School of Physical and Chemical Sciences, University of Canterbury,
Christchurch, New Zealand

The Strong CP Problem and Axions

Joseph A. Wilson

Supervisor Assoc. Prof. Jenni Adams

«Compiled on 31/07/20, 12:41»

UNIVERSITY OF CANTERBURY

Astr430 Literature Review

Abstract

Literature review on something relating to axions.

Contents

1 QCD and the Strong CP Problem

THE STANDARD MODEL of particle physics is a quantum field theory which describes all known fundamental particles and interactions to very high precision—with notable exceptions including: the absence of gravity; dark matter; and the experimentally incorrect prediction that neutrinos are massless. Alongside these shortcomings, the standard model also exhibits issues of a more philosophical nature, such as the reasons for the values of the theory's many free parameters. (For example; why are there three generations of particles? Why do the particle masses and coupling constants have the values they do?) Among these mysteries is the *strong CP problem*, a fine-tuning problem regarding the apparent symmetry of the strong force under time reversal.[1] This chapter summarises the background theory necessary to understanding the statement and origin of the strong CP problem.

The standard model was born amid the explosion of phenomenological particle physics in the early 1960s, in an effort to explain the patterns in the rapidly increasing catalogue of known particles. The introduction of the quark model explained the properties of numerous hadrons in terms of six kinds of quark, and was given the name *quantum chromodynamics* (QCD). This, together with the well-established quantum theory of electromagnetism, *quantum electrodynamics* (QED), became the first approximation to the standard model [1].

The modern statement of the standard model is that it is a Lorentz-invariant *quantum field theory* whose *gauge group* is the compact connected Lie group

$$\big(\mathrm{SU}(3) \oplus \mathrm{SU}(2) \oplus \mathrm{U}(1)\big)\big/\mathbb{Z}_6, \tag{1.1}$$

which is equipped with a particular action on matter fields. A quantum field theory is a quantised *gauge field theory*, which is a field theory whose Lagrangian is symmetric under the action of the gauge group (elaborated upon in § 1.1). Each term in the gauge group of the standard model (1.1) corresponds to a fundamental force of nature: The factor $\mathrm{SU}(3)$ is the gauge group of quantum chromodynamics, the sector of the standard model describing the interactions of colour-charged quarks and gluons via the strong force. The factor $\mathrm{SU}(2) \oplus \mathrm{U}(1)$ corresponds to the unified electroweak force, and contains another $\mathrm{U}(1)$ sub-factor corresponding to the electromagnetic force of QED.

Quantum chromodynamics was formulated quite some time after quantum electrodynamics was understood, which is a reflection that QCD is more intricate than QED. Unlike QED, quantum chromodynamics exhibits *asymptotic freedom*, meaning that the effective strength of the strong interaction between colour-changed particles *increases* with increasing separation. Thus, QCD is severely non-perturbative and evades even modern analytical treatment, except

[1] Time reversal T and charge–parity CP symmetry are equivalent assuming the combined charge–parity–time symmetry; i.e., $T = CP \mod CPT$.

at low energies. Asymptotic freedom also results in quark confinement, which contributed to the confusion of particle physicists in the 1960s—quarks could never be observed in isolation. The dissimilarity of QCD to QED can be credited largely to the fact that the gauge group $\mathrm{SU}(3)$ is non-Abelian, corresponding to the fact that the gluon force carriers of QCD are themselves colour-charged, which results in asymptotic freedom.

However, quantum chromodynamics remains a highly successful and beautiful theory. Aside from the technical problems of its non-perturbative nature, QCD also exhibits a mystery of purely theoretical concern: the *strong CP problem.* This problem is of the fine-tuning variety, but it lacks even an anthropic solution and is good reason to pursue physics beyond the standard model.

The significance and role of the gauge group is outlined in the next section where a geometrical overview of gauge field theory is presented. This overview is intended to equip the reader with a geometrical idea of the kinds of objects which comprise classical gauge theories and which underpin the standard model. Then, QCD is defined and the origin of the so-called *strong CP problem* is highlighted.

«Justify focus on classical gauge theory»

1.1 Overview of Classical Gauge Theory

Gauge theories take place on a spacetime manifold and consist of two mathematically distinct dynamical entities: a *matter field* and a *gauge field.* A classical gauge theory is completely specified by the prescription of three ingredients: the base manifold; the vector space in which the matter field takes its values; and a *gauge group* equipped with an action on the matter field. The dynamical gauge field arises naturally as a consequence of the matter field's symmetry under the action of the gauge group.

The underlying premise of gauge theory is that there may be redundancy in the mathematical description of a matter field at each point in spacetime. (For example, the complex phase of a wavefunction.) This redundancy results in mathematical degrees of freedom of the matter fields which are non-physical, called *local gauge freedoms.* (Keeping with our example: local rotations of phase.) Crucially, local gauge freedoms introduce ambiguity in the notion of the physical rate of change of matter fields about a point, because of the point-wise independence of its gauge freedom. In other words, there is no preferred directional derivative of a matter field with local gauge freedom, until the choice of a *connection* is made. The triumph of gauge theory is that, by introducing the gauge field to act as a connection, this ambiguity is promoted to a separate set of physical degrees of freedom. The implications of this are twofold: Firstly, the gauge field yields a choice of derivative (namely, the covariant derivative with respect to the gauge field), allowing the inclusion of well-defined derivatives of matter fields in the theory's equations of motion. Secondly, the gauge field has a dynamical role in the theory, describing a new kind of field: force fields (and, after quantisation, the force carrier bosons).

