PROCEEDINGS OF SCIENCE

The strong CP puzzle and axions

Christopher Smith[a,*]

[a] *Laboratoire de Physique Subatomique et de Cosmologie,*
Université Grenoble-Alpes, CNRS/IN2P3, Grenoble INP, 38000 Grenoble, France

E-mail: chsmith@lpsc.in2p3.fr

In the first part of this talk, after a brief presentation of the strong CP puzzle, the construction of axion models and their main phenomenological features are described. In the second part, the possibility to mix the Peccei-Quinn symmetry with baryon and lepton numbers is discussed, showing that the axion could ultimately play a role in other puzzles of the Standard Model like the smallness of neutrino masses or baryogenesis.

31st International Workshop on Deep Inelastic Scattering (DIS2024)
8–12 April 2024
Grenoble, France

*Speaker

PoS(DIS2024)013

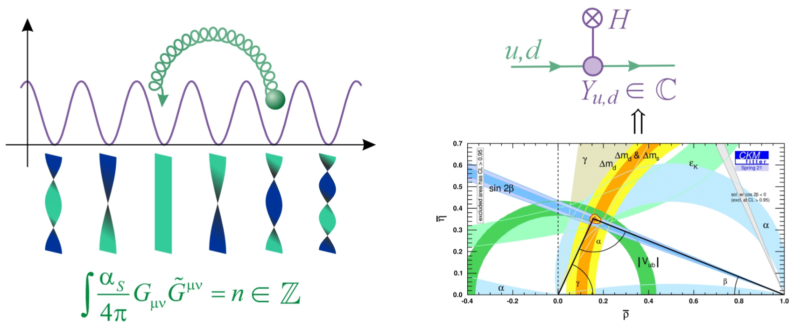

Figure 1: Illustration of the two sources for θ in the SM. In this picture, a naive way to see that the QCD topology breaks T, hence CP, is to imagine a gluon emitted from the vacuum at time t_1 and absorbed at a later time t_2. The time-reversed process is different since emitting back the gluon at t_2, there is no guarantee to end up in the same vacuum at t_1.

1. Brief introduction to the strong CP puzzle

The Standard Model (SM) is defined as containing all the possible renormalizable couplings compatible with its gauge symmetries and matter content. This should include the CP-violating

$$\mathcal{L}_{SM} \supset \frac{\alpha_S}{4\pi}\theta \, \mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu} \;, \tag{1}$$

where $\tilde{G}^{\mu\nu} = \varepsilon^{\mu\nu\rho\sigma}G_{\rho\sigma}/2$ is the dual of the gluon field strength tensor, and θ an additional free parameter. Such a coupling induces a CP-violating term to the neutron coupling to photons, $\langle\gamma n|\,\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}|n\rangle$, encoded at leading order in the non-relativistic expansion into its electric dipole moment (EDM). Experimentally, no such EDM has been found [1], and the current limit $\theta < 10^{-10}$ appears at odd with naturality. Yet, the problem is actually far more serious because, in the SM, two unrelated mechanisms contribute to θ.

First, there is a strong interaction contribution of topological origin, θ_{glue}. That it must be topological is understood from the peculiar nature of the coupling of Eq. (1): it is a total derivative. Thus, only boundary terms could make it non-zero. Usually, fields are required to vanish asymptotically, but in the presence of a gauge symmetry, one should rather ask them to be pure gauge at infinity. This is where the topology of the color group $SU(3)$ comes into play: vacuum gauge configurations fall into equivalence classes according to how many times they wind over the sphere at infinity, see Fig. 1. This would be rather innocuous but for the existence of field configurations, the instantons, tunneling between these gauge configurations. Thus, the true vacuum of QCD is a specific superposition, labelled by θ_{glue}, of these topologically distinct vacua. Importantly, these mathematical considerations are not pure speculations but do have experimental support in the mass of the η'. This state can mix with $\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$, now interpreted as a glueball made heavy by the presence of instantons [2]. Numerically, this is measured on the lattice [3] via the topological susceptibility $\chi \sim \langle 0|(\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu})(\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu})|0\rangle$. Thus, a priori, θ_{glue} should be an $O(1)$ parameter.

