Axions in string theory

Peter Svrček

Department of Physics and SLAC, Stanford University
Stanford CA 94305/94309 U.S.A.
E-mail: `svrcek@stanford.edu`

Edward Witten

Institute For Advanced Study
Princeton NJ 08540 U.S.A.
E-mail: `witten@ias.edu`

ABSTRACT: In the context of string theory, axions appear to provide the most plausible solution of the strong CP problem. However, as has been known for a long time, in many string-based models, the axion coupling parameter F_a is several orders of magnitude higher than the standard cosmological bounds. We re-examine this problem in a variety of models, showing that F_a is close to the GUT scale or above in many models that have GUT-like phenomenology, as well as some that do not. On the other hand, in some models with Standard Model gauge fields supported on vanishing cycles, it is possible for F_a to be well below the GUT scale.

Contents

1. Review and introduction

The purpose of this paper is to reassess the status of axions in string theory. We begin with a review and introduction, after which, in section 2, we make some general remarks. The rest of the paper is devoted to analyzing axions in various string models. The main conclusion is as it has been in the past: there is some tension between string models of axions and cosmological bounds, with the axion coupling parameter in many string models being larger, and hence the axion frequency smaller, than allowed by the usual cosmological arguments. However, we also discuss some string models, both old ones and new ones, that are compatible with the usual cosmological bounds. Since the early days of QCD instanton physics [1–4], it has been understood that to the action of QCD, it is possible to add a CP-violating interaction

$$I_\theta = \frac{\theta}{32\pi^2} \int d^4x \epsilon^{\mu\nu\alpha\beta} \operatorname{tr} F_{\mu\nu} F_{\alpha\beta}. \tag{1.1}$$

Here θ is a coupling parameter, and the interaction it multiplies measures the Yang-Mills instanton number. Since the instanton number, with appropriate boundary conditions, is an integer (and more generally, the instanton number is determined by the boundary conditions modulo an integer), θ is an angular parameter.

The strong CP problem [5] is the problem of explaining the extreme smallness of θ, or more precisely of $\overline{\theta}$, the effective θ after rotating away the phases of the quark bare masses. From the upper limit on the electric dipole moment of the neutron, most recently $|d_n| < 6.3 \times 10^{-26} \, e \, \mathrm{cm}$ [6], one has roughly $|\overline{\theta}| < 3 \times 10^{-10}$. From limits on the electric dipole moment of $^{199}\mathrm{Hg}$ [7], one has $|\overline{\theta}| < 1.5 \times 10^{-10}$. At the factor of two level, the limits on $\overline{\theta}$ are subject to some QCD uncertainties (for a recent discussion, see [8]) and some nuclear and atomic uncertainties in interpreting the results from $^{199}\mathrm{Hg}$.

Broadly speaking, three solutions to the strong CP problem have been proposed:

1. The up quark mass may vanish.

2. $\overline{\theta}$ can relax spontaneously to a suitably small value if a new light particle, the axion, exists.

3. Finally, it might be that CP is a valid symmetry microscopically, and is spontaneously broken in such a way that $\overline{\theta}$ naturally turns out to be small.

Concerning the first option, it has been realized since early instanton studies that if one or more quark bare masses vanishes, then instanton effects, including instanton contributions to electric dipole moments, vanish. The conceivably realistic way to implement this in the real world is to assume that $m_u = 0$. Though this seems at first to be inconsistent with estimates of quark mass ratios from hadron phenomenology [9], it is not clear that this is so once one allows [10, 11] for the fact that the combination $m_d m_s / \Lambda_{\mathrm{QCD}}$ has the same quantum numbers as m_u. However, recent analyses from lattice gauge theory claim to take such questions into account and to show that $m_u \neq 0$ [12].