In geometrical language, the matter field $\boldsymbol{\psi} \in \Gamma(\mathcal{E})$ is a section of a vector bundle $V \rightarrowtail \mathcal{E} \twoheadrightarrow \mathcal{M}$. In other words, at any point $p \in \mathcal{M}$ in spacetime the field assigns a vector, $\boldsymbol{\psi}|_p \cong V$.[2]

[2] The reason many objects in gauge theory (such as fields $\boldsymbol{\psi}$) are defined as sections of fibre bundles (and not simply as smooth maps $\boldsymbol{\psi} : \mathcal{M} \to V$) is because fibre bundles are themselves smooth manifolds which admit the

Notations used in this chapter

$\mathcal{A}, \mathcal{B}, \ldots$	objects with manifold structure
$\rightarrowtail$; $\twoheadrightarrow$; $\leftrightarrow$	injection; surjection; bijection
$F \rightarrowtail \mathcal{E} \overset{\pi}{\twoheadrightarrow} \mathcal{M}$	fibre bundle $\mathcal{E}$ over base space $\mathcal{M}$ with fibre F and projection $\pi : \mathcal{E} \twoheadrightarrow \mathcal{M}$
$\mathrm{T}_p\mathcal{M}$; $\mathrm{T}\mathcal{M}$	tangent space at $p \in \mathcal{M}$; tangent bundle of $\mathcal{M}$
$\mathrm{T}^r_s\mathcal{M}$	type $\binom{r}{s}$ tensors over $\mathcal{M}$, equal to $\left(\bigotimes_{i=1}^r \mathrm{T}\mathcal{M}\right) \otimes \left(\bigotimes_{i=1}^s \mathrm{T}^*\mathcal{M}\right)$
$\mathbb{T}\mathcal{M}$	tensor bundle of $\mathcal{M}$, equal to $\bigoplus_{r,s=0}^\infty \mathrm{T}^r_s\mathcal{M}$
$\underset{\sim}{A}$	spacetime tensor field; i.e., a section of $\mathbb{T}\mathcal{M}$
$\boldsymbol{A}$	vector field of some abstract vector space not in $\mathbb{T}\mathcal{M}$
$\underset{\sim}{\boldsymbol{A}}$	spacetime tensor field with values in some other abstract vector space
$\wedge^p V$	pth exterior power of the vector space V, consisting of p-forms
$\Gamma(\mathcal{E})$	smooth sections $\mathcal{M} \to F$ of fibre bundle $F \rightarrowtail \mathcal{E} \twoheadrightarrow \mathcal{M}$
$\Omega^p(\mathcal{M})$	space of p-forms on $\mathcal{M}$, equal to $\Gamma(\wedge^p \mathrm{T}^*\mathcal{M})$
$\Omega^p(\mathcal{M}, V)$	space of V-valued p-forms, equal to $\Gamma(\mathcal{E} \otimes \wedge^p \mathrm{T}^*\mathcal{M})$ where $V \rightarrowtail \mathcal{E} \twoheadrightarrow \mathcal{M}$

The gauge transformations of $\boldsymbol{\psi}|_p$ at a point p form a group G under composition—but this group does not describe local gauge transformations which act on the *entire* field $\boldsymbol{\psi}$. Rather, the total space of local gauge transformations is a principle G-bundle $G \rightarrowtail \mathcal{G} \twoheadrightarrow \mathcal{M}$, consisting of smooth maps $\mathcal{M} \to G$.[3] The action of $\mathcal{G}$ on the space of matter fields $\mathcal{E}$ may be denoted

$$\boldsymbol{\psi} \overset{g}{\mapsto} \boldsymbol{\psi}' = g \cdot \boldsymbol{\psi}$$

for $\boldsymbol{\psi} \in \mathcal{E}$ and $g \in \mathcal{G}$ (i.e., $\boldsymbol{\psi} : \mathcal{M} \to V$ and $g : \mathcal{M} \to G$). This specification of an action "$\cdot$" of $\mathcal{G}$ on the vector space $\mathcal{E}$ is equivalent to the choice of a linear representation $\rho : G \to \mathrm{GL}(V)$ of G on V. Thus, we may represent the group action as matrix product, $g \cdot \boldsymbol{\psi} = \rho(g)\boldsymbol{\psi}$, remembering that both $\boldsymbol{\psi}$ and g (but not ρ) vary across $\mathcal{M}$.

