\theta}$ will lead to CP-violating effects. Recalling that the vacuum functional is given by a path integral over gauge field configurations, each weighted by the classical action, one arrives at the formula

$$_+\langle\theta|\theta\rangle_- = \int_{\text{Paths}} \delta A_\mu e^{iS[A]} = \sum_\nu e^{i\nu\theta} \left[\sum_\nu \ _+\langle n+\nu|n\rangle_- \right] \ . \tag{168}$$

Although the formula for $_+\langle\theta|\theta\rangle_-$ above was derived in the A_a^0 gauge, the parameter ν entering in this formula has actually a gauge invariant meaning. One finds ('t Hooft 1976b and Polyakov 1977)

$$\nu = n_+ - n_- = \frac{g^2}{32\pi^2} \int d^4x G_a^{\mu\nu} \tilde{G}_{a\mu\nu} \ . \tag{169}$$

To prove this result requires using Bardeen's identity (Bardeen 1972) which expresses the product of $G\tilde{G}$ as a total derivative:

$$G_a^{\mu\nu} \tilde{G}_{a\mu\nu} = \partial_\mu K^\mu \ , \tag{170}$$

where the "current" K^μ is given by

$$K^\mu = \epsilon^{\mu\alpha\beta\gamma} A_{a\alpha} \left[G_{a\beta\gamma} - \frac{g}{3}\epsilon_{abc} A_{b\beta} A_{c\gamma} \right] . \tag{171}$$

For pure gauge fields $[G_{a\beta\gamma} = 0]$ and in the $A_a^0 = 0$ gauge this curent has only a temporal component:

$$K^i = 0; \ K^0 = -\frac{g}{3}\epsilon_{ijk}\epsilon_{abc} A_a^i A_b^j A_c^k = \frac{4}{3}ig\epsilon_{ijk} \operatorname{Tr} A^i A^j A^k . \tag{172}$$

Using these relations, in this gauge one can write the winding numbers $n_\pm$ as

$$n_\pm = \frac{ig^3}{24\pi^2} \int d^3r \epsilon_{ijk} \operatorname{Tr} A^i A^j A^k = \frac{g^2}{32\pi^2} \int d^3r K^0 \,|_{t=\pm\infty} . \tag{173}$$

The above formula allows one to express the winding number difference $\nu = n_+ - n_-$ as

$$\nu = n_+ - n_- = \frac{g^3}{32\pi^2} \int d^3r K^0|_{t=-\infty}^{t=+\infty} = \frac{g^2}{32\pi^2} \int d\sigma_\mu K^\mu . \tag{174}$$

Whence, Eq. (169) follows by using Gauss's theorem and Bardeen's identity.

Having identified ν as an integral over $G\tilde{G}$, one can rewrite the formula for the vacuum functional in terms of an effective action. Defining

$$S_{\text{eff}}[A] = S[A] + \theta\frac{g^2}{32\pi^2} \int d^4x G_a^{\mu\nu} \tilde{G}_{a\mu\nu} , \tag{175}$$

one sees that

$${}_+\langle\theta|\theta\rangle_- = \sum_\nu \int_{\text{Paths}} \delta A_\mu e^{iS_{\text{eff}}[A]} \delta\left[\nu - \frac{g^2}{32\pi^2} \int d^4x G_a^{\mu\nu} \tilde{G}_{a\mu\nu}\right] . \tag{176}$$

The more complicated structure of the gauge theory vacuum [θ-vacuum] effectively adds an additional term to the gauge theory Lagrangian:

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{gauge theory}} + \theta\frac{g^2}{32\pi^2} G_a^{\mu\nu} \tilde{G}_{a\mu\nu} . \tag{177}$$

Perturbation theory is connected to the $\nu = 0$ sector, since $\int d^4x G\tilde{G} = 0$. Effects of non-zero winding number differences ($\nu \neq 0$) involve **non-perturbative** contributions. These are naturally selected by the connection of the pseudoscalar density $G\tilde{G}$ with the divergence of chiral currents, through the chiral anomaly (Adler, Bell and Jackiw 1969).

Let me examine this first for QCD. Assuming there are n_f flavors whose mass can be neglected ($m_f = 0$), the axial current in QCD

$$J_5^\mu = \frac{1}{2}\sum_{i=1}^{n_f} \bar{q}_i\gamma^\mu\gamma_5 q_i \tag{178}$$

is still not conserved as a result of the chiral anomaly. One has

$$\partial_\mu J_5^\mu = n_f \frac{g_3^2}{32\pi^2} G_a^{\mu\nu} \tilde{G}_{a\mu\nu} \ . \tag{179}$$

In view of the above, chirality changes ΔQ_5, are simply related to ν:

$$\Delta Q_5 = \int d^4x \partial_\mu J_5^\mu = n_f \frac{g_3^2}{32\pi^2} \int d^4x G_a^{\mu\nu} \tilde{G}_{a\mu\nu} = n_f \nu \ . \tag{180}$$

Clearly, if $\nu \neq 0$ sectors are important in QCD, then the above changes are important and the corresponding $U(1)_A$ symmetry is **never** a symmetry of the theory. This then is the physical explanation why (in the relevant $n_f = 2$ case) the η does not have the properties of a Goldstone boson.

't Hooft ('t Hooft 1976c), by using semiclassical methods, provided an estimate of the likelyhood of the occurence of processes involving $\nu \neq 0$ transitions. Basically, he viewed the transition from an n-vacuum at $t = -\infty$ to an $(n + \nu)$-vacuum at $t = +\infty$ as a tunneling process and estimated the tunneling probability by WKB methods. 't Hooft's result ('t Hooft 1976c)

$$A[\nu] \sim e^{-S_E[\nu]} \tag{181}$$

uses as the WKB factor in the exponent the minimal Euclidean action for the gauge theory. Such a minimal action obtains if the gauge field configurations are those provided by instantons (Belavin et al 1975). These are self-dual solutions of the field equations in Euclidean space $[G_a^{\mu\nu} = \tilde{G}_a^{\mu\nu}]$ and their action is simply related to ν. For these solutions

$$S_E[\nu] = \frac{1}{4} \int d^4x_E G_a^{\mu\nu} G_a^{\mu\nu} = \frac{1}{4} \int d^4x_E G_a^{\mu\nu} \tilde{G}_a^{\mu\nu} = \frac{8\pi^2}{g_3^2} \nu \ . \tag{182}$$

What 't Hooft showed in his careful calculation ('t Hooft 1976c) is that the coupling constant that enters in $S_E[\nu]$ is actually a running coupling, with its scale set by the scale of the instanton solution involved. Further, to evaluate the amplitude in question one must integrate over all such scales. Thus, schematically, 't Hooft's result is

$$A[\nu] \sim \int d\rho \ \exp\left[-\frac{2\pi\nu}{\alpha_3(\rho^{-1})}\right] \ . \tag{183}$$

In QCD, since the gauge coupling squared $\alpha_3(\rho^{-1})$ grows for large distances, there is no particular suppression due to the tunneling factor for large size instantons. Because of this, although one cannot really calculate $A[\nu]$, one expects that

$$A[\nu \neq 0] \sim A[0] \ . \tag{184}$$

Thus, as advertized, $U(1)_A$ is not really a symmetry of QCD.

Much of the above discussion applies to the electroweak theory. However, as we shall see, there is a crucial difference. Since the electroweak theory is based on the group $SU(2) \times U(1)$, because of the $SU(2)$ factor there is also here a nontrivial vacuum structure. The $W\tilde{W}$ density connected to the index difference in

this case is directly related to the divergence of the B+L current. Focusing on this contribution, one has

$$\partial_\mu J^\mu_{\text{B+L}} = -\frac{g_2^2}{16\pi^2} N_g W_i^{\mu\nu} \tilde{W}_{i\mu\nu} \ . \tag{185}$$

Hence, the change in (B+L) in the electroweak theory is also simply connected to the (weak) index ν ('t Hooft 1976a)

$$\Delta(B+L) = \int d^4x \partial_\mu J^\mu_{\text{B+L}} = -\frac{g_2^2}{16\pi^2} N_g \int d^4x W_i^{\mu\nu} \tilde{W}_{i\mu\nu} = -2N_g\nu \ . \tag{186}$$

I note that for three generations $[N_g = 3]$ the minimal violation of the (B+L)-current is $|\Delta(B+L)| = 6$. So, even though baryon number is violated in the Standard Model the process $p \to e^+\pi^0$, which involves $\Delta(B+L) = 2$, is still forbidden! More importantly, however, the amplitude for (B+L)-violation itself is totally negligible. This amplitude, at least semiclassically, will again be given by a result similar to what was obtained in QCD (except with $\alpha_3 \to \alpha_2$). However, because the electroweak symmetry is broken, the integration over instanton sizes cuts off at sizes of order $1/v$ (or momentum scales of order M_Z). Hence, one estimates ('t Hooft 1976a)[12]

$$A[\nu]_{\text{(B+L)-violation}} \sim \exp\left[-\frac{2\pi\nu}{\alpha_2(M_Z)}\right] \sim 10^{-80\nu} \ . \tag{187}$$

I want to remark that, although the above result is negligibly small, in the early Universe (B+L)-violation in the electroweak theory can be important. This was first observed by Kuzmin, Rubakov, and Shaposhnikov (Kuzmin et al 1985), who pointed out that in a thermal bath the semiclassical estimate of 't Hooft ceases to be accurate. Effectively, in these circumstances, the gauge configurations associated with (B+L)-violating processes are not governed by a tunnelling factor, but by a Boltzman factor. As one nears the electroweak phase transitions, furthermore, this Boltzman factor tends to unity and the (B+L)-violating processes proceed essentially unsuppressed.

