Solutions of the Strong CP Problem: A Scorecard

Michael Dine

Department of Physics
University of California, Santa Cruz

Work and discussions with P. Draper, G. Festuccia and S. Thomas
Institute for Advanced Study; Technion/Weizmann/TelAviv, Feb. 2015/Rutgers March 2015

Usually speak of three solutions of the strong CP problem

1. $m_u = 0$
2. Spontaneous CP violation with nearly vanishing θ ("Nelson-Barr" or NB)
3. The axion, or the Peccei-Quinn symmetry

There are others (e.g. Hiller and Schmaltz, Anson Hook) which can be shoehorned into this classification scheme).

Among naturalness problems, the strong CP problem is special in that it is of almost no consequence. We don't have to invoke anthropic selection to realize that if the cosmological constant was a few orders of magnitude larger than observed, the universe would be dramatically different. The same is true for the value of the weak scale and of the light quark and lepton masses. But if θ were, say, 10^{-3}, nuclear physics would hardly be different than we observe, since effects of θ are shielded by small quark masses.

So while theorists may be endlessly clever in providing solutions to the problem, we might choose to be guided by a principle that the smallness of θ should be a consequence of other aspects of physical theory, or, at the least, a plausible accident of features of an underlying theory.

One goal today is to ask to look at how each of these solutions might fare under this principle. Needles to say, this will require us to be at least somewhat precise about what these solutions are and how they differ (and not least whether this characterization is in some sense exhaustive).

I come at this with a bias, and can readily twist things to get the result I want. But it should be clear that, at least within the present state of our knowledge of what a complete theory of high energy physics might look like (string theory in its various guises) we can not make a statement divorced from prejudice.

The $m_u = 0$ and axion solutions have been subjected to the greatest scrutiny under this principle, so I will devote a disproportionate amount of time to the NB solution.

Outline

- $m_u = 0$.
  1. Theoretical justifications
  2. Lattice status and a proposed calibration of lattice measurements of m_q
  3. Generalizations
- The axion: PQ *Quality*
- Nelson-Barr
  1. Loop corrections at low energy and insensitivity to high energy physics
  2. Role of axions (and reprise for $m_u = 0$)
  3. Nelson-Barr in a landscape
  4. Tuning of parameters at tree level
  5. Loop Corrections

If $m_u = 0$, one can rotate away θ. More precisely, one requires

$$\frac{m_u}{m_d} < 10^{-10}. \tag{1}$$

There are two issues with this proposal:

1. Why might m_u be so small?
2. We can measure m_u (with the help of the lattice). Is this consistent with lattice results?

Accounting for small m_u

Banks, Nir, Seiberg put forward models which, in accounting for quark flavor, gave rise to small or zero m_u.

A simple possibility is suggested by string theory, which often exhibits anomalous discrete symmetries; more precisely, the chiral content of the theory is anomalous, with the anomaly being cancelled by the non-linear transformation of an axion-like field. In the supersymmetric case, this means that one has a modulus field, coupling to the $\bar{u}$ quark as

$$e^{-\Phi} Q H_U \bar{u}. \tag{2}$$

One requires that the exponential be very small, but this is plausible. One can speculate as to whether or not a suitable discrete symmetry structure is typical of underlying theories.

How might $m_u = 0$ be consistent with known facts of hadron physics

Instantons suggestive (Georgi-McArthur). With three light quarks, generate an effective u quark mass (two point function) proportional to $m_d m_s$. Simple dimensional analysis suggests the effect goes as

$$\frac{m_d m_s}{\Lambda} \tag{3}$$

with Λ a suitable QCD scale. This could easily be of order the few MeV expected from current algebra. Kaplan and Manohar expressed this as an ambiguity in current algebra, i.e. they isolated a term and second order in quark masses which could mimic the effects of a u quark mass.

Current algebra analysis

(Follows treatment of Banks, Nir, Seiberg)

Leading order chiral Lagrangian:

$$\mathcal{L}_2 = \frac{F^2}{4}\mathrm{Tr}\left(\partial_\mu U^\dagger \partial^\mu U\right) + \frac{F^2}{4}\mathrm{Tr}\left(\chi^\dagger U\right) + \mathrm{c.c.} \tag{4}$$

$$\chi = 2MB_0\ , \quad U = e^{i\frac{\lambda^A \pi^A}{F}}\ , \tag{5}$$

M is the quark mass matrix, and B_0 is proportional to the magnitude of the chiral condensate.