Here, it is worth noting that all gauge theories are isomorphic to a gauge theory possessing only one matter field $\boldsymbol{\psi}_{\text{total}}$ with one gauge group (G, ρ). For instance, if a theory consists of different fundamental particles and forces, represented by fields $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ with gauge groups (G_ϕ, ρ_ϕ) and $(G_\varphi, \rho_\varphi)$, then we may take the total matter field $\boldsymbol{\psi}_{\text{total}} = \boldsymbol{\phi} \oplus \boldsymbol{\varphi}$ to be the direct sum of the fields in the theory, and similarly equip it with the gauge group $G_{\text{total}} = G_\phi \oplus G_\varphi$ with representation $\rho_{\text{total}} = \rho_\phi \oplus \rho_\varphi$. This new theory is physically identical to the original. In full generality, therefore, we treat our theory as possessing one matter and one gauge field, while freely speaking of its composite parts as separate fields where convenient.

A *connection* on $\mathcal{E}$ is a derivation[4] $\underset{\sim}{\nabla} : \Gamma(\mathcal{E}) \to \Omega^1(\mathcal{M}, \mathcal{E})$ from vector fields to vector-valued 1-forms, designed so that $(\underset{\sim}{\nabla}\boldsymbol{\psi})(\underset{\sim}{X})$—that is, the action of the V-valued 1-form $\underset{\sim}{\nabla}\boldsymbol{\psi}$ on a vector $\underset{\sim}{X} \in \mathrm{T}\mathcal{M}$—gives the directional derivative of $\boldsymbol{\psi}$ along $\underset{\sim}{X}$ with respect to $\underset{\sim}{\nabla}$. Employing a basis $\{\boldsymbol{e}_a\}$ of $\mathcal{E}$ and local coordinates $\{x^\mu\}$ of $\mathcal{M}$, any connection is of the form

$$\begin{aligned} \underset{\sim}{\nabla}\boldsymbol{\psi} &= \underset{\sim}{\mathrm{d}}\boldsymbol{\psi} + \underset{\sim}{\boldsymbol{A}}\boldsymbol{\psi} \\ &= \left(\partial_\mu \psi^a{}_b + A_\mu{}^a{}_b \psi^b\right) \underset{\sim}{\mathrm{d}} x^\mu \otimes \boldsymbol{e}_a \end{aligned}$$

construction of connections.

[3] We sometimes loosely refer to the fibre group G as the gauge group, but strictly this means the fibre bundle $\mathcal{G}$ (following [2]).

[4] More precisely, $\underset{\sim}{\nabla}$ is a $\mathcal{C}^\infty(\mathcal{M})$-linear derivation, meaning $\underset{\sim}{\nabla}(f\boldsymbol{u} + g\boldsymbol{v}) = f\underset{\sim}{\nabla}\boldsymbol{u} + g\underset{\sim}{\nabla}\boldsymbol{v}$ for scalar fields $f, g \in \mathcal{C}^\infty(\mathcal{M})$ and $\underset{\sim}{\nabla}(\boldsymbol{u} \otimes \boldsymbol{v}) = \underset{\sim}{\nabla}(\boldsymbol{u}) \otimes \boldsymbol{v} + \boldsymbol{u} \otimes \underset{\sim}{\nabla}(\boldsymbol{v})$.

for some matrix-valued 1-form $\underset{\sim}{\boldsymbol{A}}$. If the connection is required to transform like the matter field $\boldsymbol{\psi}$ under local gauge transformations, that is, as

$$\underset{\sim}{\nabla}\boldsymbol{\psi} \overset{g}{\mapsto} g \cdot (\underset{\sim}{\nabla}\boldsymbol{\psi}) = (g \cdot \underset{\sim}{\nabla})(g_\rho \boldsymbol{\psi}) = g_\rho \underset{\sim}{\nabla}\boldsymbol{\psi},$$

then $\underset{\sim}{\nabla}$ is called a *covariant derivative* and the connection 1-form $\underset{\sim}{\boldsymbol{A}}$ consequently obeys

$$\underset{\sim}{\boldsymbol{A}} \overset{g}{\mapsto} g \cdot \underset{\sim}{\boldsymbol{A}} = g_\rho \underset{\sim}{\boldsymbol{A}} g_\rho^{-1} - (\underset{\sim}{\mathrm{d}} g_\rho) g_\rho^{-1}.$$

Such a connection $\underset{\sim}{\nabla}_A$ is not unique; it depends on the choice of the 1-form field $\underset{\sim}{\boldsymbol{A}}$. It is exactly this connection 1-form which is promoted to a dynamical object in a gauge theory and given the name *the gauge field*. In order for the gauge field to be incorporated in the theory's equations of motion, we must have a notion of its derivative. However, $\underset{\sim}{\nabla}_A : \Gamma(\mathcal{E}) \cong \Omega^0(\mathcal{M}, \mathcal{E}) \to \Omega^1(\mathcal{M}, \mathcal{E})$ is not defined on 1-forms such as $\underset{\sim}{\boldsymbol{A}}$ until we make the natural extension[5] to a *covariant exterior derivative* $\underset{\sim}{\mathrm{d}}_A : \Omega^p(\mathcal{M}, \mathcal{E}) \to \Omega^{p+1}(\mathcal{M}, \mathcal{E})$