2.3   The Strong CP Problem

The θ-vacuum of QCD is a new source of CP-violation,[13] as a result of the effective interaction

$$\mathcal{L}_{\text{CP-violation}} = \theta\frac{\alpha_3}{8\pi} G_a^{\mu\nu} \tilde{G}_{a\mu\nu} \ , \tag{188}$$

---

[12] Here we use $\alpha_2(M_Z) = \frac{\alpha(M_Z)}{\sin^2\theta_W} \sim \frac{1}{30}$.

[13] One can show that the equivalent θ-parameter in the electroweak theory can be rotated away as a result of the chiral nature of these interactions (Krasnikov et al 1978).

which reflects the presence of the vacuum angle. It turns out, in fact, that the situation is a little bit more complicated, because of the electroweak interactions. Recall that the quark mass matrices arising as a result of the spontaneous breakdown of $SU(2) \times U(1)$ are, in general, neither Hermitian nor diagonal

$$\mathcal{L}_{\text{mass}} = -\bar{q}_{\text{L}i} M_{ij} q_{\text{R}j} - \bar{q}_{\text{R}i} (M^\dagger)_{ij} q_{Lj} \ . \tag{189}$$

These matrices can, however, be diagonalized by performing appropriate unitary transformations on the quark fields

$$q_{\text{R}} \to q'_{\text{R}} = U_{\text{R}} q_{\text{R}} \ ; \quad q_{\text{L}} \to q'_{\text{L}} = U_{\text{L}} q_{\text{L}} \ . \tag{190}$$

It is easy to check that part of the above transformations involves a $U(1)_A$ transformation. In fact, the $U(1)_A$ piece of these transformations is just

$$q_{\text{R}} \to q'_{\text{R}} = \exp\left[\frac{i}{2n_f} \text{Arg det } M\right] q_{\text{R}} \equiv \exp\left[\frac{i}{2}\alpha\right] q_{\text{R}}$$

$$q_{\text{L}} \to q'_{\text{L}} = \exp\left[-\frac{i}{2n_f} \text{Arg det } M\right] q_{\text{L}} \equiv \exp\left[-\frac{i}{2}\alpha\right] q_{\text{L}} \ . \tag{191}$$

It turns out that such $U(1)_A$ transformations engender a change in the vacuum angle (Jackiw and Rebbi 1976). Thus they effectively add a contribution to Eq. (188), beyond that of the QCD angle θ.

  To prove this contention (Jackiw and Rebbi 1976), one has to examine carefully what is the result of a chiral $U(1)_A$ transformation. Although the current J_5^μ connected to $U(1)_A$ has an anomaly, it is always possible to construct a conserved current by using the current K^μ which enters in Bardeen's identity (Bardeen 1972). Recalling Eqs. (170) and (179), it is obvious that the desired conserved chiral current $\tilde{J}_5^\mu$ is

$$\tilde{J}_5^\mu = J_5^\mu - \frac{n_f \alpha_3}{4\pi} K^\mu \ . \tag{192}$$

The charge which generates chiral transformations, $\tilde{Q}_5$, needs to be time-independent. By necessity, it must therefore be related to $\tilde{J}_5^\mu$—the conserved current:

$$\tilde{Q}_5 = \int d^3x \tilde{J}_5^0 \ . \tag{193}$$

Although $\tilde{Q}_5$ is time-independent, this charge is not invariant under large gauge transformations, since K^μ is itself not a gauge-invariant current like J_5^μ. One finds

$$\Omega_1 \tilde{Q}_5 \Omega_1 = \Omega_1 \left[Q_5 - \frac{n_f \alpha_3}{4\pi} \int d^3x K^0 \right] \Omega_1 = \tilde{Q}_5 + n_f \ . \tag{194}$$

  Consider the action of a large gauge transformation Ω_1 on a chirally rotated θ-vacuum state $e^{i\alpha\tilde{Q}_5}|\theta\rangle$. One has

$$\Omega_1 \left[e^{i\alpha\tilde{Q}_5}|\theta\rangle \right] = \Omega_1 e^{i\alpha\tilde{Q}_5} \Omega_1^{-1} \Omega_1 |\theta\rangle$$

$$= e^{i(\alpha n_f + \theta)} \left[e^{i\alpha\tilde{Q}_5}|\theta\rangle \right] \ . \tag{195}$$

It follows from the above, immediately, that a chiral $U(1)_A$ rotation indeed shifts the vacuum angle (Jackiw and Rebbi 1976):

$$e^{i\alpha \tilde{Q}_5}|\theta\rangle = |\theta + \alpha n_f\rangle \ . \tag{196}$$

For the electroweak theory, the chiral rotation one needs to perform to diagonalize the quark mass matrices has a parameter $\alpha = \frac{1}{n_f}\det M$. Whence, it follows that the effective CP-violating Lagrangian term arising from the structure of the gauge theory vacuum is

$$\mathcal{L}^{\text{eff}}_{\text{CP-violation}} = \bar{\theta}\frac{\alpha_3}{8\pi}G^{\mu\nu}_a \tilde{G}_{a\mu\nu} \ , \tag{197}$$

where

$$\bar{\theta} = \theta + \text{Arg det } M \ . \tag{198}$$

The effective CP-violating parameter $\bar{\theta}$ is the sum of a QCD contribution—the vacuum angle θ—and an electroweak piece–Arg det M—related to the phase structure of the quark mass matrix.

The interaction (197) is C even, and T and P odd. Thus it violates CP also. It turns out, as we shall see below, that unless $\bar{\theta}$ is very small $[\bar{\theta} \leq 10^{-10}]$ this interaction produces an electric dipole moment for the neutron which is beyond the present experimental bound for this quantity. It is difficult to understand why a parameter like $\bar{\theta}$, which is a sum of two very different contributions, should be so small. This conundrum is known as the strong CP problem.

Before discussing the strong CP problem further, let me first indicate how to calculate the contribution of the effective Lagrangian (197) to the electric dipole moment of the neutron. This is most easily done by transforming the $\bar{\theta}$ interaction from an interaction involving gluons to one involving quarks. For simplicity, let me concentrate on the two-flavor case ($n_f = 2$) and take, again for simplicity, $m_u = m_d = m_q$. In this case, it is easy to see that the chiral $U(1)_A$ transformation

$$\begin{pmatrix} u \\ d \end{pmatrix} \rightarrow \exp\left[i\frac{\bar{\theta}\gamma_5}{4}\right]\begin{pmatrix} u \\ d \end{pmatrix} \tag{199}$$

will get rid of the $\bar{\theta}G\tilde{G}$ term. However, the above transformation will, at the same time, generate a CP-violating γ_5-dependent mass term for the u and d quarks:

$$\mathcal{L}^{\text{eff}}_{\text{CP-violation}} = i\bar{\theta}m_q\left[\bar{u}\frac{\gamma_5}{2}u + \bar{d}\frac{\gamma_5}{2}d\right] \ . \tag{200}$$

One can use the above effective Lagrangian directly to calculate the neutron electric dipole moment. One has, in general

$$d_n \bar{n}\sigma_{\mu\nu}k^\nu\gamma_5 n = \langle n|T(J^{\text{em}}_\mu i\int d^4x \mathcal{L}^{\text{eff}}_{\text{CP-violation}})|n\rangle \ . \tag{201}$$

To arrive at a result for d_n one inserts a complete set of states $|X\rangle$ in the matrix element above and tries to estimate which set of states $|X\rangle$ dominates. In the literature there are two calculations along these lines. Baluni (Baluni 1979)

uses for $|X\rangle$ the odd parity $|N^-_{1/2}\rangle$ states which are coupled to the neutron by $\mathcal{L}^{\text{eff}}_{\text{CP-violation}}$. Crewther *et al.* (Crewther et al 1979), instead, do a soft pion calculation (effectively $|X\rangle \sim |N\pi_{\text{soft}}\rangle$). The result of these calculations are rather similar and lead to an expression for d_n whose form could have been guessed. Namely

$$d_n \sim \frac{e}{M_n}\left(\frac{m_q}{M_n}\right)\bar{\theta} \sim \begin{cases} 2.7 \times 10^{-16}\ \bar{\theta}\ \text{ecm (Baluni 1979)} \\ 5.2 \times 10^{-16}\ \bar{\theta}\ \text{ecm (Crewther et al 1979)} \end{cases} \qquad (202)$$

The present bound on d_n (Particle Data Group 1996) is, at 95% C.L.,

$$d_n < 1.1 \times 10^{-25}\ \text{ecm} . \qquad (203)$$

Whence, to avoid contradictions with experiment, the parameter $\bar{\theta}$ must be less than 2×10^{-10}. Why this should be so is a mystery. This is the strong CP problem.

2.4   The Chiral Solution to the Strong CP Problem

About twenty years ago, Helen Quinn and I (Peccei and Quinn 1977) suggested a possible dynamical solution to the strong CP problem. If our mechanism holds in nature then $\bar{\theta}$ actually vanishes, and there is no need to explain a small numbr like 10^{-10} cropping up in the theory.[14] To "solve" the strong CP problem, Quinn and I postulated that the Lagrangian of the Standard Model was invariant under an additional global $U(1)$ chiral symmetry—$U(1)_{\text{PQ}}$. This required imposing certain constraints on the Higgs sector of the theory, but otherwise appeared perfectly possible. Because the $U(1)_{\text{PQ}}$ symmetry is a chiral symmetry, if this symmetry were exact, it is trivial to see that the $\bar{\theta}G\tilde{G}$ term can be eliminated, since the chiral rotation $\exp\left[-i\frac{\theta}{n_f}\tilde{Q}^{\text{PQ}}_5\right]$ gives

$$\exp\left[-i\frac{\bar{\theta}}{n_f}\tilde{Q}^{\text{PQ}}_5\right]|\bar{\theta}\rangle = |0\rangle . \qquad (204)$$

That is, by a $U(1)_{\text{PQ}}$ transformation the effective vacuum angle $\bar{\theta}$ is set to zero and this parameter is no longer present in the theory. Phyically, however, if $U(1)_{\text{PQ}}$ is an extra global symmetry of the Standard Model, it is not possible for this symmetry to remain unbroken. What Quinn and I showed (Peccei and Quinn 1977) was that, even if $U(1)_{\text{PQ}}$ is spontaneously broken, one still is able to eliminate the $\bar{\theta}G\tilde{G}$ term.