A second contribution to θ comes from the Higgs Yukawa coupling to quarks, $\mathbf{Y}_{u,d}$. Those are complex matrices, hence after symmetry breaking, anomalous rephasing of the quark fields are required to make the quark masses real, generating $\theta_{flav} = \arg\det \mathbf{Y}_u + \arg\det \mathbf{Y}_d$. Here also, we have experimental support for $\mathbf{Y}_{u,d}$ being complex since weak CP violation from the CKM matrix is well-established [4], see Fig. 1. Thus, θ_{flav} is also a free $O(1)$ parameter.

Now, the non-observation of strong CP violation requires a near perfect cancellation, to one part in ten billion, between the QCD topological contribution and the Higgs Yukawa contributions, $\theta = \theta_{glue} + \theta_{flav} < 10^{-10}$. Arguing that somehow QCD has to be defined on that specific vacuum where CP is conserved is clearly not sufficient since this only sets θ_{glue} to zero. As the Higgs and QCD sectors are entirely separated in the SM, with the Higgs boson not even colored, the situation is way more puzzling than having simply the Lagrangian parameter θ accidentally small. This mysterious cancellation is so puzzling that significant departures from the SM are envisioned, with the axion solution being the prime candidate. Yet, before turning to that, let us briefly discuss some alternative solutions that have appeared along the years.

Solution 1: Massless quark

Historically, the first solution to the strong CP puzzle is also the simplest. If one quark flavor is massless, say the up quark, then its right and left components can be rephased independently. But any mismatch between the rephasing of these two chiralities generates a shift in θ thanks to the chiral anomaly. Thus, θ ceases to be physical and can be rotated to zero.

Nowadays, chiral perturbation theory combined with lattice QCD forbid this solution, with the smallest quark mass m_u definitely greater than zero [5]. Yet, it should be remarked that what matters to solve the strong CP puzzle is the absence of a Lagrangian mass term for the up quark, whose physical mass also receives instanton contributions [6]. Though QCD alone cannot suffice, if one can concoct a model in which some instanton contributions do saturate the measured up-quark mass, then QCD becomes CP conserving again (see e.g. Ref. [7] and references there).

Solution 2: GUT paradigm

Per se, Grand Unified Theories (GUT) do not solve the strong CP puzzle, but they may alleviate it somehow. First, some GUT theories can force Yukawa matrices to be hermitian. This is typically the case when a left-right symmetry is active [8]. At the GUT scale, this ensures $\theta_{flav} = 0$. Of course, since CP is violated in the electroweak sector, this cannot be the end of the story [9]. Finite contributions from the CKM phase are tiny though, below 10^{-16}. Infinite contributions, arising first at seven loops, make θ a running parameter. Those sum up to even tinier corrections below 10^{-18}. All that remains thus is the issue of θ_{glue}. Inspired by how the electroweak θ term gets rotated away in the SM thanks to the residual but anomalous $U(1)$ symmetry corresponding to baryon plus lepton numbers, $\mathcal{B} + \mathcal{L}$, a similar mechanism may be invoked [10]. To our knowledge, a full model with all these ingredients, able to account for the unification of the couplings and realistic fermion masses has not been proposed yet.

Solution 3: Infinities

The reality of the strong CP puzzle rests on how to deal with no less than three infinities. First, the impact of $\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$ vanishes at any finite order of perturbation theory. Second, being a total derivative, it is the boundary at infinity that matters. Third, the vacuum $|\theta\rangle$ is an infinite sum over the topologically distinct sectors $|k\rangle$, i.e., $|\theta\rangle = \sum_k e^{ik\theta}|k\rangle$.

Messing up with how these infinities are treated individually or in conjunction with one another can drastically change the outcome. The main constraint here is to make sure $\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$ does contribute to the η' mass. Another difficulty to design a paradigm within QCD to drive θ to zero is that QCD with a finite θ probably does not look much like the QCD we know, since it may not even be confining [11]. Thus, phenomenology does not appear to behave smoothly as a function of θ. Finally, lattice simulations are not always of great help. While the topological susceptibility, and thereby the η' mass, are well reproduced, the impact of the θ term is more elusive, with e.g. the neutron EDM matrix element being still compatible with zero [12].