$$\underset{\sim}{\mathrm{d}}_A = \underset{\sim}{\mathrm{d}} + \underset{\sim}{\boldsymbol{A}},$$

to allow the construction of, among other things, the curvature 2-form, or *gauge field strength*

$$\begin{aligned} \underset{\sim}{\boldsymbol{F}} &:= \underset{\sim}{\mathrm{d}}_A \underset{\sim}{\boldsymbol{A}} \\ &= \underset{\sim}{\mathrm{d}} \underset{\sim}{\boldsymbol{A}} + \underset{\sim}{\boldsymbol{A}} \wedge \underset{\sim}{\boldsymbol{A}} = (\underset{\sim}{\mathrm{d}} \underset{\sim}{A}{}^a{}_b + \underset{\sim}{A}{}^a{}_c \wedge \underset{\sim}{A}{}^c{}_b)\, \boldsymbol{e}_a \otimes \boldsymbol{e}^b, \end{aligned}$$

where we have written the r.h.s. with respect to a basis $\{\boldsymbol{e}_a\}$ and dual basis $\{\boldsymbol{e}^a\}$ of $\mathcal{E}$ (note that V is equipped with the Euclidean inner product; $\boldsymbol{e}^a(\boldsymbol{e}_b) = \delta^a_b$). The field strength $\underset{\sim}{\boldsymbol{F}}$ is useful because it is geometrical in the sense that it transforms in the same matter as the matter fields $\underset{\sim}{\boldsymbol{F}} \mapsto g \cdot \underset{\sim}{\boldsymbol{F}} = g_\rho \underset{\sim}{\boldsymbol{F}} g_\rho^{-1}$ under gauge transformations, even though $\underset{\sim}{\boldsymbol{A}}$ does not.

«The gauge field strength also satisfies the Bianchi identity $\underset{\sim}{\mathrm{d}}_A \underset{\sim}{\boldsymbol{F}} = 0$.»

We have seen how the gauge field $\underset{\sim}{\boldsymbol{A}}$ and its strength $\underset{\sim}{\boldsymbol{F}}$ arise when we require the notion of a spacetime derivative for a matter field which possesses gauge freedom, and are now acquainted with the dynamical objects of a gauge theory. The matter field $\boldsymbol{\psi}$, its derivative $\underset{\sim}{\nabla}_A \boldsymbol{\psi}$ and the strength of the gauge field $\underset{\sim}{\boldsymbol{F}}$ are constructions which all transform regularly under gauge transformations. All that remains to be specified in our theory are the equations of motion, which are to be expressed in terms of these three geometrical objects in a gauge invariant manner.

1.1.1 Lagrangians in Field Theories

For classical gauge theories, the equations of motion may be specified as the extremisation of an action

$$S[\boldsymbol{\psi}, \underset{\sim}{\nabla}_A \boldsymbol{\psi}, \underset{\sim}{\boldsymbol{F}}] = \int_{\mathcal{M}} \underset{\sim}{\mathscr{L}}[\boldsymbol{\psi}, \underset{\sim}{\nabla}_A \boldsymbol{\psi}, \underset{\sim}{\boldsymbol{F}}],$$

[5] The extension is obtained by requiring the graded Leibniz property $\underset{\sim}{\mathrm{d}}_A(\underset{\sim}{\phi} \otimes \boldsymbol{\psi}) = \underset{\sim}{\mathrm{d}}\underset{\sim}{\phi} \otimes \boldsymbol{\psi} + (-1)^p \underset{\sim}{\phi} \wedge \underset{\sim}{\nabla}\boldsymbol{\psi}$ for $\underset{\sim}{\phi} \in \Omega^p(\mathcal{M})$ and $\boldsymbol{\psi} \in \Gamma(\mathcal{E})$, analogous to the usual exterior derivative.

where $\underset{\sim}{\mathscr{L}}$ is a local Lagrangian density. In order that the equations of motion are physically well-defined, we require the Lagrangian possesses the relevant symmetries: gauge symmetry, so that the equations of motion are gauge invariant; and Lorentz symmetry (which is automatic if $\underset{\sim}{\mathscr{L}} = \mathscr{L}\,\underset{\sim}{\mathrm{vol}}$ is expressed as a volume form) [3, § 7.1]. The equations of motion are invariant under adjustments to the Lagrangian density by a total derivative $\underset{\sim}{\mathrm{d}}\alpha$, since by Stokes' theorem these contribute only to boundary terms. Therefore, a Lagrangian which is said to possess gauge symmetry is still generally permitted to transform as $\underset{\sim}{\mathscr{L}} \mapsto \underset{\sim}{\mathscr{L}} + \underset{\sim}{\mathrm{d}}\alpha$ under gauge transformations, if the fields vanish at the boundary.