To see this, it is useful to focus on the associated Nambu-Goldstone boson resulting from the spontaneous breakdown of the $U(1)_{\text{PQ}}$ symmetry. This excitation is the axion, first discussed by Weinberg and Wilczek (Weinberg and

---

[14] Even incorporating a $U(1)_{\text{PQ}}$ symmetry into the theory it turns out that CP violating effects in the electroweak interactions do not allow $\bar{\theta}$ to totally vanish. However, the effective $\bar{\theta}$ induced back through weak CP-violation is tiny ($\bar{\theta} \sim 10^{-15}$) (Georgi et al 1986) and well within the bound provided by the neutron electric dipole moment.

Wilczek 1978) in connection with the $U(1)_\mathrm{PQ}$ symmetry. It turns out that the axion is not quite massless, so it is really a pseudo-Goldstone boson (Weinberg 1972). This is a consequence of the $U(1)_\mathrm{PQ}$ symmetry having an anomaly due to QCD interactions. One finds (Weinberg and Wilczek 1978) that the axion mass is of order

$$m_a \sim \frac{\Lambda_\mathrm{QCD}^2}{f} \ , \tag{205}$$

where Λ_QCD typifies the scale of the QCD interactions, while f is the scale of the $U(1)_\mathrm{PQ}$ breakdown. If $f \gg \Lambda_\mathrm{QCD}$, then axions turn out to be very much lighter than ordinary hadrons.

If we denote the axion field by $a(x)$, it turns out that imposing a $U(1)_\mathrm{PQ}$ symmetry on the standard model effectively serves to replace the CP-violating $\bar\theta$ parameter by the dynamical CP-conserving axion field:

$$\bar\theta \to \frac{a(x)}{f} \ . \tag{206}$$

To understand why this is so, recall that since the axion is the Nambu-Goldstone boson of the broken $U(1)_\mathrm{PQ}$ symmetry, this field translates under a $U(1)_\mathrm{PQ}$ transformation:

$$a(x) \xrightarrow{U(1)_\mathrm{PQ}} a(x) + \alpha f \ , \tag{207}$$

where α is the parameter associated with the $U(1)_\mathrm{PQ}$ transformation. Because of Eq. (207), the axion field can only enter in the Lagrangian of the theory through derivative terms. Even though the detailed axion interactions are somewhat model-dependent, this property allows one to understand how to augment the Lagrangian of the Standard Model so that it becomes $U(1)_\mathrm{PQ}$ invariant.

Focussing only on the possible additional contributions due to the inclusion of the axion field, one is lead to the following effective Lagrangian for the theory

$$\mathcal{L}_\mathrm{SM}^\mathrm{eff} = \mathcal{L}_\mathrm{SM} + \bar\theta \frac{\alpha_3}{8\pi} G_a^{\mu\nu} \tilde{G}_{a\mu\nu} - \frac{1}{2} \partial_\mu a \partial^\mu a$$
$$+ \mathcal{L}_\mathrm{axion}^\mathrm{int} \left[\frac{\partial_\mu a}{f} \ ; \psi \right] + \frac{a}{f} \xi \frac{\alpha_3}{8\pi} G_a^{\mu\nu} \tilde{G}_{a\mu\nu} \ . \tag{208}$$

The third term above is the kinetic energy term for the axion field, while the fourth term in Eq. (208) schematically indicates the kind of interactions the axion field can participate in with the other fields [ψ] in the theory. The last term above, as can be noticed, does not involve a derivative of the axion field, thereby violating the usual expectations for Nambu-Goldstone fields. The reason why this term is included, however, is clear. The $U(1)_\mathrm{PQ}$ symmetry is anomalous[15]

$$\partial_\mu J_\mathrm{PQ}^\mu = \xi \frac{\alpha_3}{8\pi} G_a^{\mu\nu} \tilde{G}_{a\mu\nu} \ . \tag{209}$$

This anomaly must be reflected in the effective Lagrangian (208) when one performs a chiral $U(1)_\mathrm{PQ}$ transformation. This is guaranteed by having the last

---

[15] Here ξ is a model-independent number of $O(1)$ (see, for example, Peccei 1989).

term in Eq. (208), since it precisely reproduces the anomaly when the axion field undergoes the $U(1)_{\rm PQ}$ transformation (207).

The last term in Eq. (208), whose origin is intimately connected to the chiral anomaly, because it contains the axion field directly (and not its derivative) provides a potential for the axion field. As a result, it is not true anymore that all values of the vacuum expectation value (VEV) of $a(x)$ are allowed.[16] The minimum of $V_{\rm eff}$ in the vacuum is simply

$$\left\langle \frac{\partial V_{\rm eff}}{\partial_a} \right\rangle = -\frac{\xi}{f}\frac{\alpha_3}{8\pi}\langle G_a^{\mu\nu}\tilde{G}_{a\mu\nu}\rangle \, \big|_{\langle a\rangle \neq 0} \ . \tag{210}$$

What Quinn and I showed (Peccei and Quinn 1977), in essence, is that the periodicity of $\langle G\tilde{G}\rangle$ in the effective vacuum angle $\theta_{\rm eff}$ for the Lagrangian of Eq. (208)

$$\theta_{\rm eff} = \bar{\theta} + \frac{\xi}{f}\langle a(x)\rangle \ , \tag{211}$$

requires that $\theta_{\rm eff} = 0$, or

$$\langle a(x)\rangle = -\frac{f}{\xi}\bar{\theta} \ . \tag{212}$$

As a result of Eq. (212), only the physical axion field

$$a(x)_{\rm phy} = a(x) - \langle a(x)\rangle \tag{213}$$

interacts with the gluon field strengths, eliminating altogether the $\theta G\tilde{G}$ term. Thus, indeed, imposing an additional $U(1)_{\rm PQ}$ symmetry in the Standard Model, even in the case this symmetry is spontaneously broken, solves the strong CP problem.

As we remarked earlier, the axion is actually massive because of the anomaly in the $U(1)_{\rm PQ}$ current. This follows readily from the effective Lagrangian (208). The second derivative of the effective potential $V_{\rm eff}$, which arose precisely because of the chiral anomaly in the $U(1)_{\rm PQ}$ symmetry, when evaluated at its minimum value $\langle a(x)\rangle$ gives for the axion mass squared the value

$$m_a^2 = \left\langle \frac{\partial^2 V_{\rm eff}}{\partial a^2} \right\rangle \bigg|_{\langle a\rangle} = -\frac{\xi}{f}\frac{\alpha_3}{8\pi}\frac{\partial}{\partial a}\langle G_a^{\mu\nu}\tilde{G}_{a\mu\nu}\rangle \bigg|_{\langle a\rangle} \sim \frac{\Lambda_{\rm QCD}^2}{f} \ . \tag{214}$$

Using the above results, it is clear that the effective theory incorporating $U(1)_{\rm PQ}$ and axions no longer suffers from the strong CP problem. All that remains as a signal of this erstwhile problem is the direct interaction of the (massive) axion field with the gluonic pseudoscalar density.

$$\mathcal{L}_{\rm SM}^{\rm eff} = \mathcal{L}_{\rm SM} + \mathcal{L}_{\rm axion}^{\rm int}\left[\frac{\partial_\mu a_{\rm phys}}{f} \ ; \ \psi\right] - \frac{1}{2}\partial_\mu a_{\rm phys}\partial^\mu a_{\rm phys}$$
$$- \frac{1}{2}m_a^2 a_{\rm phys}^2 + \frac{a_{\rm phys}}{f}\xi\frac{\alpha_3}{8\pi}G_a^{\mu\nu}\tilde{G}_{a\mu\nu} \ . \tag{215}$$

---

[16] This would be true if $\mathcal{L}_{\rm SM}^{\rm eff}$ only contained interactions involving $\partial_\mu a$, since these cannot fix a value for the VEV of $a, \langle a\rangle$.