The second option originated in the work of Peccei and Quinn [5], who postulated that there is a U(1) symmetry (often called a PQ symmetry), that is conserved except for gauge anomalies, and acts by chiral rotations on one or more quarks. If unbroken, such a symmetry would imply the vanishing of some quark masses, leading us back to the solution to the strong CP problem with $m_u = 0$. It was assumed in [5], however, that the quarks get masses from coupling to a scalar field that carries the U(1) symmetry and has an expectation value. In this case, as was soon noted [13, 14], the PQ breaking leads to a light spin zero particle — the "axion" — that gets mass only from QCD instanton effects. It originally was assumed that PQ breaking was tied to electroweak symmetry breaking, but subsequently models were constructed [15, 16] in which PQ breaking occurs at a much higher scale. This leads to an "invisible" axion that interacts extremely weakly. Experiments that might detect such an axion were proposed in [17]. For a review of the current status of searches for axionic dark matter, see [18]. Recently, significant limits have also come in searches for axions from the sun [19].

The third idea was developed in [20–22]. There is no problem in beginning with an underlying CP-invariant model, so that the bare value of θ vanishes, and then breaking CP

spontaneously. However, the effective θ measured at low energies receives a contribution from the phase of the determinant of the quark mass matrix. The quark mass matrix is complex (as we know from studies of CP violation in weak interactions). The trick in using an underlying CP symmetry to make $\overline{\theta}$ small is to ensure that the determinant of the quark mass matrix is real and positive (to sufficient accuracy), although that matrix is complex. This is a little delicate, but can be done in a technically natural, and even fairly elegant, way [20 – 24].

In this paper, we consider primarily the axion hypothesis. As explained in early reviews [25, 26], in addition to laboratory constraints if F_a is close to the weak scale, this hypothesis is subject to a large variety of astrophysical constraints. If the axion coupling parameter F_a (see section 2 for its definition) is less than about $10^9 \, \mathrm{GeV}$, then the axion couples too strongly and too many axions are produced in various astrophysical environments, causing red giants to cool too rapidly, for example. There is also astrophysical trouble if the axion coupling is too *weak*, in other words if F_a is too large [27 – 29]. With standard cosmological assumptions, if F_a is greater than about $10^{12} \, \mathrm{GeV}$, the early universe produces too much axionic dark matter relative to what we see.

Since the upper bound on F_a leads to some tension with string theory, we will describe it in more detail. It is assumed that the early universe starts out at very high temperatures with a random value of the axion field. It is hard, according to the argument, for the initial value of the axion field to be anything but random, since the "correct" value of the axion field which minimizes the energy depends on the phases of the light quark masses, and these are irrelevant in the very early universe. (In addition, inflationary fluctuations might have randomized the initial value of the axion field.) The axion field has a potential energy, which at low temperatures is of order $F_\pi^2 m_\pi^2$, that is determined by QCD effects and is irrelevant in the very early universe. (To evaluate the bound on F_a precisely, one must take into account the temperature dependence of the effective potential.) The natural frequency of oscillation of the axion field is of order $F_\pi m_\pi / F_a$. Once the universe cools enough so that this frequency exceeds the Hubble constant, the axion field begins to oscillate in its potential. The oscillations describe a bose-condensed ensemble of nonrelativistic axion particles, and once the field begins to oscillate, the number density of axion particles diminishes as the universe expands. The larger is F_a, the later the oscillations begin, and the greater is the energy density of axions at the end. For F_a greater than about $10^{12} \, \mathrm{GeV}$, the energy density of axions in today's universe, calculated on these assumptions, would exceed the dark matter density that we actually observe. The net effect of the astrophysical and cosmological bounds is to place F_a in a range from about 10^9 to $10^{12} \, \mathrm{GeV}$ — smaller, as we will discuss, than is natural in many string models.

An obvious question about the axion hypothesis is how natural it really is. Why introduce a global PQ "symmetry" if it is not actually a symmetry? What is the sense in constraining a theory so that the classical Lagrangian possesses a certain symmetry if the symmetry is actually anomalous?[1]

---

[1] A similar question can be asked about the hypothesis that $m_u = 0$, as QCD has no additional symmetries when $m_u = 0$. Nevertheless, technically natural models that lead to a solution of the strong CP problem via

It could be argued that the best evidence that PQ "symmetries" are natural comes from string theory, which produces them without any contrivance. Soon after the discovery of the Green-Schwarz anomaly cancellation mechanism [30] expanded the possible scope of string phenomenology, it was recognized [31] that the terms in the low energy effective action that lead to anomaly cancellation also cause certain light string modes to behave as axions (for reviews, see chapter 14.3.2 of [32], or section 7.6 of [25]). In certain cases, some of the would-be axions also get masses from a Higgs mechanism, and one is left over with a global PQ symmetry. These statements can be unified by saying that the string compactifications always generate PQ symmetries, often spontaneously broken at the string scale and producing axions, but sometimes unbroken.