The second order operator

$$\mathcal{L} \supset r_1 \left(\mathrm{Tr}(\chi^\dagger U \chi^\dagger U) - \mathrm{Tr}(\chi^\dagger U)^2 \right), \tag{6}$$

can mimic the effects of a u quark mass. Indeed, by a redefinition of χ, r_1 can be eliminated, providing an effective contribution to m_u of order $m_d m_s$. Alternatively, having fixed the ambiguity by requiring – for instance – that M is proportional to the UV quark mass matrix in a precisely specified scheme, a value of r_1 (of order 10^{-3} and a small value of the bare m_u would be compatible with the observed pseudoscalar meson masses.

The KM transformation is not a symmetry of QCD. In a lattice computation with fixed bare masses and small lattice spacing, one can take M to be the bare mass matrix. Then, for example,

$$m_\pi^2 = \beta_1(m_u + m_d) + \beta_2 m_s(m_u + m_d) + \mathcal{O}(m_{u,d}^2)\ . \tag{7}$$

Typical lattice calculations are done with $m_u(a) = m_d(a) \equiv \hat{m}$. The parameters β_1 and β_2 can be extracted on the lattice by varying $\hat{m}$ and $m_s(a)$ independently; e.g.,

$$\frac{\beta_2}{\beta_1} \approx \frac{m_{\pi_1}^2 - m_{\pi_2}^2}{m_{\pi_2}^2 m_{s_1} - m_{\pi_1}^2 m_{s_2}}\ . \tag{8}$$

Require

$$\frac{\beta_2}{\beta_1} \approx 5 \ \mathrm{GeV}^{-1}. \tag{9}$$

Lattice results are quoted in terms of Gasser-Leutweyler parameters. From quoted lattice results we estimate

$$\frac{\beta_2}{\beta_1} \approx (1) \ \mathrm{GeV}^{-1}. \tag{10}$$

The error bars are large, but the ratio is far too small to account for the effects of the *u* quark mass. However, β_2/β_1 is a fundamental prediction of QCD and it would be interesting to see a dedicated study with increased precision. It would provide another demonstration of $m_u \neq 0$, as well as a probe of the contribution of small instantons to the chiral lagrangian.

Summary of lattice results for light quark masses

Current results from lattice simulations (summarized by the FLAG working group)

$$m_u = 2.16\ (9)(7)\mathrm{MeV} \quad m_d = 4.68\ (14)(7)\mathrm{MeV} \quad m_s = 93.5(2.5)\mathrm{MeV}$$

Numbers are in $\overline{MS}$ scheme at 2 GeV.

So m_u is many standard deviations from zero.

Aside: A lattice test

Arguably this is the end of the story. But the calculations are complicated and the stakes are high (indeed, the question of whether m_u is consistent with zero would seem the most interesting question for light quark physics on the lattice).

One possible test of error estimates. Singular parts of certain Green's functions, in the massless limit, calculable from instantons. Roughly objects like:

$$G(x) = \langle \bar{u}(x) u(x) \bar{d}(x) d(x) \bar{s}(0) s(0) \rangle \tag{12}$$

With two colors, leading term behaves as

$$G(x) = C \frac{\Lambda^{16/3}}{|x|^{11/3}} \tag{13}$$

Coefficient is known; there is a systematic expansion in powers of $g(x)$. Analysis for $N = 3$ slightly more involved because a borderline case.

Challenging, but given small masses and lattice spacings

Axion *Quality*

Global symmetries should arise only as accidents of gauge symmetry and the structure of low dimension terms in an effective action. It has been recognized almost from the beginning that this is a challenge for the axion solution of the strong CP problem.

We can define an axion quality factor, Q_a, as

$$Q_a = \frac{f_a \frac{\partial V_{pq-v}(a)}{\partial a}}{m_\pi^2 f_\pi^2} \tag{14}$$

where V_{pa-v} is the PQ-violating potential. Solving the strong CP problem requires

$$Q_a < 10^{-10} \tag{15}$$

In a conventional effective field theory analysis (i.e. finite number of degrees of freedom above f_a), this is quite a challenge. If

$$\langle \Phi \rangle = f_a e^{ia/f_a} \tag{16}$$

symmetry violating operators like

$$\frac{\Phi^{n+4}}{M_p^n} \tag{17}$$

make too large a contribution to Q_a even for $f_a = 10^{11}$ GeV unless $n > 7$. We might try to achieve this with a discrete Z_N, but this requires $N = 11$ at least, which certainly violates our minimalist principle.