As is obvious from the above equation, the physics of axions depends on the scale of $U(1)_{\rm PQ}$ breaking f. In the original model Helen Quinn and I put forth (Peccei and Quinn 1977), we associated f quite naturally with the scale of electroweak symmetry breaking $v = (\sqrt{2}\,G_F)^{-1/2}$. To impose the $U(1)_{\rm PQ}$ symmetry on the Standard Model we had to have two distinct Higgs doublets, Φ_1 and Φ_2, with different $U(1)_{\rm PQ}$ charges. The axion field then turns out to be the common phase field of Φ_1 and Φ_2 which is orthogonal to the weak hypercharge (Peccei 1989). Isolating just this contribution in Φ_1 and Φ_2, one has

$$\Phi_1 = \frac{v_1}{\sqrt{2}} \exp\left[ix\frac{a}{f}\right]\begin{pmatrix}1\\0\end{pmatrix} \;;\;\; \Phi_2 = \frac{v_2}{\sqrt{2}}\exp\left[i\frac{a}{xf}\right]\begin{pmatrix}0\\1\end{pmatrix}\;. \tag{216}$$

Here $x = v_2/v_1$, is the ratio of the two Higgs VEV's and the $U(1)_{\rm PQ}$ symmetry breaking scale f is given by

$$f = \sqrt{v_1^2 + v_2^2} = (\sqrt{2}\,G_F)^{-1/2} \simeq 250\ {\rm GeV}\;. \tag{217}$$

The Φ_1 field has weak hypercharge of $-1/2$, while the Φ_2 field has weak hypercharge of $+1/2$. Hence, in the Yukawa interactions Φ_1 couples the $u_{{\rm R}j}$ fields to the left-handed quark doublets, while Φ_2 couples $d_{{\rm R}j}$ to these same fields

$$\mathcal{L}_{\rm Yukawa} = -\Gamma_{ij}^u(\bar{u},\bar{d})_{{\rm L}i}\Phi_1 u_{{\rm R}j} - \Gamma_{ij}^d(\bar{u},\bar{d})_{{\rm L}i}\Phi_2 d_{{\rm R}j} + {\rm h.c.} \tag{218}$$

In view of Eq. (216), it is clear that the above interaction is $U(1)_{\rm PQ}$ invariant. The shift of the axion field by αf [cf Eq. (207)] under a $U(1)_{\rm PQ}$ transformation is compensated by an appropriate rotation of the right-handed quark fields. Specifically, under a $U(1)_{\rm PQ}$ transformation one has

$$a_{\rm phys} \xrightarrow{PQ} a_{\rm phys} + \alpha f$$
$$u_{{\rm R}j} \xrightarrow{PQ} \exp\left[-i\alpha x\right]u_{{\rm R}j}$$
$$d_{{\rm R}j} \xrightarrow{PQ} \exp\left[-i\frac{\alpha}{x}\right]d_{{\rm R}j}\;. \tag{219}$$

It is clear from the above that this $U(1)_{\rm PQ}$ transformation encompasses also a $U(1)_A$ transformation. As a result, one can use $U(1)_{\rm PQ}$ to send $\bar{\theta} \to 0$, as advertized.

Unfortunately, weak interaction scale axions [with $f \sim 250$ GeV; $m_a \sim 100$ keV] of the type which ensue in the model suggested by Helen Quinn and myself, or in variations thereof, have been ruled out experimentally. I do not want to review all the relevant data here, as this is done already fully elsewhere (Peccei 1989). An example, however, will give a sense of the strength of this assertion. If weak scale axions were to exist, one expects a rather sizable branching ratio for the decay $K^{\pm} \to \pi^{\pm} a$ (Bardeen et al 1987)

$$BR(K^{\pm} \to \pi^{\pm} a) \sim 3 \times 10^{-5}\;. \tag{220}$$

Experimentally, however, the process $K^+ \to \pi^+$ "Nothing", which would reflect the axion decay of the K^+ meson, has a bound roughly three orders of magnitude lower (Asano et al 1981)

$$BR(K^+ \to \pi^+ + \text{ Nothing}) < 3.8 \times 10^{-8} \ . \tag{221}$$

One can bypass this bound by modifying the $U(1)_{\text{PQ}}$ properties of the Higgs fields involved. However, these variant model themselves run into other experimental troubles (Peccei 1989).

Although weak scale axions do not exist, it is still possible that the strong CP problem is solved because of the existence of a $U(1)_{\text{PQ}}$ symmetry. The dynamical adjustment of $\bar\theta \to 0$ works **independently** of what is the scale, f, of the spontaneous symmetry breaking of $U(1)_{\text{PQ}}$. Obviously if $f \gg (\sqrt{2}\, G_F)^{-1/2}$, the resulting axions are extremely light ($m_a \sim \Lambda_{\text{QCD}}^2/f$), extremely weakly coupled (couplings $\sim f^{-1}$) and very long lived ($\tau_a \sim f^5$) and thus are essentially **invisible**. A variety of invisible axion models have been suggested in the literature (Kim et al 1979) and they offer an interesting, if perhaps unconventional, resolution of the strong CP problem. Fortunately, as we shall see, these models are actually testable.

If $f \gg (\sqrt{2}\, G_F)^{-1/2}$, it is clear that the spontaneous breakdown of $U(1)_{\text{PQ}}$ must occur through a VEV of a field which is an $SU(2) \times U(1)$ singlet. Thus, in invisible axion models, the axion is essentially the phase associated with an $SU(2) \times U(1)$ singlet field σ.[17] Keeping only the axion degrees of freedom, one has

$$\sigma = \frac{f}{\sqrt{2}} e^{ia/f} \ . \tag{222}$$

It turns out that astrophysics and cosmology give important constraints on the $U(1)_{\text{PQ}}$ breaking scale f, or equivalently the axion mass (Peccei 1989)

$$m_a \simeq 6 \left[\frac{10^6 \text{ GeV}}{f} \right] \text{ eV} \ . \tag{223}$$

These constraints restrict the available parameter space for invisible axion models and suggest ways in which these excitations, if they exist, could be detected. Let me briefly discuss these matters.

The astrophysical bounds on axions arise because, if f is not large enough, axion emission removes energy from stars, altering their evolution. These bounds are reviewed in great details in a recent monograph by Raffelt (Raffelt 1996). Although these bounds are somewhat dependent on the type of invisible axion model one is considering, typically invisible axions avoid all astrophysical constraints if

$$f \geq 5 \times 10^9 \text{ GeV} \ ; \ \ m_a \leq 10^{-3} \text{ eV} \ . \tag{224}$$

Cosmology, on the other hand, provides an upper bound on f (Preskill et al 1983). At the $U(1)_{\text{PQ}}$ phase transition in the early Universe, at temperatures

---

[17] The field σ need not necessarily be an elementary scalar field (Kim 1979).

$T \sim f$, the effects of the QCD anomaly are not yet felt and the axion vacuum expectation value $\langle a \rangle$ is not alligned dynamically to cancel the $\bar{\theta}$ term. This cancellation only occurs as the Universe cools towards temperatures T of order $T \sim \Lambda_{\mathrm{QCD}}$. The axion VEV $\langle a \rangle$, as the temperature decreases, is driven to the correct minimum in an oscillatory fashion. These coherent, zero momentum, axion oscillations contribute to the Universe's energy density. If f is too large, in fact, the energy density due to axions can overclose the Universe. Demanding that this not happen gives a bound (Preskill et al 1983):

$$f \leq 10^{12} \text{ GeV} ; \;\; m_a \geq 6 \times 10^{-6} \text{ eV} . \tag{225}$$

This bound has some uncertainties, related to cosmology (for a discussion see, for example, Peccei 1996), but otherwise is not very dependent on the properties of the invisible axions themselves.

If axions contribute substantially to the Universe's energy density, the value of f (or m_a) will be close to the above bound. If this is the case, axions could be the source for the dark matter in the Universe. Remarkably, then, it may be actually possible, experimentally, to detect signals for these invisible axions. The basic idea, due to Sikivie (Sikivie 1983), is to try to convert axions, trapped in the galactic halo, into photons in a laboratory magnetic field.

If invisible axions constitute the dark matter of our galactic halo, they would have a velocity typical of the virial velocity in the galaxy, $v_a \sim 10^{-3} c$. Further, as the dominant components of the energy density of the Universe, axions would have a typical energy density in the halo of order

$$\rho_a^{\mathrm{halo}} \sim 5 \times 10^{-25} \text{ g/cm}^3 \sim 300 \text{ MeV/cm}^3 . \tag{226}$$

As a result of the (electromagnetic) anomaly, axions have an interaction with the electromagnetic field given by the effective Lagrangian (Peccei 1989)

$$\mathcal{L}_{a\gamma\gamma}^{\mathrm{eff}} = \frac{\alpha}{\pi f} K_{a\gamma\gamma} a \mathbf{E} \cdot \mathbf{B} . \tag{227}$$

Here $K_{a\gamma\gamma}$ is a model dependent parameter of $O(1)$. As a result of the above interaction, in the presence of an external magnetic field a galactic axion can convert into a photon.

Specifically, the electric field produced by an axion of energy $E_a \simeq m_a$ in the presence of a magnetic field $\mathbf{B}_0$ can be deduced from the modified wave equation

$$\left(\boldsymbol{\nabla}^2 - \frac{\partial^2}{\partial t^2} \right) \mathbf{E} = \frac{\alpha}{\pi f} K_{a\gamma\gamma} \mathbf{B}_o \frac{\partial^2 a}{\partial t^2} . \tag{228}$$