This convincingly shows that axions and PQ symmetries are natural, but in the original models, the resulting axion mass parameter F_a was too large for the cosmological bounds. This problem motivated a number of early contributions. In [33], it was proposed to circumvent the problem by taking an *unbroken* PQ symmetry from the string and spontaneously breaking it at lower energies, producing an axion of smaller F_a (but less simply described as a string mode than the axions found in [31]). In [34], it was proposed that, for anthropic reasons, the usual cosmological bound on F_a might be invalid, avoiding the contradiction of this bound with most of the string models. The clash between the cosmological bound on F_a and the simplest string theory predictions for it was also emphasized in the review article [25].

More recent developments have led to the emergence of many possible new string-based models of particle physics. In this paper, we will reassess the string theory predictions for F_a in the light of these new developments. We first reconsider the heterotic string. The value of F_a for the "model-independent" axion of the weakly coupled heterotic string has been computed in [35] and more precisely recently in [36]. The value is $F_a = 1.1 \times 10^{16}\,\mathrm{GeV}$ if one assumes the usual running of α_s up to the string scale. We work out analogous predictions for the "model-dependent" heterotic string axions. We then go on to consider other models, such as strongly coupled $E_8 \times E_8$ heterotic strings (axions were previously discussed in these models in [37]), M-theory on a manifold of G_2 holonomy, and models based on D-branes and orientifolds.

A general conclusion is that, in many models, it is difficult to push F_a drastically below $1.1 \times 10^{16}\,\mathrm{GeV}$ and easier to increase it closer to the reduced Planck mass, $\sqrt{\frac{1}{8\pi G_N}} \sim 2.4 \times 10^{18}\,\mathrm{GeV}$. (It has been argued earlier that it is hard to make F_a — or its analog for other light spin zero modes — much *higher* than the Planck scale [38]. We do not

---

$m_u = 0$ do exist. For an example, see [11]. Incidentally, m_u can be rigorously measured, in principle, from certain OPE coefficients that violate the chiral symmetry of up quarks (violation of the same symmetry by instantons is softer at short distances); an example of a precise definition of m_u is as follows. Take the current operator $J = \overline{u}\gamma_\mu u$ and let S be the scalar operator $S = \overline{u}u$. In the operator product expansion $J(x)J(0) \sim \cdots + f(x)S(0) + \cdots$, the operator S appears with a nonzero coefficient function $f(x)$ whether m_u is zero or not. One can define m_u as the coefficient of $1/|x|^2$ in $f(x)$, or more precisely, as the coefficient of $(-\ln|x|)^a/|x|^2$, where the exponent a comes from the usual one-loop anomalous dimension of the operator S. If $m_u = 0$, then $f(x)$ vanishes in perturbation theory (but receives instanton contributions) and is less singular as x goes to zero than $(-\ln|x|)^a/|x|^2$.

have anything new to say about this.) Examples include most models with GUT-like phenomenology, and also models in which QCD gauge fields are supported on $D3$-branes. On the other hand, there are also models in which F_a can be lower. One early approach [33] to lowering F_a below the GUT scale remains potentially valid from a modern point of view and can be compatible with GUT phenomenology, though possibly not with low energy supersymmetry. If one is willing to abandon GUT phenomenology, there is another possibility, in which QCD gauge fields are supported on a "vanishing cycle." We discuss various examples, beginning in section 4. These models have string or Kaluza-Klein scales well below the Planck scale, and hence their very early cosmology may be exotic. Related ideas have been discussed in [39] (which appeared while the present paper was in gestation).