Recently, the focus has been on the order of the infinite limits [13] for the volume over which correlation functions are defined and the summation over k. Sketchily, the idea is that if the volume is sent to infinity before the sum over k, then each k region becomes infinitely large, and all our observables are forced to live in a given vacuum, see Fig. 1. In our opinion, whether such a non-standard reversal in the order of the limits is permitted is not established though. So, for the time being, our point of view will be that the strong CP puzzle appears quite resilient, but given the complexity of the QCD dynamics, one cannot rule out the possibility that it resolves itself spontaneously in the future.

2. Axions to solve the strong CP puzzle

The main ingredient of the axion solution to the strong CP puzzle, along with its phenomenology, can be illustrated starting from:

$$\mathcal{L} \supset \frac{\alpha_S}{4\pi}\theta\, \mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu} + (y\phi\bar{\psi}_L\psi_R + h.c.) + \partial_\mu\phi^\dagger\partial^\mu\phi - V(\phi^\dagger\phi)\ . \tag{2}$$

The complex scalar field ϕ does not belong to the SM, though the colored fermion ψ may.

The Peccei-Quinn step:

The first step is to enforce the $U(1)_{PQ}$ symmetry generated by $\phi \to \exp(i\theta)\phi$ by giving different charges to ψ_L and ψ_R. The associated Noether current then has an anomaly, $\partial_\mu J^\mu_{PQ} \sim \mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$. At the same time, $V(\phi^\dagger\phi)$ is assumed to break $U(1)_{PQ}$ spontaneously, so the scalar field can be represented as

$$\phi(x) = (f_a + \rho(x))\exp(ia(x)/f_a)\ , \tag{3}$$

where f_a is the vacuum expectation value and the axion a is the Goldstone boson whose shifts span the circular vacuum. By Goldstone theorem, a is coupled to $\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$ since the current satisfies $\langle 0|J^\mu_{PQ}|a(p)\rangle = if_a p^\mu$. The ρ field has a mass of $O(f_a)$, so integrating it out, the Lagrangian at low energy is generically

$$\mathcal{L} \supset \frac{1}{2}\partial_\mu a\partial^\mu a + \frac{\alpha_S}{4\pi}\left(\theta + \frac{a}{f_a}\right)\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu} + \frac{1}{f_a}\partial_\mu a J^\mu_{PQ} + \dots\ . \tag{4}$$

The invisibility step:

Whether it is via $\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$, itself coupled to η', or via SM fermions present in J^μ_{PQ}, the axion is coupled to normal matter. Its non-observation immediately requires f_a to be well above the electroweak scale v_{EW}. There are then essentially two paradigms to make the axion "invisible".

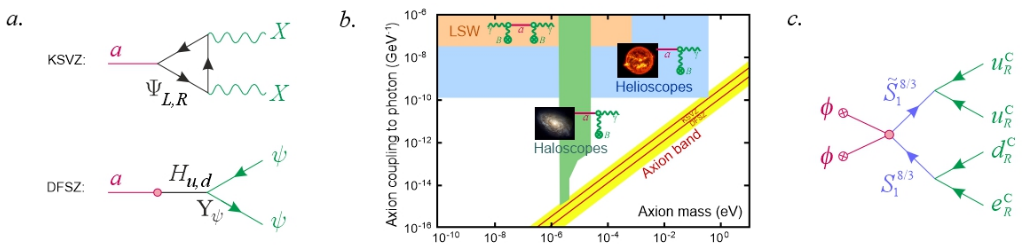

Figure 2: a. Leading axion couplings to SM particles in the KSVZ and DFSZ scenarios. b. Schematic representation of the axion exclusion plot, with the axion band from the theoretical correlation between $g_{a\gamma\gamma}$ and m_a. The main search strategies are light-shining through the wall-type experiments (LSW), helioscopes using the sun as an axion source, or haloscopes aiming for the DM axions of the galactic halo. c. Spontaneous proton decay in the model of Eq. (12).

The KSVZ scenario [14] introduces both ϕ and $\psi_{L,R}$ as new degrees of freedom. To preserve the renormalizability of the SM, $\psi_{L,R}$ are then vectorlike under all the SM symmetries. Generically, the Lagrangian at the electroweak scale is then (see Fig. 2)

$$\mathcal{L} \supset \frac{a}{16\pi^2 f_a} \sum_{X=G,W,B} g_X^2 N_X X_{\mu\nu} \tilde{X}^{\mu\nu} , \tag{5}$$

with N_X functions of the SM gauge charges of ψ.