Experimentally, the generated electromagnetic energy can be detected by means of a resonant cavity. When the cavity is tuned to the axion frequency $w_a \simeq m_a$, one should get a narrow line on top of the noise spectrum. On resonance, the axion to photon conversion power is given by the expression (Sikivie 1983)

$$P_{\mathrm{axion}} = \frac{\rho_a}{m_a} \cdot V B_o^2 \cdot \left[\frac{\alpha}{\pi f} K_{a\gamma\gamma} \right]^2 C_{\mathrm{overlap}} \cdot Q_{\mathrm{eff}} . \tag{229}$$
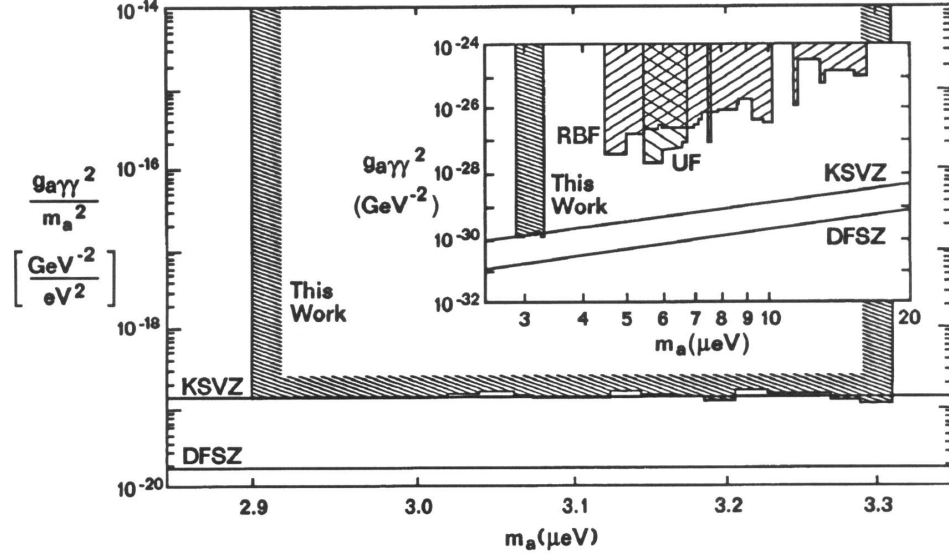
Fig. 2. Result of the Livermore experiment, along with limits from some previous axion searches

In the above, the first factor gives the expected number of axions per unit volume, the second details the magnetic energy stored in the cavity, the third contains the coupling strength squared, $g_{a\gamma\gamma}^2 = \left[\frac{\alpha}{\pi f} K_{a\gamma\gamma}\right]^2$. Finally, $C_{\text{overlap}} \simeq 0.7$ is an effectiveness factor for the cavity and Q_{eff} is the least value between the Q of the cavity itself $[Q \sim 10^6]$ and the Q due to the energy spread in the spectrum of halo axion, $Q_a \simeq [v_a^2/c^2]^{-1} \sim 10^6$.

Halo axions produce microwave photons, since 4×10^{-6} eV $\equiv 1$ GHz. Two pilot experiments carried out in the late 80's (de Panfilis et al 1987) had limited magnetic energy $[VB_o^2 \simeq 0.5 \text{ m}^3 \text{ (Tesla)}^2]$ and relatively noisy amplifiers. These experiments set limits for $g_{a\gamma\gamma}^2$ about 2 to 3 orders of magnitude above the theoretical expectations. Presently, there are two second generation experiments underway, one at the Lawrence Livermore National Laboratory and the other in Kyoto. The Livermore experiment uses a very large $VB_o^2 \sim 12 \text{ m}^3 \text{ (Tesla)}^2$ and low noise "state-of-the-art" amplifiers. Although the signal expected at 1 GHz is tiny, $P_{\text{axion}} \sim 5 \times 10^{-22}$ Watts, this experiment has already excluded a set of invisible axion masses, at the level of strength expected theoretically. These recent results (Hagmann et al 1998), along with some of the older data are shown in Fig. 2. The Kyoto experiment (Matsuki et al 1991) uses a moderate $VB_o^2 \sim 0.2 \text{ m}^3 \text{ (Tesla)}^2$. However, it utilizes an extremely clever technique for counting the number of photons converted from axions—using Rydberg atoms—

which makes up for the small VB_o^2. The Kyoto experiment is presently in a testing phase. One hopes that when both the Livermore and Kyoto experiments are completed, in 3-5 years time, they will have settled the important question of whether axions exist or not.

2.5   Do Real Nambu-Goldstone Bosons Exist?

We have known for almost 40 years that when a global symmetry group G breaks down spontaneously to a subgroup H $[G \rightarrow H]$, dim G/H massless Nambu-Goldstone bosons (Nambu 1980 and Goldstone 1981) appear in the spectrum of the theory. However, we have no **real** physical examples still of this phenomena. To be fair, pions are an excellent example of states which are **nearly** Nambu-Goldstone bosons. However, although there is no question that pions are the Nambu-Goldstone excitations associated with the breakdown of the $SU(2)_V \times SU(2)_A$ approximate global symmetry of QCD to $SU(2)_V$, pions have a small mass since the u and d quarks are not exactly massless.

For a while, it was believed that it was impossible for real physical Nambu-Goldstone bosons to exist in nature. The argument was simple. Because these particles are massless, their existence seemed to be precluded by the fact that the only long-range forces we know in nature are gravity and electromagnetism. However, in the early 1980's it was realized that the existence of $m = 0$ Nambu-Goldstone bosons does not pose a contradiction, so that one can actually contemplate the interesting possibility that such states may actually exist.

This idea came up first as a result of studying the possibility that lepton number may be spontaneously violated. Chikashige, Mohapatra, and I (Chikashige, Mohapatra and Peccei 1981a) dubbed the Nambu-Goldstone boson associated with the spontaneous breakdown of lepton number a Majoron. Soon thereafter, others (Wilczek 1982) considered theories where one had a global family number which could also be spontaneously broken, resulting in other types of real Nambu-Goldstone bosons, given the name of Familons.

In this subsection I want to explain briefly why, in general, real Nambu-Goldstone bosons, are not dangerous excitations to have in a theory. After having done so, I then want to discuss briefly a specific type of Majoron model, to illustrate some of the consequences of these kind of models. Succintly, the reason why Nambu-Goldstone bosons do not run afoul with present limits on possible additional long-range forces is due to a little theorem of Gelmini, Nussinov, and Yanagida (Gelmini, Nussinov and Yanagida 1983), which shows that the exchange of Nambu-Goldstone bosons leads only to a long-range tensor force.

The proof of the Gelmini-Nussinov-Yanagida theorem is very simple. One is interested in the potential produced by the exchange of a Nambu-Goldstone boson between two fermions. Recall that Nambu-Goldstone boson fields, π, always shift under a broken symmetry transformation [cf Eq. (207) for the axion]. Therefore, one has

$$\pi(x) \xrightarrow{\xi} \pi(x) + v_\pi \xi \; . \tag{230}$$

Here ξ is a parameter in G/H and v_π is a scale parameter associated with the symmetry breakdown in question. As a result of Eq. (230), clearly Nambu-Goldstone fields must always be derivatively coupled. Hence, the most general coupling of a Nambu-Goldstone boson π to two fermions f_1 and f_2 takes the form

$$\mathcal{L}_{\mathrm{NGB}}^{\mathrm{fermion}} = i \frac{\partial_\mu \pi}{v_\pi} \bar{f}_1 [a\gamma_\mu + b\gamma_\mu \gamma_5] f_2 + \mathrm{h.c.} \ , \tag{231}$$

where a and b are numerical coefficients. If one uses the fermion equations of motion, one can reduce the above to a more useful form involving the π field directly

$$\mathcal{L}_{\mathrm{NGB}}^{\mathrm{fermion}} = \frac{\pi}{v_\pi} \bar{f}_1 [a(m_1 - m_2) + b(m_1 + m_2)\gamma_5] f_2 + \mathrm{h.c.} \ , \tag{232}$$

where m_1 and m_2 are the masses of the fermion fields f_1 and f_2, respectively.

In calculating the potential due to π-exchange between two fermions one needs, at each vertex, to use the interaction Lagrangian above with $f_1 = f_2$. Obviously, for two equal fermions, the effective coupling of a Nambu-Goldstone boson is always a **pseudoscalar** coupling. Thus Nambu-Goldstone boson exchange cannot really generate coherent long-range forces, since a pseudoscalar coupling in the non-relativistic limit reduces to a $\boldsymbol{\sigma} \cdot \mathbf{p}$ coupling. More precisely, the effective diagonal coupling of a Nambu-Goldstone boson, π, to a fermion, f, is given by

$$\mathcal{L}_{\mathrm{NGB}}^{\mathrm{diag}} = ig_\pi \frac{m_f}{v_\pi} \bar{f}\gamma_5 f \pi \ , \tag{233}$$

where g_π is a, dimensionless, coupling constant. In the non-relativistic limit, the above reduces to

$$\mathcal{L}_{\mathrm{NGB}}^{\mathrm{diag}} \to g_\pi \chi_f^* \frac{\boldsymbol{\sigma} \cdot \boldsymbol{\nabla}}{v_\pi} \chi_f \pi \ , \tag{234}$$

where χ_f is a Pauli spinor. Such an interaction gives an exchange potential between two fermions which is spin-dependent and tensorial, with an $1/r^3$ not an $1/r$ fall off

$$V_{\mathrm{NGB-exchange}}^{\mathrm{eff}} = \frac{g_\pi^2/4\pi}{v_\pi^2} \left\{ \frac{\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 - 3(\boldsymbol{\sigma}_1 \cdot \hat{r})(\boldsymbol{\sigma}_2 \cdot \hat{r})}{r^3} + \frac{4\pi}{3}\delta^3(r)\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 \right\} \ . \tag{235}$$

There have been analyses in the literature (Feinberg and Sucher 1979) of the size of possible non-magnetic dipole-dipole interactions in matter, precisely of the type one would obtain from the exchange of a real Nambu-Goldstone boson. These bounds effectively limit how small the scale v_π can be. One finds no contradiction with experiment (Chikashige, Mohapatra and Peccei 1981a) provided that

$$\frac{v_\pi}{g_\pi} \geq \mathrm{TeV} \ . \tag{236}$$

Thus, one can contemplate having real Nambu-Goldstone bosons of global symmetries which are broken down at scales not much bigger than the weak scale! If

v_π/g_π is much above the bound (236), clearly one expects no measurable effects in matter. Furthermore, if v_π/g_π is large, these Nambu-Goldstone bosons are also hard to directly produce, since the effective coupling for producing them from a fermion f scales like $g_\pi m_f/v_\pi$.