From a modern point of view, PQ symmetries generated from string theory are always explicitly violated by instantons of some kind (in addition to the low energy QCD instantons); candidates include worldsheet instantons [40], brane instantons [41], gauge instantons from other factors of the gauge group, and gravitational instantons. Such effects, of course, can be exponentially small. One of the necessary conditions for solving the strong CP problem via string-derived instantons is that explicit breaking of the relevant PQ symmetry by non-QCD effects must be much smaller (at least 10^{10} times smaller, to make $\overline{\theta}$ small enough) than breaking due to QCD instantons. Thus, in the various models, in addition to estimating F_a, we also estimate the actions of the relevant instantons. Suppressing the instantons is a significant constraint on models and in many cases favors having supersymmetry survive at least somewhat below the GUT scale. (This is the main reason that supersymmetry enters our discussion. Though for convenience we consider compactification manifolds that preserve supersymmetry, this will not play an essential role in most of the analysis. The Barr mechanism [33] may be an exception.)

What about other solutions to the strong CP problem in string theory? If we leave aside lattice gauge theory evidence [12] that $m_u \neq 0$, one could try to embed $m_u = 0$ in string theory by leaving a string-based PQ symmetry *unbroken*. The hard part would probably be to ensure that the PQ symmetry enforces $m_u = 0$ without setting to zero any other quark or charged lepton masses. As for the third approach of explaining the strong CP problem by assuming that CP is spontaneously broken in a way that leaves the determinant of the quark mass matrix real and positive, string theory passes the first hurdle in that in almost all string theory compactifications, there is a locus in moduli space at which CP is unbroken, and thus CP can be interpreted as a spontaneously broken symmetry. (For example, see section 16.5.1 in [32].) This still leaves much ground to cover, and we do not know if a mechanism such as that of [21, 33] can be embedded in string theory.

As we have already noted, many string models give F_a too large for the standard cosmological bounds, and in fact the value $F_a = 1.1 \times 10^{16}$ GeV first given in [35, 36], or something relatively near it, arises in a number of different models. This value corresponds to an axion frequency close to 130 KHz. Most models we consider lead to F_a in the range from roughly 10^{15} GeV to the reduced Planck mass. At the reduced Planck mass, the axion frequency is a little less than 1 KHz.

While these values clash with standard cosmological reasoning, proposals have been made that would relax the cosmological bounds. One proposal involves anthropic argu-

ments [34]; others involve arranging so that QCD was actually strongly coupled in the early universe [42], late entropy production due to particle decays [43], or drastically reducing the energy scale of inflation [44, 45]. A variant of the last proposal (stressed to us by P. Steinhardt) is the cyclic model of the universe [46], in which the universe might never reach the high temperatures needed to create an excess of axionic dark matter. The anthropic proposal would lead one to expect that axionic dark matter would be significant in the universe, a point made in [34] and reconsidered recently [47]. A model with F_a large also has to survive certain other cosmological constraints [36]. At any rate, an experiment capable of finding or excluding axionic dark matter in the most relevant frequency range – from about 1 KHz to a few MHz — would greatly clarify things. It has been suggested [48, 49] that this might be accomplished with an experiment using LC circuits.

2. General remarks

The QCD instanton number can be written in many equivalent ways. It is

$$N = \frac{1}{32\pi^2} \int d^4x \epsilon^{\mu\nu\alpha\beta} \operatorname{tr} F_{\mu\nu} F_{\alpha\beta} = \frac{1}{16\pi^2} \int d^4x \operatorname{tr} F_{\mu\nu} \widetilde{F}^{\mu\nu}, \tag{2.1}$$

where $\widetilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta} F^{\alpha\beta}$. We use the sign convention $\epsilon^{0123} = 1$. The trace is taken in the three-dimensional representation of SU(3). Alternatively, if we introduce the two-form $F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu$, then

$$N = \frac{1}{8\pi^2} \int \operatorname{tr} F \wedge F. \tag{2.2}$$

Finally, if we write $F_{\mu\nu} = F_{\mu\nu}^a T_a$, where T_a is a basis of the Lie algebra, normalized to $\operatorname{tr} T_a T_b = \frac{1}{2}\delta_{ab}$, then

$$N = \frac{1}{64\pi^2} \int d^4x \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu}^a F_{\alpha\beta}^a. \tag{2.3}$$