The other scenario starts from a two-Higgs doublet model (THDM) [15]. For a type II fermionic sector, and provided the scalar potential leaves room for the $U(1)_{PQ}$ symmetry, the electroweak symmetry breaking makes the pseudoscalar Higgs boson massless [16], so $a = A^0$ but $f_a \sim v_{EW}$. To make an invisible axion, the DFSZ idea [17] is to add the complex scalar field of Eq. (3) together with a coupling $\phi^2 H_u H_d$, such that the physical axion field becomes $a \to a + (v_{ew}/f_a)A^0$. All the couplings of a to SM fields come from that of A^0, see Fig. 2, hence to leading order,

$$\mathcal{L} \supset \frac{a}{f_a} \sum_{\psi=u,d,e} \chi_\psi m_\psi \bar{\psi}\gamma^5\psi , \tag{6}$$

where χ_ψ depend on the vacuum expectation values of $H_{u,d}$. Of course, at the loop level, both the KSVZ and DFSZ include couplings to fermions and gauge bosons.

The hadronic step:

No matter the precise implementation, the outcome is a massless field coupled to $\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$, and possibly also to $F_{\mu\nu}\tilde{F}^{\mu\nu}$:

$$\mathcal{L} \supset \frac{\alpha_S}{4\pi}\left(\theta + \frac{a}{f_a}\right) \mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu} + \frac{1}{4} g_{a\gamma\gamma}^{SD} a F_{\mu\nu}\tilde{F}^{\mu\nu} + \ldots . \tag{7}$$

The θ term then necessarily relaxes to zero. Indeed, once QCD hadronization sets in, the circular vacuum space of the potential gets tilted, and a falls to a specific vacuum which happens to be that where CP is conserved. Technically, $\phi\, \mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$ becomes a potential $V(\phi)$, whose minimum is at $\langle\phi\rangle = 0$. A recurrent difficulty here is that this tilting is not that strong, so the PQ symmetry needs to be near perfect. This is the axion quality problem: even Planck-scale tilting of the potential could jeopardize the final drive to the CP-conserving solution [18].

There are other important consequences (see e.g. Ref. [19] for a review). First, a becomes massive since the vacuum space is no longer flat, with $f_a m_a \approx f_\pi m_\pi$. Second, a mixes with the light neutral mesons π^0, η, and η' since $\operatorname{tr} G_{\mu\nu}\tilde{G}^{\mu\nu}$ mixes with the singlet η_0. This immediately pushes f_a to very large values, e.g. from $K \to \pi(\pi \to a)$. Third, even if in a specific model, $g_{a\gamma\gamma}^{SD} = 0$, the axion inherits a $\gamma\gamma$ coupling $g_{a\gamma\gamma}^{LD}$ from that of π^0, η, and η'. Barring fine-tuning, $g_{a\gamma\gamma} = g_{a\gamma\gamma}^{SD} + g_{a\gamma\gamma}^{LD}$ is then correlated to the mass, $g_{a\gamma\gamma}/m_a \approx 10^{-1\pm1}$, see Fig. 2. Fourth, the axion is sufficiently long-lived to become a cold dark matter (DM) candidate. If the axion field is initially misaligned with the CP-conserving vacuum, its falling to it corresponds to collective oscillations whose zero-mode energy density can match the observed dark matter abundance. With m_a related to the topological susceptibility, and from lattice estimates of the temperature dependence of the latter, $10 - 1000\ \mu eV$ emerges as a preferred m_a range for $O(1)$ initial misalignment angles.

The discovery step?

All the ingredients are now in place. The axion has its characteristic coupling to $F_{\mu\nu}\tilde{F}^{\mu\nu} = E\cdot B$, correlated to its mass. Search strategies exploit this coupling mostly as an axion-photon conversion in an intense B field, and their constraints are then represented in the $g_{a\gamma\gamma}$ vs. m_a plane, see Fig. 2. There are many such experiments currently going on or planned in the near future (including a local initiative in Grenoble [20]), along with many efforts aiming at the axion couplings to leptons or nucleons (see [21] for an up-to-date list and corresponding exclusion plots). It should be said also that Fig. 2 represents the 'vanilla' scenarios, but many alternatives exist, with models out of the axion band (using larger representations, clockwork, some Z_n symmetry, mirror QCD,..., see e.g. Refs. [22]), or out of the preferred DM mass range (with some kinetic energy [23] or topological defects [24]).