2.6   Majorons.

I want to illustrate the above discussion by briefly considering the simplest example of spontaneously broken Lepton number and its associated Majoron. As we discussed earlier, Lepton number is a classical global symmetry of the Standard Model. Even at the quantum level, because the $\nu \neq 0$ amplitudes are highly suppressed, this remains an almost exact symmetry. However, there is no reason why Lepton number should remain a symmetry of the theory, once one considers extensions of the Standard Model. Indeed, the simplest extension of the Standard Model introduces right-handed neutrino fields $\nu_{\mathrm{R}i}$ for each family. Because these fields are $SU(2) \times U(1)$ singlets, one can write a Majorana (fermion-fermion) mass term involving these fields of the form

$$\mathcal{L}_{\mathrm{mass}} = -\frac{(M_{\mathrm{R}})_{ij}}{2}\nu_{\mathrm{R}i}^T C \nu_{\mathrm{R}j} + \text{h.c.} , \tag{237}$$

with C being the charge conjugation matrix introduced earlier ($C = 1$ in the Majorana representation). Obviously $\mathcal{L}_{\mathrm{mass}}$ does not respect Lepton number, since its two terms carry Lepton number +2 and -2, respectively.

One can restore Lepton number as a symmetry in the above example by introducing an appropriately transforming Higgs field. In this case, what one needs is a complex $SU(2) \times U(1)$ singlet field σ, which carries Lepton number -2 (Chikashige, Mohapatra and Peccei 1981a). Clearly the interaction Lagrangian

$$\mathcal{L}_{\mathrm{CMP}} = -\frac{h_{ij}}{\sqrt{2}}\left[\nu_{\mathrm{R}i}^T C \nu_{\mathrm{R}j}\sigma + \text{h.c.}\right] \tag{238}$$

is L invariant by construction. If the dynamics of the theory forces σ to acquire a VEV, $\langle\sigma\rangle = \frac{1}{\sqrt{2}}V$, then the above Lagrangian reproduces the effect of having an explicit Majorana mass term for the right-handed neutrino fields. In this case, one has

$$(M_{\mathrm{R}})_{ij} = h_{ij}V , \tag{239}$$

and Lepton number is spontaneously broken. Hence this theory must also contain an explicit Nambu-Goldstone boson—the Majoron. This is the model which I first studied with Chikashige and Mohapatra (Chikashige, Mohapatra and Peccei 1981a).

As was the case for the axion, the Majoron can also be identified here as the phase field associated with the complex field σ. Focusing only on the Majoron, χ, degrees of freedom, one can write

$$\sigma \simeq \frac{V}{\sqrt{2}}e^{i\chi/V} . \tag{240}$$

If the interaction (238) was the only interaction that the $\nu_{\mathrm{R}i}$ fields had, then clearly χ would couple only to these fields. However, once one introduces right-handed neutrino fields, one cannot avoid coupling $\nu_{\mathrm{R}i}$ to the usual leptonic doublet fields $(\nu, e)_{\mathrm{L}i}$ via the ordinary Higgs doublet field. As a result of these couplings, the Majoron field χ also ends up by having a (small) interaction with the left-handed neutrino fields. However, if the right-handed Majorana mass M_{R} (or, equivalently, the VEV of the σ-field V) is large, the Majoron still predominantly couples to the right-handed neutrinos.

Let us see how this goes in detail. As a result of the spontaneous breaking of both Lepton number and $SU(2) \times U(1)$, the neutrino fields have both a Dirac (fermion-antifermion) and a Majorana mass term:

$$\mathcal{L}_{\mathrm{mass}} = -\frac{1}{2}(M_{\mathrm{R}})_{ij}\left[\nu_{\mathrm{R}i}^{T}\nu_{\mathrm{R}j}\right] - (M_D)_{ij}[\bar{\nu}_{\mathrm{L}i}\nu_{\mathrm{R}j}] + \mathrm{h.c.} , \tag{241}$$

with the Dirac mass matrix M_D being proportional to the doublet Higgs VEV.[18] If the eigenvalues of M_R are much greater than those of M_D, then the neutrino mass matrix

$$\mathcal{M} = \begin{pmatrix} 0 & M_{\mathrm{D}} \\ M_{\mathrm{D}} & M_{\mathrm{R}} \end{pmatrix} \tag{242}$$

has a set of large eigenvalues, corresponding to the eigenvalues of M_R, and a set of extremely small eigenvalues, associated with the matrix M_D^2/M_R. This is the famous see-saw mechanism (Yanagida et al 1979). As a result, one ends up with a spectrum of neutrinos with both superheavy states and superlight states:

$$\mathcal{L}_{\mathrm{mass}} \simeq -\frac{1}{2}\left[\frac{M_{\mathrm{D}}^2}{M_{\mathrm{R}}}\right]_{ij} \bar{\eta}_{1i}\eta_{1j} - \frac{1}{2}[M_{\mathrm{R}}]_{ij}\bar{\eta}_{2i}\eta_{2j} . \tag{243}$$

The light neutrinos η_{1i} are mostly left-handed, while the heavy neutrinos η_{2i} are mostly right-handed.

The mass mixing discussed above, has a counterpart in the interactions of the Majoron. Although the field χ mostly couples to the heavy fields η_{2i}, there will also be a small coupling of χ to η_{1i}. That is, the Majoron χ as a result of the neutrino mass mixing actually has also a small coupling to the ordinary left-handed neutrinos. Specifically (Chikashige, Mohapatra and Peccei 1981a), one finds

$$\mathcal{L}_{\mathrm{Majoron}}^{\mathrm{int}} = -\frac{h_{ij}}{2}\bar{\eta}_{2i}i\gamma_5\eta_{2j}\chi - \left(\frac{hM_{\mathrm{D}}^2}{M_{\mathrm{R}}^2}\right)_{ij}\bar{\eta}_{1i}i\gamma_5\eta_{2j}\chi . \tag{244}$$

It follows that the Majoron coupling to the light neutrinos is of order $M_{\mathrm{D}}^2/M_{\mathrm{R}}^2 \sim m_\nu/M_{\mathrm{R}}$, where m_ν is the mass (matrix) for the light neutrinos. The Majoron has an even weaker coupling to ordinary matter, which is induced at one-loop

---

[18] Naively, one would expect M_D to be similar to the mass matrix M_ℓ for the charged leptons.

order via mixing of the χ with the Z^o. One finds (Chikashige, Mohapatra and Peccei 1981a)

$$\mathcal{L}_{\text{matter}}^{\text{eff}} = i\frac{m_f}{v_\chi}\bar{f}\gamma_5 f\chi \tag{245}$$

with the scale v_χ of order $v_\chi \sim (G_F m_\nu)^{-1} \gg$ TeV. So clearly the Majoron in this model easily satisfy the constraints imposed on additional dipole-dipole interactions in matter (Feinberg and Sucher 1979).

If Majorons exist, it is possible for the heaviest of the light neutrinos to decay into the other neutrinos by Majoron emission. The process $\nu_i \rightarrow \nu_j\chi$, if it were fast enough, would serve to open up a region of neutrino masses forbidden by cosmology. For stable neutrinos, one knows that neutrinos in the mass range from a few eV to a few GeV (Kolb and Turner 1990) overclose the Universe. However, these bounds cease to apply for unstable neutrinos. If the lifetime τ for the neutrino decay $\nu_i \rightarrow \nu_j\chi$ is much shorter than the Universe's lifetime T_o, then effectively one can redshift the ν_i energy beyond its mass m_{ν_i}. Hence the contribution of these neutrinos to the energy density of the Universe is reduced to (Chikashige, Mohapatra and Peccei 1981b)

$$\rho_{\nu_i} \sim m_{\nu_i}\left[\frac{\tau}{T_o}\right]^{1/2} T_\nu^3 \tag{246}$$

where T_ν is the neutrino temperature now, $T_\nu \sim T_\gamma \sim 3^o$ K.

The lifetime τ for the process $\nu_i \rightarrow \nu_j\chi$ naively scales as (Chikashige, Mohapatra and Peccei 1981b)

$$\tau(\nu_i \rightarrow \nu_j\chi) \sim \frac{1}{m_{\nu_i}}\left[\frac{M_{\text{R}}}{M_{\text{D}}}\right]^4 \tag{247}$$

and can be made short enough if M_{R} is not too large. However, in the simplest Majoron model discussed here (Chikashige, Mohapatra and Peccei 1981a) this lifetime is lengthened by a further factor of $\left[\frac{M_{\text{R}}}{M_{\text{D}}}\right]^4$ (Schechter and Valle 1984), making it very doubtful that $\tau < T_o$. More elaborate models (Gelmini and Roulet 1995) restore the simple formula (247) and the possibility that, through Majoron decay, neutrinos with masses in the "forbidden" cosmological range could exist. This is not entirely an academic exercise, as the existing bounds on m_{ν_μ} and m_{ν_τ} [$m_{\nu_\mu} \leq 170$ keV; $m_{\nu_\tau} < 24$ MeV (Particle Data Group 1996)] allow these particles to have masses precisely in this range.

2.7   Global Symmetries and Gravity.

In the preceding subsections I have discussed various interesting global symmetries, which may be associated with the interactions of the Standard Model, and have explored a bit the consequences of these symmetries. There are, however, some arguments one can adduce from the analysis of gravitational interactions

which bring into question the whole notion of having theories with exact global symmetries. I want to end my lectures by discussing this point briefly.