Axions in string theory

Witten pointed out early on that string theory provides a possible resolution to this conundrum.

This is most easily understood in the framework of supersymmetry. Typically string models possess moduli, Φ, whose imaginary component obeys a discrete shift symmetry:

$$\Phi = x + ia; \quad a \to a + 2\pi \tag{18}$$

This insures, for example, that any superpotential is a function of $e^{-\Phi}$ at large x. Here x might be $\frac{8\pi^2}{g^2}$ for some gauge coupling g.

So the issue becomes: why or whether the theory sits in an asymptotic region of the moduli space where e^{-x} is very small. One can put forward various scenarios (and this is consistent at least with the fact that the observed gauge couplings are small), but reliable computations are not possible at present.

The Nelson-Barr mechanism

Invokes spontaneous CP violation to argue “bare θ" is zero. Constructs a mass matrix such that CP breaking gives a large CKM angle (as observed, $\delta = 1.2$) with $\arg \det m_q = 0$.

Such a structure is perhaps made plausible by string theory, where CP is a (gauge) symmetry, necessarily spontaneously broken. Some features of the required mass matrices appear, e.g., in Calabi-Yau compactifications of the heterotic string.

Simple realization of the NB structure

Complex scalars η_i with complex (CP-violating) vev's.
Additional vectorlike quark with charge $1/3$.

$$\mathcal{L} = \mu \bar{q} q + \lambda_{if} \eta_i \bar{d}_f q + y_{fg} Q_f \bar{d}_g \phi \tag{19}$$

where ϕ is Higgs; y, λ, μ real.

$$M = \begin{pmatrix} \mu & B \\ 0 & m_d \end{pmatrix} \tag{20}$$

$B_f = \lambda_{if} \eta_i$. This has real determinant.

The structure is reminiscent of an E_6 gauge theory, which has the requisite vector-like quarks and singlets.

Requirements for a successful NB Solution

1. Symmetries: It is important that η_i not couple to $\bar{q}q$, for example. So, e.g., η's complex, subject to a Z_N symmetry.
2. Coincidences of scale: if only one field η, CKM angle vanishes (can make d quark mass matrix real by an overall phase redefinition). Need at least two, and their vev's (times suitable couplings) have to be quite close:

$$\delta_{CKM} \propto \frac{B_{small}}{B_{large}} \tag{21}$$

3. Similarly, μ (which might represent vev of another field) can not be much larger than η_i, and if much smaller the Yukawa's and B's have to have special features.

Constraints on the Overall Scale

Before considering radiative effects, possible higher dimension operators in $\mathcal{L}$ constrain the scales η_i, μ. E.g.

$$\frac{\eta_i^* \eta_j}{M_p} \bar{q} q \tag{22}$$

requires $\frac{|\eta|}{M_p} < 10^{-10}$.

Barr-Nelson With/Without Supersymmetry

Without supersymmetry, highly tuned. Two light scalars and μ (or three light scalars). Far worse than θ.

Even ignoring that, require close coincidence of scales.

Supersymmetry helps. Allows light scalars. Coincidences still required (and typically more chiral multiplets to achieve desired symmetry breakings – typically seven in total). Some of the high dimension operators better controlled (e.g. if μ, η_i much larger than susy breaking scale, don't have analogs of the $\eta_i^* \eta_j \bar{q} q$ operator).

Axions in the NB Scenario

What does it mean that the “bare" θ is naturally zero in a model which is CP-conserving at some underlying level? String theory provides a realization. Here one might mean that the vev's of the moduli are CP conserving, i.e. that the various axions have vanishing vev. These axions might be presumed to be heavier than the conventional QCD axion (otherwise they would provide a PQ resolution of strong CP). Such masses could arise from strong string effects, or other strong gauge groups.

So NB might be considered a particular limit of the PQ picture. Here it is not necessary that the quality be particular good, provided that $\arg\det m_q \approx 0$ and the axion coupling to the fields which break CP is weak enough.

How plausible is $\theta_{bare} = 0$

Thinking of "θ_{bare}" as the expectation value of some axion-like field, one can ask: how likely is it that this quantity vanishes. One model: flux landscapes. Here, "KKLT" we might consider a model with a superpotential

$$W = e^{-\Phi/b} + W_0. \tag{23}$$

Looking at supersymmetric stationary points, the value of the axion depends on the phase of W_0. $\theta_{bare} = 0$ requires that W_0 is real.