3. Baryonic axions

Introducing the axion to solve the strong CP puzzle is a rather high price to pay, not least because one is forced to add also new fermions (like in KSVZ), scalars (like in DFSZ), or even a whole new dark sector. Its capability to explain the observed dark matter relic density strengthen its case, but in a spirit of minimality, it would be desirable to further its phenomenological role. This is where the interplay with baryon and lepton numbers, $\mathcal{B}$ and $\mathcal{L}$, could prove fruitful. Indeed, at the root of the axion mechanism is the PQ symmetry, under which some colored fermions must transform non-trivially. As such, this symmetry is nothing more than a flavor $U(1)$ symmetry, exactly like $\mathcal{B}$ and $\mathcal{L}$. In practice, these symmetries naturally mix, as we will see. This opens the way for the axion to play a role in the generation of neutrino masses and leptogenesis, or even directly in baryogenesis. Given the accidental closeness of the DM and baryonic relic densities, having a model relating both would certainly look promising. This is still a long way off. In the present section, the goal is to summarize some recent advances made incorporating $\mathcal{B}$ and $\mathcal{L}$ violation within axion models.

Reparametrization freedom

Before embarking into model building, it is crucial to realize that axion models can be defined in very different though equivalent forms. This is well-known in the context of Chiral Perturbation

Theory, but not always fully appreciated for axions. Specifically, starting from the renormalizable Lagrangian of Eq. (2), the usual linear representation is

$$\phi \sim f_a + \rho + ia : \mathcal{L}_{linear} = \frac{\alpha_S \theta}{4\pi} \operatorname{tr} G_{\mu\nu}\tilde{G}^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi - im\frac{a}{f_a}\bar{\psi}\gamma^5\psi + \dots . \tag{8}$$

The Lagrangian stays renormalizable after symmetry breaking, with the axion having pseudoscalar couplings to fermions. Importantly, no process is anomalous in this description, even at loop level. This is evident in the original Peccei-Quinn model since it is simply matched onto the THDM.

In view of the topology of the vacuum (i.e., a circle), one can instead adopt the polar representation for ϕ. The Lagrangian is then no longer manifestly renormalizable

$$\phi \sim (f_a + \rho)\exp(ia/f_a) : \mathcal{L}_{polar} = \frac{\alpha_S \theta}{4\pi} \operatorname{tr} G_{\mu\nu}\tilde{G}^{\mu\nu} + \bar{\psi}(i\not{D} - m\exp(ia\gamma^5/f_a))\psi + \dots . \tag{9}$$

The interest of that representation is that ρ decouples and can trivially be integrated out. Also, it shows explicitly how the Goldstone boson makes the fermion mass term compatible with the chiral $U(1)_{PQ}$ symmetry.

In the previous representation, a $U(1)_{PQ}$ transformation changes $\theta \to \theta + \alpha$ by first shifting $a \to a + f_a\alpha$, which then requires a compensating chiral, hence anomalous, rotation of the fermion field $\psi \to \exp(i\alpha/2)\psi$. It is possible to skip this last step by adopting a reparametrization of the fermion field to make it neutral under $U(1)_{PQ}$ (compare with Eq. (4)):

$$\psi \to \exp(-ia\gamma^5/(2f_a))\psi : \mathcal{L}_{derivative} = \frac{\alpha_S}{4\pi}\left(\theta + \frac{a}{v}\right) \operatorname{tr} G_{\mu\nu}\tilde{G}^{\mu\nu} + \bar{\psi}\left(i\not{D} - m + \frac{\partial_\mu a\gamma^\mu\gamma^5}{f_a}\right)\psi + \dots . \tag{10}$$

This is the derivative representation, in which the axion appears essentially as a dynamical θ term, but with derivative interactions otherwise, as appropriate for a Goldstone boson.