Furthermore, the Lagrangian of a quantum field theory enjoys an enlarged criterion of gauge symmetry: it is also permitted to transform as

$$S \mapsto S + n \cdot 2\pi\hbar, \tag{1.2}$$

where $n \in \mathbb{Z}$ may vary discretely under different gauges. This is because of the origin of a QFT's equations of motion in the path integral. For instance, the quantum mechanical amplitude that the fields $\boldsymbol{\psi}$ and $\underset{\sim}{\boldsymbol{F}}$ satisfy prescribed boundary conditions on $\partial\Omega$ surrounding some region of spacetime $\Omega \subseteq \mathcal{M}$ is given by the path integral

$$\mathscr{A} = \int_{\partial\Omega} \mathcal{D}[\boldsymbol{\psi}, \underset{\sim}{\boldsymbol{F}}] \exp\left\{\frac{i}{\hbar}\int_{\Omega} \underset{\sim}{\mathscr{L}}[\boldsymbol{\psi}, \underset{\sim}{\nabla}_A\boldsymbol{\psi}, \underset{\sim}{\boldsymbol{F}}]\right\}, \tag{1.3}$$

where the intended meaning of $\mathcal{D}[\cdots]$ is an integration over all field configurations on Ω. If the Lagrangian were to undergo a gauge transformation (1.2), the amplitude $\mathscr{A} \mapsto \mathscr{A}\exp(n \cdot 2\pi i)$ would be left invariant. In other words, physical consistency of a QFT does not require the single-valuedness of S, but only of $\exp(iS/\hbar)$.

We also require that the quantum field theory associated to a Lagrangian be *renormaliseble.* This restricts the form of the Lagrangian considerably.

«transition»

1.1.2 The Dirac Lagrangian

1.1.3 The Yang–Mills Lagrangian and the Topological θ-Term

Taking Lorentz and gauge invariance in into account, the only admissible QFT Lagrangians which may be constructed from the gauge field $\underset{\sim}{\boldsymbol{A}}$ alone are linear combinations of

$$\langle \underset{\sim}{\boldsymbol{F}} \wedge \star\underset{\sim}{\boldsymbol{F}} \rangle \equiv \frac{1}{2}\left\langle \boldsymbol{F}_{\mu\nu}, \boldsymbol{F}_{\rho\sigma}\right\rangle_{\mathrm{Ad}} g^{\mu\rho}g^{\nu\sigma}\underset{\sim}{\mathrm{vol}}, \quad \text{and} \quad \langle \underset{\sim}{\boldsymbol{F}} \wedge \underset{\sim}{\boldsymbol{F}} \rangle \equiv \frac{1}{4}\left\langle \boldsymbol{F}_{\mu\nu}, \boldsymbol{F}_{\rho\sigma}\right\rangle_{\mathrm{Ad}} \epsilon^{\mu\nu\rho\sigma}\underset{\sim}{\mathrm{vol}},$$

where $\star$ is the Hodge dual and $\langle\, ,\, \rangle_{\mathrm{Ad}}$ is a gauge-invariant inner product on the Lie algebra $\mathfrak{g}$ of the gauge group G. (Recall that $\underset{\sim}{\boldsymbol{F}}$ is a $\mathfrak{g}$-valued form, so an inner product on $\mathfrak{g}$ is needed, and it must be gauge- or 'Ad-invariant' [3, § 7.1.2].) The inner product $\langle\, ,\, \rangle_{\mathrm{Ad}}$ on $\mathfrak{g}$ is not unique; it depends on a choice of *coupling constants.* In particular, if the gauge group G is the direct sum of n simple Lie groups,[6] then $\langle\, ,\, \rangle_{\mathrm{Ad}}$ is specified by the choice of exactly n coupling constants,

[6] Any compact connected Lie group is either of this form, or is a finite quotient of such a group, if $\mathrm{U}(1)$ is counted as 'simple'. [3, § 2.4.3].

one corresponding to each factor of G [3, § 2.5]. Physically, the coupling constants determine the relative interaction strengths of the forces associated to each factor of G, and must enter the theory as free parameters determined by experiment. (For instance, the gauge group of the standard model (1.1) has three such coupling constants for the strong $\mathrm{SU}(3)$, weak $\mathrm{SU}(2)$, and electromagnetic $\mathrm{U}(1)$ interactions.)

The term $\langle \underset{\sim}{\boldsymbol{F}} \wedge \star \underset{\sim}{\boldsymbol{F}} \rangle$ is known as the *Yang–Mills Lagrangian*, and is a major component in the standard model, describing boson force carrier self-interactions (such as gluon self-interactions). The Yang–Mills Lagrangian yields the equation of motion $\underset{\sim}{\mathrm{d}}_A \star \underset{\sim}{\boldsymbol{F}} = 0$. For the Abelian gauge group of electromagnetism $G = \mathrm{U}(1)$, this, together with the Bianchi identity $\underset{\sim}{\mathrm{d}}_A \underset{\sim}{\boldsymbol{F}} = 0$, are the source-free Maxwell's equations.