Perhaps the simplest way to see why gravitational interactions may cause trouble is to focus on the "No Hair" theorem for black holes. Basically this theorem (see, for example, Banks 1990) asserts that black holes can be characterized only by a few fundamental quantities, like mass and spin, but possess otherwise no other quantum numbers. Because black holes can absorb particles which carry global charge, while carrying no global charge themselves, it appears that through these processes one can get an explicit violation of whatever symmetry is associated with the global charge. That is, global charge can be lost when particles carrying this charge are swallowed by a black hole.

One can parametrize the effect of the breaking of global symmetries by gravitational interactions by adding to the low-energy Lagrangian non-renormalizable terms, scaled by inverse powers of the Planck mass $M_{\mathrm{P}} \sim 10^{19}$ GeV. These terms, of course, should be constructed so as to explicitly violate the symmetries in question. Schematically, therefore, the full Lagrangian of the theory, besides containing the usual Standard Model terms, should also include some effective non-renormalizable interactions containing various operators O_n, breaking explicitly the Standard Model global symmetries:

$$\mathcal{L}_{\mathrm{grav.\ int.}}^{\mathrm{eff}} = \sum_m \frac{1}{M_{\mathrm{P}}^n}\, O_n \; . \tag{248}$$

Here the dimension of the operators O_n, which explicitly breaks some of the Standard Model global symmetries, is $n + 4$.

Let me make two remarks. First, the Lagrangian (248) can often be augmented by other effective interactions which themselves break certain global symmetries even more strongly than gravity. For instance, an explicit mass term for the right-handed neutrinos [cf Eq. (237)] does break L directly and more strongly than the operators in Eq. (248) do. This said, however, in what follows I will concentrate only on the gravitational effects embodied in Eq. (248).

Because $M_{\mathrm{P}} \sim 10^{19}$ GeV $\gg (\sqrt{2}\, G_F)^{-1/2} \sim 250$ GeV, the naive expectation is that Eq. (248) cannot be that important, except at superheavy scales. This turns out to be true for the interactions themselves, but fails when one considers the effect of Eq. (248) on the Nambu-Goldstone sector. To demonstrate the first point, let me consider the example of (B+L)-violation. The dominant, $d = 6$, (B+L)-violating interaction induced by gravity schematically has the form (Weinberg, Wilczek and Zee 1979)

$$\mathcal{L}_{(B+L)-violation} \sim \frac{1}{(M_{\mathrm{P}})^2} u_i^c d_j u_k^c e f_{ijk} \; . \tag{249}$$

Such a term leads to a proton lifetime, for the process $p \to e^+ \pi^0$, of order

$$\tau(p \to e^+ \pi^0) \sim (M_{\mathrm{P}})^4 \sim 10^{46} \text{ years} \; , \tag{250}$$

much greater than the present experimental bound on this process discussed earlier [Eq. (155)]. So the breaking of B+L provided through gravitational effects is indeed irrelevant.

The situation is, however, different when one considers the Nambu-Goldstone sector. Let us consider again the simple example of spontaneously broken Lepton number with its associated Majoron. To the Lepton number conserving potential, which forces the $SU(2) \times U(1)$ singlet field σ to acquire a VEV, one must now add non-renormalizable Lepton number violating terms induced by the gravitational interactions. The simplest such term involves a dimension 5 operator. Thus, one is invited to study the potential

$$V_{\text{total}} = \lambda \left(\sigma^\dagger \sigma - \frac{V^2}{2} \right)^2 - \frac{\lambda'}{M_{\text{P}}} (\sigma^\dagger \sigma)^2 [\sigma + \sigma^\dagger] \ . \tag{251}$$

The first term above is clearly invariant under the Lepton number transformation $\sigma \to e^{-2ia}\sigma$. This is not so for the term which scales as M_{P}^{-1}. Writing, as before,

$$\sigma \simeq \frac{V}{\sqrt{2}} \exp \left[i \frac{\chi}{V} \right] \ , \tag{252}$$

one sees that the effect of including the gravitational corrections is to produce a mass term for the erstwhile Nambu-Goldstone field χ. One finds, for the Majoron, a mass

$$m_\chi^2 = \frac{\lambda'}{2\sqrt{2}} V^2 \left(\frac{V}{M} \right) \ . \tag{253}$$

Note that the size of the Majoron mass depends on the value of V, the scale of the spontaneous breakdown of Lepton number. For example, if we took $V \sim$ TeV—the lowest it can be according to the bound of Eq. (236)—then $m_\chi \sim V(V/M_{\text{P}})^{1/2} \sim 10^{-8} \ V \simeq 10$ KeV. If V is larger, the mass of the Majoron grows as

$$m_\chi \sim 10 \left[\frac{V}{\text{TeV}} \right]^{3/2} \text{KeV} \ . \tag{254}$$

Clearly, if the Majoron is massive, some of its physical properties are altered substantially. For instance, it could happen that the decay $\nu_i \to \nu_j \chi$ is actually kinematically forbidden! Of course, the above results are predicated on the assumption that the global Lepton number symmetry is violated explicitly by a dim 5 interaction. If the violation were due to a higher dimensional operator of dimension d, then one finds for the Majoron mass the formula

$$m_\chi \sim V \left(\frac{V}{M_{\text{P}}} \right)^{\frac{d-4}{2}} \tag{255}$$

which leads to masses which become smaller the larger d is.

These considerations are particularly troubling for the $U(1)_{\text{PQ}}$ solution to the strong CP problem (Holman et al 1992). Not only potentially do gravitational effects give an additional contribution to the axion mass, but they can also alter

the QCD potential so that $\bar{\theta}$ does not finally adjust to zero! One can understand what is going on by schematically sketching the form of the effective axion potential in the absence and in the presence of the $U(1)_{\mathrm{PQ}}$ breaking gravitational interactions (Barr and Seckel 1992). Without gravity, a useful parametrization for the physical axion effective potential, which follows from examining the contributions of instantons (Peccei and Quinn 1977), is

$$V_{\mathrm{axion}} = -\Lambda_{\mathrm{QCD}}^4 \cos a_{\mathrm{phys}}/f \ . \tag{256}$$

This potential displays the necessary periodicity in a_{phys}/f, has a minimum at $\langle a_{\mathrm{phys}} \rangle = \bar{\theta}_{\mathrm{eff}} = 0$, and leads to an axion mass $m_a = \Lambda_{\mathrm{QCD}}^2/f$.

Including gravitational effects changes the above potential by adding a sequence of terms involving operators of different dimensions. Let us just consider one such term and examine the potential (Barr and Seckel 1992)

$$\tilde{V}_{\mathrm{axion}} = -\Lambda_{\mathrm{QCD}}^4 \cos \frac{a_{\mathrm{phys}}}{f} - \frac{cf^d}{M_{\mathrm{P}}^{d-4}} \cos \left[\frac{a_{\mathrm{phys}}}{f} - \delta \right] \ . \tag{257}$$

Here c is some dimensionless constant and δ is a CP-violating phase which enters through the gravitational interactions. This potential modifies the formula for the axion mass, giving now

$$m_a^2 \simeq \frac{\Lambda_{\mathrm{QCD}}^4}{f^2} + c \frac{f^{d-2}}{M_{\mathrm{P}}^{d-4}} \ . \tag{258}$$

For f in the range of interest for invisible axions, the second term above coming from the gravitational effects dominates the QCD mass estimate for the axion, unless c is extraordinarily small and/or the dimension d is rather large. More troublesome still, $\tilde{V}_{\mathrm{axion}}$ now no larger has a minimum at $\langle a_{\mathrm{phys}} \rangle = 0$. Rather one finds a minimum of $\tilde{V}_{\mathrm{axion}}$ for values of

$$\bar{\theta}_{\mathrm{eff}} = \frac{\langle a_{\mathrm{phys}} \rangle}{f} \simeq c \sin \delta \frac{f^d}{M_{\mathrm{P}}^{d-4} \Lambda_{\mathrm{QCD}}^4} \ . \tag{259}$$

That is, the gravitational effects (provided there is a CP violating phase associated with them) induce a non-zero $\bar{\theta}$, even in the presence of a $U(1)_{\mathrm{PQ}}$ symmetry! To satisfy the bound $\bar{\theta} \leq 10^{-10}$ again necessitates that d be large and/or that the constant c be extraordinarily small.

To date there is no clear resolution to this problem and it could be that these considerations actually vitiate the chiral solution to the strong CP problem. Since this is the most appealing solution to this conundrum, this is somewhat troubling. Nevertheless, it is worth noting a number of points. First, one does not really understand quantum gravity. Thus it is possible that when matters are better understood the effective global symmetry breaking interactions we introduced may in fact not be there at all, or be tremendously suppressed. Second, there are some encouraging results in this direction coming from string theory. Axions associated with broken chiral symmetries arise very naturally in string

theory (Witten 1984). Furthermore, CP is conserved, at least in higher dimensions in string theory (Choi, Kaplan and Nelson 1993), so perhaps it is possible that $\sin\delta = 0$. Finally, there are arguments that for large compactification radii, the effective $U(1)_{PQ}$ symmetries are broken very little in strings, so that the tiny number needed for c $[c \le 10^{-51}]$ may not be out of the question (see, for example, Choi 1997).

Irrespective of the above considerations, one should note that if the gravitational effects induce values of $\bar{\theta} < 10^{-10}$, so that the strong CP problem is still solved by imposing a $U(1)_{PQ}$ symmetry, then also the axion mass is approximately given by its QCD form. Thus, perhaps the best way to resolve these thorny theoretical questions is to find experimental evidence for the existence of invisible axions, with the canonical properties!