Though Feynman rules, diagrams, symmetry properties, and non-relativistic expansions [25] are different in the three representations $\mathcal{L}_{linear}$, $\mathcal{L}_{polar}$, and $\mathcal{L}_{derivative}$, observables have to be the same whatever the chosen form. In most applications, the derivative representation is taken as a starting point. For most low-energy applications, this is perfectly fine, but there are issues when $\mathcal{B}$ and/or $\mathcal{L}$ violation is present, or when weak interactions are considered. Let us discuss these points in turns.

Electroweak couplings

To illustrate the issue with the weak interaction [26], let us adopt in a simplified setting with only a single fermion, as in Eq. (2). Starting from a Yukawa term like $y\phi\bar{\psi}_L\psi_R$, the PQ charges are only defined up to a common constant. Setting $Q_\phi = 1$, we must have $Q_{\psi_L} = \alpha$ and $Q_{\psi_R} = \alpha - 1$. The free parameter α reflects the existence of another $U(1)$ flavor symmetry, corresponding to the conservation of the ψ-number. Now, if ψ is charged under some gauge interaction X, the effective axion Lagrangian can be constructed as

$$\begin{aligned}\mathcal{L}_{derivative} \supset\ & \frac{a}{f_a}\frac{g_X^2}{16\pi^2}(Q_{\psi_L}C_2^{\psi_L} - Q_{\psi_R}C_2^{\psi_R}) \operatorname{tr} X_{\mu\nu}\tilde{X}^{\mu\nu} \\ & + \frac{\partial_\mu a}{f_a}\left((Q_{\psi_R} + Q_{\psi_L})\bar{\psi}\gamma^\mu\psi + (Q_{\psi_R} - Q_{\psi_L})\bar{\psi}\gamma^\mu\gamma^5\psi\right) .\end{aligned} \tag{11}$$

For QCD and QED, the Casimir invariant $C_2^{\psi_L} = C_2^{\psi_R}$, hence α drops out of the anomalous term. It also disappears from the axial current, but remains in the vector one. However, that current being conserved, it cannot contribute to observables. Hence, the ψ-number ambiguity is irrelevant.

The situation changes for a chiral interaction, for which $C_2^{\psi_L} \neq C_2^{\psi_R}$. For example, if ψ_L is a weak doublet, but ψ_R a singlet, then $aW_{\mu\nu}\tilde{W}^{\mu\nu}$ is entirely tuned by the free parameter α. This reflects the fact that the $W_{\mu\nu}\tilde{W}^{\mu\nu}$ term comes entirely from the anomaly in the ψ-number symmetry current. The $aW_{\mu\nu}\tilde{W}^{\mu\nu}$ coupling then depends on how much of that current is present in the PQ current. One could think of setting $\alpha = 0$ to somewhat remove the ψ-number current, but this totally kills the $aW_{\mu\nu}\tilde{W}^{\mu\nu}$ coupling, which is not consistent either since clearly, $a \to W^+W^-$ or $a \to ZZ$ can occur if ψ_L has weak interactions, see Fig. 2.

Actually, this dependence on α is spurious, because the local anomalous terms $aW_{\mu\nu}\tilde{W}^{\mu\nu}$ and $aB_{\mu\nu}\tilde{B}^{\mu\nu}$ in $\mathcal{L}_{derivative}$ are spurious. The reason is that in the presence of chiral gauge interactions, neither the axial nor the vector current are conserved: the triangle graphs involving both these currents are anomalous. Crucially, the anomalous terms from these triangle graphs exactly cancel with the local anomalous terms. This is expected since we could instead work in the linear representation, in which no anomaly ever arises. In other words, anomalies are an artifact of the derivative representations. At the end of the day, the axion couplings thus come from the non-anomalous part of the triangle graphs, and precisely match the result one would get starting from the pseudoscalar interaction in either $\mathcal{L}_{linear}$ or $\mathcal{L}_{polar}$. What makes the situation a bit peculiar for QED and QCD is that this non-anomalous remainder is *parametrically* identical, in the $m \to \infty$ limit, to the anomalous term (this is the Sutherland-Veltman theorem). But this does not work for chiral gauge bosons.