The other term $\langle \underset{\sim}{\boldsymbol{F}} \wedge \underset{\sim}{\boldsymbol{F}} \rangle$ is known as the *topological θ-term*, for reasons which will become apparent. The θ-term breaks both time-reversal symmetry T and parity P (since $\epsilon^{\mu\nu\rho\sigma} \mapsto -\epsilon^{\mu\nu\rho\sigma}$ under T or P). It is *topological* because it does not depend on the spacetime metric $g^{\mu\nu}$. Its integral over spacetime is therefore a topological invariant,

$$\frac{1}{8\pi^2} \int_{\mathcal{M}} \langle \underset{\sim}{\boldsymbol{F}} \wedge \underset{\sim}{\boldsymbol{F}} \rangle = n \in \mathbb{Z},$$

where n is known as the *Pontryagin number*, the *second Chern class* [4, § 1] or simply the 'winding number' [2, § 2.2] of the gauge field configuration $\underset{\sim}{\boldsymbol{A}}$.

For an Abelian gauge group $\mathrm{U}(1)$, the space of gauge fields $\Omega^1(\mathcal{M}, \mathfrak{g})$ is path-connected to the identity, and the winding number n is zero for all $\underset{\sim}{\boldsymbol{A}} \in \Omega^1(\mathcal{M}, \mathfrak{g})$. Consequently, there is no θ-term in classical electromagnetism or QED. However, for non-Abelian gauge groups, $\Omega^1(\mathcal{M}, \mathfrak{g})$ may have non-trivial topology, with path-disconnected regions which can be indexed by $n \in \mathbb{Z}$. The presence of the θ-term in QCD can be viewed as a consequence of the non-Abelian nature of the gauge group $\mathrm{SU}(3)$. Gauge field configurations $\underset{\sim}{\boldsymbol{A}} \in \Omega^1(\mathcal{M}, \mathfrak{g})$ which are outside of the identity-connected component are known as *instanton configurations*. Instantons and are an important aspect of QCD and general Yang–Mills theories, which give rise to topologically distinct vacuum states and non-perturbative dynamical effects [5].

The choice of the symbol θ in the name reflects the cyclic property of a coefficient θ given to the topological θ-term: If the Lagrangian density includes the term

$$\underset{\sim}{\mathscr{L}}_\theta[\underset{\sim}{\boldsymbol{F}}] = \frac{\theta}{8\pi^2} \langle \underset{\sim}{\boldsymbol{F}} \wedge \underset{\sim}{\boldsymbol{F}} \rangle ,$$

then its contribution to the action in the path-integral (1.3) is θn. Since $e^{i\theta n} = e^{i(\theta + 2\pi)n}$, the coefficient $\theta \in [0, 2\pi)$, known as the *θ-parameter*, is physically periodic.

1.2 QCD and the Strong CP Problem

Quantum chromodynamics describes the strong interactions between hadronic matter. Pure QCD is a quantum gauge theory whose matter field describes quarks and whose gauge field describes gluons. Pure QCD can be interpolated with QED or the standard model, in which case Quarks are fermionic Dirac fields which come in pairs of *up type* and *down type* quarks, each transforming under different representations of weak hypercharge $\mathrm{U}(1)$ (but which are identical in pure QCD). Pure QCD is agnostic to the number of quark pairs, but when included in the standard model, three copies (each called a *generation*) of quark pairs are used, meaning the standard model contains six quark fields in total . «??»

One of the attractive qualities of QCD is that, given the gauge group $G = \mathrm{SU}(3)$ and its action on a matter field ψ, the theory is almost completely specified given reasonable physical assumptions. That is, there is not much freedom for QCD to be slightly different without being a drastically different theory.

The Lagrangian of pure, N-quark QCD is a sum of the Dirac Lagrangian and the Yang–Mills Lagrangian,

$$\begin{aligned}\mathcal{L}_{\mathrm{QCD}} &= \bar{\boldsymbol{\psi}}(i\boldsymbol{\gamma}^\mu \nabla_\mu - m)\boldsymbol{\psi} - \langle \underset{\sim}{\boldsymbol{F}} \wedge \star \underset{\sim}{\boldsymbol{F}} \rangle \\ &\equiv \sum_{q=1}^{N} \bar{\psi}_a^{(q)} \left(i\gamma^{\mu a}{}_b \nabla_\mu - m\delta^a{}_b\right)\psi^b_{(q)} - \frac{1}{4} F^a{}_{b\mu\nu} F_a{}^{b\mu\nu}\end{aligned}$$

where $\boldsymbol{\psi} \equiv \bigoplus_{q=1}^{N} \boldsymbol{\psi}^{(q)}$ is the fermion field separated into its quark components, each of which is a Dirac spinor (transforming under the spin-$\frac{1}{2}$ representation of the Lorentz group). The Dirac matrices $\boldsymbol{\gamma}^\mu$ are basis elements of the Clifford algebra $\mathcal{C}l_{1,3}(\mathbb{C})$ satisfying $\boldsymbol{\gamma}^{(\mu}\boldsymbol{\gamma}^{\nu)} = \eta^{\mu\nu}\mathbf{1}$ and $\underset{\sim}{\nabla}\boldsymbol{\psi} \equiv (\nabla_\mu \boldsymbol{\psi})\, \underset{\sim}{\mathrm{d}}x^\mu$ is the covariant derivative with respect to the gluon gauge field, $\underset{\sim}{\boldsymbol{A}}$.