In writing the instanton number, we normalize the gauge fields so that the covariant derivative $D_\mu = \partial_\mu + iA_\mu$ is independent of the gauge coupling g, which instead appears in the gauge kinetic energy

$$-\frac{1}{2g^2} \int d^4x \operatorname{tr} F_{\mu\nu} F^{\mu\nu} = -\frac{1}{4g^2} \int d^4x F_{\mu\nu}^a F^{\mu\nu\, a}. \tag{2.4}$$

(We write the action in Lorentz metric with signature $-+++$.) It follows that the action of an instanton, that is a field with $F = \widetilde{F}$ and $N = 1$, is

$$I = \frac{8\pi^2}{g^2} = \frac{2\pi}{\alpha_s}, \tag{2.5}$$

where $\alpha_s = g^2/4\pi$.

The axion is a spin zero field a with a PQ shift symmetry $a \to a + \text{constant}$ that is broken only, or primarily, by the effects of QCD instantons. The coupling of a to QCD instantons is

$$\frac{r}{32\pi^2} \int d^4x \, a \, \epsilon^{\mu\nu\alpha\beta} \operatorname{tr} F_{\mu\nu} F_{\alpha\beta}, \tag{2.6}$$

where r is a constant. In all of the string and M-theory models, the axion turns out to be a periodic variable; we normalize it so that the period is 2π. As N is an integer, the action is well-defined mod 2π if r is an integer, as is the case in the string models.

The axion field also has a kinetic energy, $\frac{1}{2}F_a^2 \partial_\mu a \partial^\mu a$, for some constant F_a. The low energy effective action of the axion thus includes the terms

$$\Delta I = \int d^4x \left(-\frac{F_a^2}{2} \partial_\mu a \partial^\mu a + \frac{ra}{32\pi^2} \epsilon^{\mu\nu\alpha\beta} \operatorname{tr} F_{\mu\nu} F_{\alpha\beta} \right). \tag{2.7}$$

F_a is called the axion decay constant or coupling parameter, because axion couplings are proportional to $1/F_a$. For example, if we introduce a rescaled axion field $\widetilde{a} = F_a a$, then the kinetic energy becomes canonical, and the axion coupling is proportional to $1/F_a$:

$$\Delta I = \int d^4x \left(-\frac{1}{2} \partial_\mu \widetilde{a} \partial^\mu \widetilde{a} + \frac{r\widetilde{a}}{32\pi^2 F_a} \epsilon^{\mu\nu\alpha\beta} \operatorname{tr} F_{\mu\nu} F_{\alpha\beta} \right). \tag{2.8}$$

As explained in the introduction, the main result of the present paper is that in a wide range of string models, F_a is in the range from roughly the GUT scale to the reduced Planck scale, and thus is above the range favored by the usual cosmological bounds.

To see that existence of an axion solves the strong CP problem, first note that if an axion is present, the physics is independent of the QCD vacuum angle θ. The θ-dependence of the action is an additional term,

$$\frac{\theta}{32\pi^2} \int d^4x \, \epsilon^{\mu\nu\alpha\beta} \operatorname{tr} F_{\mu\nu} F_{\alpha\beta}, \tag{2.9}$$

and if an axion field a is present, this term can be eliminated by shifting a by a constant, $a \to a - \theta/r$. In effect, the existence of an axion promotes θ to a dynamical field ra; the vacuum expectation value of a must be determined to minimize the energy.