Merging the PQ symmetry with $\mathcal{B}$ and $\mathcal{L}$

In realistic models, beside the ψ-number discussed above for KSVZ-type models, there are two additional ambiguities due to $\mathcal{B}$ and $\mathcal{L}$, since those are always conserved by the SM Yukawa couplings. It is important to keep track of these ambiguities whenever $\mathcal{B}$ and/or $\mathcal{L}$ violating interactions are turned on [27].

Let us give a simple example. We add to the KSVZ model the coupling $\phi\bar{\nu}_R^{\rm C}\nu_R + h.c.$. This merges lepton number with the PQ charges, because one is forced to set not only the PQ charge of ν_R to $-1/2$, but also that of the lepton weak doublet $\ell_L = (\nu_{eL}, e_L)$ and singlet e_R because of $\bar{e}_R Y_e \ell_L$ and $\bar{\nu}_R Y_\nu \ell_L$. Notice that if one constructs $\mathcal{L}_{derivative}$ using these charges, spurious couplings of the axion to charged leptons appear, showing once again how plagued by cancellations is that representation. More importantly, had we frozen those PQ charges to zero, not accounting for the ambiguity due to $\mathcal{L}$, one may have wrongly concluded that the $\phi\bar{\nu}_R^{\rm C}\nu_R$ coupling is incompatible with the PQ symmetry. This is obviously wrong since nothing forbids it in the linear representation. What is true though is that whenever ϕ is coupled to normal matter, it is impossible to incorporate a PQ symmetry if more than two different combinations of $\mathcal{B}$ and $\mathcal{L}$ are broken. For example, if $\phi\bar{\nu}_R^{\rm C}\nu_R$ is included to merge the seesaw and axion mechanisms, then only a single other type of $\mathcal{B}$ and/or $\mathcal{L}$ violation is possible before the axion becomes massive. In this context, one should also pay attention to electroweak instantons, generating the effective $(\ell_L q_L^3)^3$ interaction, that can induce axion quality problems [27].

To force ϕ to carry other combinations of $\mathcal{B}$ and $\mathcal{L}$ is not so easy though. Clearly, a direct coupling to three quarks (plus a lepton) would not be renormalizable. The strategy systematically analyzed in Ref. [28] is to do this via various types of leptoquarks and diquarks. Then, one can merge the PQ symmetry with any two combinations of $\mathcal{B}$ and $\mathcal{L}$ allowed by the gauge symmetries. For example, if one adds two weak singlet scalar states of hypercharge 8/3 and couplings

$$\mathcal{L} = \mathcal{L}_{KSVZ} + \phi\bar{\nu}_R^C\nu_R + S_1^{8/3}\bar{d}_R e_R^C + \tilde{S}_1^{8/3}\bar{u}_R^C u_R + \phi^2 S_1^{8/3\dagger}\tilde{S}_1^{8/3} + h.c. , \tag{12}$$

then a single global $U(1)$ merging PQ, $\mathcal{B}$ and $\mathcal{L}$ remains. This symmetry prevents any other couplings, and when ϕ breaks it spontaneously, neutrinos become massive and proton decay arises, see Fig. 2. Even though these models are hardly economical, they illustrate how varied the phenomenological possibilities are, with features like spontaneous proton decay, neutron-antineutron oscillations, neutrino masses and neutrinoless double-beta decay, a possibility to account for the neutron lifetime puzzle (for an ALP with a mass conveniently set right in-between that of the proton and neutron), or with axion-induced neutron-antineutron transitions.

4. Conclusion

Currently, the axion remains our best mechanism to solve the strong CP puzzle. Further, it is becoming one of our best candidates as constituent of the elusive dark matter, as the absence of signals of new physics at colliders disfavors the weakly-interacting massive particle paradigm. Though from a model-building perspective, the axion mechanism is a rather simple and somewhat limited departure from the SM, what is remarkable is that it could also hold the key to other long-standing puzzles. For instance, the similarity in the PQ breaking and seesaw scale points towards a role of the axion in the smallness of neutrino masses, and that between baryonic and dark matter relic densities advocates for an axionic baryogenesis of some kind. With experimental efforts now closing in on the parameter space expected in the simplest scenarios, a discovery could be around the corner. This could revolutionize both particle physics and cosmology, and open the door to many further advances in our understanding of Nature.