At this point, we see that the issue which is given the name *the strong CP problem* offers itself as a question about the naturalness the QCD Lagrangian: As encountered in section 1.1.3, there is another term which may be included in the Lagrangian with no additional physical assumptions; the topological θ-term. In this sense, the most general pure QCD Lagrangian is

$$\mathcal{L}_{\mathrm{QCD}} = \bar{\boldsymbol{\psi}}(i\boldsymbol{\gamma}^\mu \nabla_\mu - m)\boldsymbol{\psi} - \langle \underset{\sim}{\boldsymbol{F}} \wedge \star \underset{\sim}{\boldsymbol{F}} \rangle + \frac{\theta}{8\pi^2} \langle \underset{\sim}{\boldsymbol{F}} \wedge \underset{\sim}{\boldsymbol{F}} \rangle\,, \tag{1.4}$$

which includes the CP-violating θ-term $\langle \underset{\sim}{\boldsymbol{F}} \wedge \underset{\sim}{\boldsymbol{F}} \rangle$.

The inclusion of the θ-term is 'natural' because we lack reason to exclude it on a theoretical basis: it is Lorentz and gauge invariant, etc., just like its CP-symmetric counterpart, $\langle \underset{\sim}{\boldsymbol{F}} \wedge \star \underset{\sim}{\boldsymbol{F}} \rangle$. In particular, the weak interaction is explicitly parity-violating, and the standard model is asymmetric under all combinations of charge conjugation symmetry, C; parity P; and time-reversal symmetry T except for combined CPT symmetry—and therefore the θ-term's CP violation is not satisfactory reason for its exclusion.

If the θ-parameter is nonzero, then the strong force is predicted to violate CP in general. After integration into the standard model, a physical implication of CP violation in QCD is that the neutron is expected to possess an electric dipole moment of approximate magnitude $|d_n| \approx 10^{-18}\, e\,\mathrm{cm}$. In reality, current measurements [6] of the neutron's electric dipole moment place an upper bound of $|d_n| \lesssim 10^{-26}\, e\,\mathrm{cm}$, and this implies a stringent constraint on the responsible θ-term, $|\theta| \lesssim 10^{-10}$ [7].

2 Solutions to the Strong CP Problem

At first sight, the strong CP problem may not appear to be a problem at all. After all, QCD is a theory which possibly—but not necessarily—admits a term which gives rise to CP-violating interactions in the strong force and consequently to an electric dipole moment of the neutron. Empirical data shows that the neutron's electric dipole moment (and hence the CP-violating term) is non-existent. From a phenomenological perspective, it is satisfactory to simply leave the CP-term out of the theory's Lagrangian and end the story here. Indeed, a trivial way to 'resolve' the strong CP problem is to simply require that CP be a symmetry of the strong force. However, this is tautological and only reinforces the problem of why CP symmetry appears to be preserved in some sectors of the standard model while it is broken in others.

Furthermore, there is a strong sense that the inclusion of the CP-violating term is natural while its exclusion requires theoretical justification. As mentioned in the previous section, the standard model is not CP symmetric, and hence the CP-violating θ-parameter is expected to be of $\mathcal{O}(1)$. Similarly, the advent of instantons in QCD demonstrated that the θ-term does not necessarily vanish [5, 8].«?????»

The strong CP problem differers from other fine-tuning problems in the standard model «such as?» in that it is of almost no consequence to everyday physics, and therefore has no anthropic solution. Whereas variations of the cosmological constant, the weak scale or the quark and lepton masses predict a universe very different to our own, variation of the θ-parameter hardly affects nuclear physics at all, because it is suppressed by quark masses.

2.1 The Massless Quark Solution

In the pure QCD Lagrangian (1.4), which includes the mass term $m\bar{\psi}\psi$, quarks possess mass given by the parameter m. This is *not* the mechanism by which quarks exhibit mass in the standard model—instead, quarks obtain mass via the *Higgs mechanism*, whereby ψ is coupled to the Higgs field by *Yukawa interaction* terms in the Lagrangian [3, § 7.6.6]. The Yukawa terms contain quark–Higgs coupling parameters m_q which can be arranged into the *quark mass matrix* $\mathcal{M}$. «In an abuse of notation, the QCD sector of the standard model may be regarded as being given by the effective Lagrangian

$$\mathcal{L}_{\mathrm{QCD}} = \bar{\psi}(i\gamma^\mu \nabla_\mu - \mathcal{M})\psi - \langle \underset{\sim}{\boldsymbol{F}} \wedge \star \underset{\sim}{\boldsymbol{F}} \rangle + \frac{\theta}{8\pi^2} \langle \underset{\sim}{\boldsymbol{F}} \wedge \underset{\sim}{\boldsymbol{F}} \rangle$$

»