In a landscape, this is likely to be extremely rare. Roughly speaking requires that all CP-odd fluxes (presumably 1/2) should vanish. Exponential suppression.

Loop Corrections at Low Energies

Note, first, that loop corrections to θ in the Standard Model are highly suppressed. Focussing on divergent corrections, one requires Higgs loops. These involve the Hermitian matrices

$$A = y_d^\dagger y_d; \quad B = y_u^\dagger y_u \tag{24}$$

Contributions to θ are proportional to traces of the form

$$Tr(ABA^2B\dots) \tag{25}$$

one additional matrix factor for each loop. It is easy to check that the first complex combination involves six matrices, e.g.

$$\mathrm{Tr}(ABA^2B^2) \tag{26}$$

but this and its complex conjugate both appear with the same coefficient. It is necessary to add a $U(1)$ gauge loop (which distinguishes u and d) to have the possibility of a complex traces. [Ellis, Gaillard]

Loop Corrections in Nelson-Barr

In the non-supersymmetric case, in the simplest model, potential corrections arise at three loops. These will constrain modestly the various coupling constants.

If there are new light degrees of freedom, these can be more problematic. This is well-known for the case of supersymmetry, where phases in gaugino masses, and in the squark masses can lead to large corrections.

Constraints on new degrees of freedom (discussions with S. Thomas)

While we have noted that NB does not require new degrees of freedom, what is perhaps more interesting is that it does not admit (or does not admit without great difficulty) new degrees of freedom.

E.g. Extra Higgs doublets: phases in $H_U H_D$ terms in potential (and others). In non-susy case, potentially huge contributions from $\eta_i \eta_j^* H_U H_D$, etc. Lead to large relative phases in quark mass matrices.

Supersymmetry breaking and Nelson-Barr

Many possible phases once allow soft breaking **Note: these effects don't decouple for large susy-breaking scale**. E.g. is susy breaking described by Goldstino superfield, X, superpotential couplings

$$\frac{\mathcal{O}_d}{M_p^{d-2}} X \tag{27}$$

where $\langle \mathcal{O} \rangle$ is complex can lead to large phases in soft breakings. Similarly phases in W. Phases in gaugino masses feed directly into θ.

Phases in neutrino masses

In the nonsupersymmetric case, we have seen that the scale of CP violation must be rather low, possibly less than 10^8 GeV. This is a low scale compared to the scales at which we generally suspect neutrino masses are generated. So NB would strongly point to a real PMNS matrix. Leptogenesis unlikely as the origin of the matter-antimatter asymmetry.

Conclusions

Each proposed solution to the strong CP problem raises troubling questions. We have argued, indeed, that θ is of so little importance that any solution should be an outcome of some other constraint on the physical theory. Solutions which require many additional degrees of freedom, intricate symmetries, or significant fine tuning have little plausibility.

1. A very light u quark might be a consequence of horizontal symmetries, or might arise as a result of the anomalous discrete symmetries which seem ubiquitous in string theory. However, there are now lattice computations which appear to definitively rule out the possibility.
2. The axion raises the issues of the quality of the PQ symmetry. String theory suggests a plausible answer, but our understanding is limited.
3. Nelson-Barr: The basic premise, that if the underlying theory is CP conserving, the “bare" θ vanishes, is open to question; it requires an understanding of how certain moduli are stabilized, and in a landscape would seem unlikely.

Still more issues for Nelson-Barr:

These questions aside, without supersymmetry, the NB models are constrained to relatively low scales and highly tuned to achieve $\arg\det m_q$, while reproducing observed KM angle. We have noted that one requires several fields, with very similar expectation values, well below the Planck scale.

Even in a supersymmetric framework, one requires several fields (typically more than six), a close coincidence of scales, and large (but not as large as for the axion) discrete symmetries. None of these features are obvious ingredients to explain some other question, though conceivably in theories of flavor some of them might arise.

Other solutions

1. Hiller-Schmaltz: essentially a variant on Barr-Nelson. CP violation arises in kinetic terms of supersymmetric fields. Requires strong couplings to obtain large δ_{CKM} and poses other issues microscopically.
2. Anson Hook: Z_2 relates two copies of the Standard Model; additional fields transform under an approximate chiral $U(1)$. Only one θ parameter. Breaking of Z_2 allows “other" QCD to remove θ. Interesting low energy consequences. Similarities to massless u quark solution.

I will leave it to you to make a final scoresheet, and a viewpoint on which solution of the strong CP problem is the most likely.