References

[1] C. Abel *et al.* Phys. Rev. Lett. **124** (2020) no.8, 081803 [2001.11966].

[2] G. 't Hooft, Phys. Rev. Lett. **37** (1976) 8; Phys. Rev. D **14** (1976) 3432; E. Witten, Nucl. Phys. B **156** (1979) 269.

[3] G. S. Bali *et al.* [RQCD], JHEP **08** (2021) 137 [2106.05398].

[4] J. Charles *et al.* [CKMfitter Group], Eur. Phys. J. C **41** (2005) 1 [hep-ph/0406184], updated results and plots available at: http://ckmfitter.in2p3.fr.

[5] H. Leutwyler, PoS **CD09** (2009) 005 [0911.1416].

[6] H. Georgi and I. N. McArthur, HUTP-81/A011.

[7] M. Carena, *et al.* Phys. Rev. D **100** (2019) 9, 094018 [1904.05360].

[8] G. Senjanovic and V. Tello, Int. J. Mod. Phys. A **38** (2023) 15n16, 2350067 [2004.04036].

[9] J. R. Ellis and M. K. Gaillard, Nucl. Phys. B **150** (1979) 141; I. B. Khriplovich and A. I. Vainshtein, Nucl. Phys. B **414** (1994) 27 [hep-ph/9308334]. , see also C. Smith and S. Touati, Nucl. Phys. B **924** (2017) 417 [1707.06805], and references there.

[10] M. Shifman and A. Vainshtein, Mod. Phys. Lett. A **32** (2017) 14, 1750084 [1701.00467].

[11] G. 't Hooft, Nucl. Phys. B **190** (1981) 455.

[12] C. Alexandrou *et al.*, Phys. Rev. D **103** (2021) 5, 054501 [2011.01084]; J. Liang *et al.* [χQCD], Phys. Rev. D **108** (2023) 9, 094512 [2301.04331].

[13] W. Y. Ai *et al.*, Phys. Lett. B **822** (2021) 136616 [2001.07152].

[14] J. E. Kim, Phys. Rev. Lett. **43** (1979) 103; M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **166** (1980) 493.

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

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

[17] M. Dine, W. Fischler and M. Srednicki, Phys. Lett. B **104** (1981), 199-202; A. R. Zhitnitsky, Sov. J. Nucl. Phys. **31** (1980) 260.

[18] M. Kamionkowski and J. March-Russell, Phys. Lett. B **282** (1992) 137 [hep-th/9202003].

[19] L. Di Luzio *et al.*, Phys. Rept. **870** (2020) 1 [2003.01100].

[20] T. Grenet *et al.* [2110.14406].

[21] https://cajohare.github.io/AxionLimits/

[22] L. Di Luzio, F. Mescia and E. Nardi, Phys. Rev. Lett. **118** (2017) 3, 031801 [1610.07593]. T. Higaki *et al.*, Phys. Lett. B **755** (2016) 13 [1512.05295]. L. Di Luzio *et al.*, JHEP **05** (2021) 184 [2102.00012]. M. K. Gaillard *et al.* Eur. Phys. J. C **78** (2018) 11, 972 [1805.06465].

[23] R. T. Co, L. J. Hall and K. Harigaya, Phys. Rev. Lett. **124** (2020) 25, 251802 [1910.14152].

[24] D. J. E. Marsh, Phys. Rept. **643** (2016) 1 [1510.07633].

[25] C. Smith, Eur. Phys. J. C **84** (2024) 12 [2302.01142].

[26] J. Quevillon and C. Smith, Eur. Phys. J. C **79** (2019) 10, 822 [1903.12559]; J. Quevillon, C. Smith and P. N. H. Vuong, JHEP **08** (2022) 137 [2112.00553]; and for an application to axion-like particles, F. Arias-Aragón, J. Quevillon and C. Smith, JHEP **03** (2023) 134 [2211.04489].

[27] J. Quevillon and C. Smith, Phys. Rev. D **102** (2020) 7, 075031 [2006.06778]; Eur. Phys. J. Plus **137** (2022) 1, 141 [2010.13683].

[28] F. Arias-Aragón and C. Smith, Phys. Rev. D **106** (2022) 5, 055034 [2206.09810].