In the full quantum theory of N massive quarks, the Lagrangian is invariant under transformations defined by N parameters $a_i \in \mathbb{R}$ and the mappings

$$\psi_{(q)} \mapsto e^{\frac{i}{2}\gamma_5 a_q}\psi_{(q)}, \qquad m_q \mapsto e^{-ia_q} m_q, \qquad \theta \mapsto \theta - \sum_{q=1}^{N} a_q, \tag{2.1}$$

where $\psi_{(q)}$ are the quark fields and m_q are their masses. The fact that the θ-parameter may be redefined by shifting the quark mass phases means that θ is not directly observable. This gauge freedom can be fixed by defining the quantity

$$\bar{\theta} = \theta - \arg\det \mathcal{M} = \theta - \arg \prod_{i=1}^{N} m_i.$$

The gauge-fixed $\bar{\theta}$-parameter is invariant under (2.1), as the two right-hand terms transform by the addition of $\mp \sum_{i=1}^{N} a_i$, respectively. The significance of this is that only variations in $\bar{\theta}$ can possibly have physical effects. In particular, $\bar{\theta}$ becomes singular when one of the quark masses is zero. This means that the $\bar{\theta}$-parameter, and hence the θ-term, are entirely unphysical in the case of a massless quark,.

At first sight, this economical resolution to the strong CP problem appears to be in contradiction with the experimentally determined masses of the quarks, all of which are non-zero (including the up quark, whose mass is estimated to be $\sim 2\,\text{MeV}$). However, it was realised in the mid-1980s that the mass of the up quark has two contributions in the standard model Lagrangian: not only the Yukawa mass m_u introduced above (the 'bare mass'), but also a non-perturbative contribution m_{eff} from topological effects (such as instantons) [9]. Importantly, it was plausible that this secondary source of the up quark's mass is of order $m_{\text{eff}} \approx 2\,\text{MeV}$, allowing m_u to vanish by another mechanism.

The massless up quark hypothesis remained controversial until the advancement of *lattice gauge theory* made numerical simulations of non-perturbative effects in QCD possible. By 2020, consensus was reached that the instanton contribution m_{eff} is not sufficiently large, and hence that the massless up quark hypothesis is false [9–11]. Instead, another mechanism is required to explain the smallness of the $\bar{\theta}$-parameter and solve the strong CP problem.

2.2 Peccei–Quinn Theory

Perhaps the most famous proposed resolution to the strong CP problem is the Peccei–Quinn theory of the *axion*, first proposed in 1977 [12]. The axion is well-known in cosmology because of the its allure as a dark matter candidate.

Peccei–Quinn theory explains the small value of θ by its promotion to a dynamical field whose potential is minimised when $\theta = 0$. In other words, a new $\mathrm{U}(1)$ symmetry is introduced (whose action is to rotate θ) and is adjoined to the gauge group of QCD. The QCD Lagrangian is modified by the addition of a potential with minimum at $\theta = 0$, which is said to *spontaneously break* the Peccei–Quinn $\mathrm{U}(1)$ symmetry. The gauge field associated to this $\mathrm{U}(1)$ symmetry is named the *axion field*, and the gauge bosons obtained upon quantisation are named *axions*.

«

1. Predictions of PC theory

2. Theoretical limitations
3. Experimental evidence

»

3 Axions and Cosmology

Bibliography

[1] S. Weinberg, "The Making of the standard model", The European Physical Journal C-Particles and Fields **34**, 5–13 (2004).

[2] D. Tong, *Cambridge, Department of Applied Mathematics and Theoretical Physics, Lecture Notes: Gauge Theory*, URL: `http://www.damtp.cam.ac.uk/user/tong/gaugetheory.html`. Last visited on 2020/05/24, 2018.

[3] M. Hamilton, *Mathematical Gauge Theory: With Applications to the Standard Model of Particle Physics*, Universitext (Springer International Publishing, 2017).

[4] E. Witten, "Quantum field theory and the Jones polynomial", Communications in Mathematical Physics **121**, 351–399 (1989).

[5] S. Vandoren and P. van Nieuwenhuizen, *Lectures on instantons*, 2008.

[6] C. Abel et al., "Measurement of the Permanent Electric Dipole Moment of the Neutron", Phys. Rev. Lett. **124**, 081803 (2020).

[7] M. Tanabashi et al., "Review of Particle Physics", Phys. Rev. D **98**, 030001 (2018).

[8] G. Gabadadze and M. Shifman, "QCD vacuum and axions: What's happening?", International Journal of Modern Physics A **17**, 3689–3727 (2002).

[9] C. Alexandrou et al., *Ruling out the massless up-quark solution to the strong CP problem by computing the topological mass contribution with lattice QCD*, 2020.

[10] M. Dine, P. Draper, and G. Festuccia, "Instanton effects in three flavor QCD", Phys. Rev. D **92**, 054004 (2015).

[11] S. Aoki et al., *Review of lattice results concerning low-energy particle physics*, 2016.

[12] R. D. Peccei and H. R. Quinn, "CP Conservation in the Presence of Pseudoparticles", Phys. Rev. Lett. **38**, 1440–1443 (1977).