Acknowledgments

References

Adler, S.(1970): in **Lectures on Elementary Particles and Quantum Field Theory**, 1970 Brandeis Summer Institute, eds. S. Deser, M. Grisaru and H. Pendelton (MIT Press, Cambridge, MA).
Adler, S. (1969): Phys. Rev. **177**, 2426; Bell, J. S., Jackiw, R. (1969): Nuovo Cimento **60**, 47.
Asano, Y., *et al.* (1981): Phys. Lett. **107B**, 159.
Baluni, V. (1979): Phys. Rev. D**19**, 2227.
Banks, T. (1990): Physicalia **12**, 19.
Bardeen, W. A. (1974): Nucl. Phys. B**75**, 246.
Bardeen, W. A., Peccei, R. D., Yanagida, T. (1987): Nucl. Phys. B**279**, 401.
Barr S. M., Seckel, D. (1992): Phys. Rev. D**46**, 539.
Belavin, A. A., Polyakov, A. M., Schwartz, A. A., Tyupkin, Y. S. (1975): Phys. Lett. B**59**, 85.
Buchanan, C. D., Cousins, R., Dib, C. O., Peccei, R. D., Quackenbush, J. (1992): Phys. Rev. D**45**, 4088; Dib, C. O., Peccei, R. D. (1992): Phys. Rev. D**46**, 2265.
Cabibbo, N. (1963): Phys. Rev. Lett. **12**, 531; Kobayashi, M., Maskawa, T. (1973): Prog. Theor. Phys. **49**, 652.
Callan, C. G., Dashen, R., Gross, D. (1976): Phys. Lett. **63B**, 172.
Chikashige, Y., Mohapatra, R. N., Peccei, R. D. (1981a): Phys. Lett. **96B**, 265.
Chikashige, Y., Mohapatra, R. N., Peccei, R. D. (1981b): Phys. Rev. Lett. **45**, 1926.
Choi, K. W. (1997): Phys. Rev. D**56**. 6588.
Choi, K. W., Kaplan D. B., Nelson, A. E. (1993): Nucl. Phys. B**391**, 515.
Christenson, J. H., Cronin, J. W., Fitch, V. L., Turlay, R. (1964): Phys. Rev. Lett. **13**, 138.
Crewther, R. J. (1978):, Acta Phys. Austria Suppl. **19**, 47, Proceedings of the XVII Int. Universität Wochen für Kernphysik, Schladming, Austria, 1978.

Crewther, R., di Vecchia, P., Veneziano, G., Witten, E. (1979): Phys. Lett. **88B**, 123; (1980): **91B**, 487(E).

de Panfilis, S., *et al.* (1987): Phys. Rev. Lett. **59**, 839; Wuensch, W. U., *et al.* (1989): Phys. Rev. D**40**, 3153; Hagmann, C. *et al.* (1990): Phys. Rev. D**42**, 1297.

Donoghue, J. F., Golowich, E., Holstein, B. R. (1992): **Dynamics of the Standard Model** (Cambridge University Press, Cambridge, UK).

Feinberg, G., Sucher, J. (1979): Phys. Rev. D**20**, 1717.

Gelmini, G., Nussinov, S., Yanagida, T. (1983): Nucl. Phys. **B219**, 31.

Gelmini, G., Roulet, E. (1995): Rept. Prog. Phys. **58**, 1207.

Georgi, H., Kaplan, D. B., Randall, L. (1986): Phys. Lett. **169B**, 73.

Hagmann, C. *et al.* (1998): Phys. Rev. Lett. **80**, 2043.

Heisenberg, W., Z. (1932): Phys. **77**, 1; see also Cassen, B., Condon, E. U. (1936): Phys. Rev. **60**, 846; Wigner, E. P. (1937): Phys. Rev. **51**, 106.

Holman R., *et al.* (1992): Phys. Lett. B**282**, 132; Kamionkowski, M., March Russel, J. (1992): Phys. Lett. B**282**, 137; Barr, S. M., Seckel, D. (1992): Phys. Rev. D**46**, 539.

Jackiw. R., Rebbi, C. (1976): Phys. Rev. Lett. **37**, 2.

Kim, J. (1979): Phys. Rev. Lett. **43**, 103.

Kim, J. (1979): Phys. Rev. Lett. **43**, 103; Shifman, M. A., Vainshtein, A. I., Zakharov, V. I. (1980): Nucl. Phys. B**166**, 453; Dine, M., Fischler, W., and Srednicki, M. (1981): Phys. Lett. **104B**, 199: Zhitnisky, A. P. (1980): Sov. Jour. Nucl. Phys. **31**, 260.

Kolb, E. W., Turner, M. (1990): **The Early Universe** (Addison Wesley, Redwood City, California).

Krasnikov, N., Rubakov, V., Tokarev, V. (1978): Phys. Lett. **79B**, 423; see also, Anselm A. A., Johansen, A. A. (1994): Nucl. Phys. B**412**, 553.

Kuzmin, V., Rubakov, V., Shaposhnikov, M. (1985): Phys. Lett. **55B**, 36.

Lee, T. D., Yang, C. N. (1956): Phys. Rev. **104**, 254.

Low, F. E.(1967): **Symmetries and Elementary Particles** (Gordon and Breach, New York).

Matsuki, S., Yamamoto, K. (1991): Phys. Lett. **263B**, 523; Ogawa, I., Matsuki, S., Yamamoto, K. (1996): Phys. Rev. D**53**, R1740.

Nambu, Y. (1960): Phys. Rev. Lett. **4**, 380; Goldstone, J. (1961): Nuovo Cimento **19**, 1425; see also, Goldstone, J., Salam, A., Weinberg, S. (1962): Phys. Rev. **127**, 965.

Pal, P. B., Wolfenstein, L. (1982): Phys. Rev. D**25**, 788.

Particle Data Group, Barnett, R. M. *et al.* (1996): Phys. Rev. D**54**, 1.

Pauli, W. (1955): in **Niels Bohr and the Development of Physics**, ed. W. Pauli (Pergamon Press, New York); Schwinger, J. (1951): Phys. Rev. **82**, 914; Lüders, G. (1954): Dansk Mat. Fys. Medd **28**, 5; Lüders, G., Zumino, B. (1958): Phys. Rev. **110**, 1450.

Peccei, R. D. (1987): in **Concepts and Trends in Particle Physics**, eds. H. Latal and H. Mutter, in Proceedings of the XXV Int. Universität Wochen für Kernphysik, Schladming, Austria, 1986 (Springer-Verlag, Berlin).

Peccei, R. D. (1989): in **CP Violation**, ed. C. Jarlskog (World Scientific, Singapore).

Peccei, R. D. (1996): Jour. Korean Phys. Soc. (Proc. Suppl.) **29**, S199.

Peccei, R. D., Quinn, H. R. (1977): Phys. Rev. Lett. **38**, 1440; Phys. Rev. D**16**, 1791.

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

Raffelt, G. (1996): **Stars as Laboratories for Fundamental Physics** (University of Chicago Press, Chicago).

Schechter J., Valle, J. (1984): Phys. Rev. D**25**, 774.

Sikivie, P. (1983): Phys. Rev. Lett. **51**, 1415; (1985): Phys. Rev. D**32**, 2988; Krauss, L. *et al.* (1985): Phys. Rev. Lett. **55**, 1797.

Slansky, R. (1981): Phys. Reports **79C**, 1.

Streater, R. F., Wightman, A. S.(1964): **PCT, Spin and Statistics, and All That**, (W. A. Benjamin, New York).

't Hooft, G. (1976a): Phys. Rev. Lett. **37**, 8.

't Hooft, G. (1976b): Phys. Rev. Lett. **37**, 8; Phys. Rev. D**14**, 3432; Polyakov, A. (1977): Nucl. Phys. B**120**, 429; for a review, see, for example, Schäfer, T., and Shuryak, E. V. (1998): Rev. Mod. Phys. **70**, 323.

't Hooft, G. (1976c): Phys. Rev. D**14**, 3432.

Weinberg, S. (1972): Phys. Rev. Lett. **29**, 1698.

Weinberg, S. (1975): Phys. Rev. D**11**, 3583.

Weinberg, S. (1978): Phys. Rev. Lett. **40**, 223; Wilczek, F. (1978): Phys. Rev. Lett. **40**, 271.

Weinberg, S. (1979): Phys. Rev. Lett. **43**, 1556; Wilczek, F., Zee, A. (1979): Phys. Rev. Lett. **43**, 1571.

Wigner, E. P. (1932): Nach. der Gess. Wiss. Göttingen **32**, 35; see also, **Group Theory and its Applications to the Quantum Mechanics of Atomic Spectra** (Academic Press, New York, 1954).

Wigner, E. P. (1952): Proc. Nat. Acad. Sci. (US) **38**, 449; Weyl, H. (1929): Z. Phys. **56**, 330.

Wilczek, F. (1982): Phys. Rev. Lett. **49**, 1549; Reiss, D. B. (1982): Phys. Lett. **115B**, 217.

Witten, E. (1984): Phys. Lett. **149B**, 551.

Wu, C. S. *et al.* (1957): Phys. Rev. **105**, 1413; **106**, 1361; Friedman, J. I., Telegdi, V. L. (1957): Phys. Rev. **105**, 1681; Garwin, R. L., Lederman, L. M., Weinrich, M. (1957): Phys. Rev. **105**, 1425.

Yanagida, T. (1979): in Proceedings of the Workshop on Unified Theory and the Baryon Number of the Universe, KEK, Japan ; Gell-Mann, M., Ramond, P., Slansky, R. (1979): in **Supergravity**, ed. by P. Van Neuwenhuisen (North Holland, Amsterdam).