To compute the potential energy as a function of a, one must know how to calculate the vacuum energy of QCD as a function of θ. Because the up and down quark masses are so small, this can conveniently be done using current algebra, as reviewed in [50], section 23.6. In the two flavor case (the u and d quarks are so much lighter than the s quark that the latter can be neglected), we describe low energy pion physics by an SU(2)-valued field U with an effective Lagrangian

$$L = -\frac{F_\pi^2}{16} \operatorname{tr} \partial_\mu U \partial^\mu U^{-1} + \frac{v}{2} \operatorname{tr} \left(MU + \overline{M}U^{-1} \right). \tag{2.10}$$

Here, M is the quark mass matrix, which at $\theta = 0$ we can take to be $M = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix}$, with m_u, m_d being real and positive; also, to get the right pion mass and couplings, $F_\pi = 184\,\mathrm{MeV}$, and $v(m_u + m_d) = F_\pi^2 m_\pi^2/4$. One can include θ by replacing (for example) $m_u \to e^{i\theta}m_u$. So, upon promoting θ to a field ra, we take

$$M = \begin{pmatrix} m_u e^{ira} & 0 \\ 0 & m_d \end{pmatrix}. \tag{2.11}$$

The effective potential for the light field a is obtained by minimizing the potential $V(U,a) = \frac{v}{2} \operatorname{tr}(MU + \overline{M}U^{-1})$ as a function of U for fixed a. The minimum energy is at $a = 0$; the

term quadratic in a turns out to be

$$V(a) = \frac{a^2}{8} r^2 F_\pi^2 m_\pi^2 \frac{m_u m_d}{(m_u + m_d)^2}.$$ (2.12)

So allowing for the normalization of the axion kinetic energy $\frac{1}{2} \int d^4x F_a^2 (\partial_\mu a)^2$, the axion mass is

$$m_a = \frac{r F_\pi m_\pi}{2 F_a} \frac{\sqrt{m_u m_d}}{m_u + m_d}.$$ (2.13)

With the estimate [9] that $m_u/m_d \cong 1/1.8$, one has

$$m_a \cong 5.4 \times 10^{-10} \, \text{eV} \cdot \frac{1.1 \times 10^{16} \, \text{GeV}}{F_a/r}.$$ (2.14)

The angular frequency of axion oscillations is $\omega = m_a c^2 / \hbar$, and the ordinary frequency is

$$\nu = \frac{\omega}{2\pi} = 130 \, \text{KHz} \cdot \frac{1.1 \times 10^{16} \, \text{GeV}}{F_a/r}.$$ (2.15)

We will see that in many string models, F_a/r is naturally fairly near 1.1×10^{16} GeV, the value found [35, 36] for the model-independent axion of perturbative heterotic strings.

In string models, there are inevitably, apart from low energy QCD instantons, some other instantons — such as worldsheet or membrane instantons — that violate the PQ symmetry. Suppose that the axion couples to a stringy instanton of action S_{inst} with a natural mass scale M. If there is no suppression due to supersymmetry, the instanton will generate a potential of the general form $-M^4 \exp(-S_{\text{inst}}) \exp(i(a + \psi))$ for some phase ψ. Since we defined the axion field so that the low energy QCD contribution to the axion potential is minimized at $a = 0$, ψ really arises from a mismatch in phase between the high scale instantons and the analogous phase in the low energy contribution, which is affected by things such as the light quark masses. Lacking a special theory of the light quark masses, we interpret ψ as an arbitrary phase that is not likely to be particularly small. The axion potential induced by such instantons and anti-instantons is of order

$$\widetilde{V}(a) = -2M^4 \exp(-S_{\text{inst}}) \cos(a + \psi).$$ (2.16)

Upon minimizing $V(a) + \widetilde{V}(a)$, we get in order of magnitude

$$r^2 a \sim M^4 \exp(-S_{\text{inst}})/F_\pi^2 m_\pi^2.$$ (2.17)

Since ra is the effective QCD theta angle, we require $|ra| < 10^{-10}$ to agree with experimental constraints, so we need

$$\exp(-S_{\text{inst}}) < 10^{-10} r \frac{F_\pi^2 m_\pi^2}{M^4}.$$ (2.18)

This is a severe constraint. For example, we if set M to the reduced Planck mass 2.4×10^{18} GeV, we need

$$S_{\text{inst}} > 200.$$ (2.19)

This is a significant constraint on models, as has been stressed in [44] (where it was argued that this constraint favors the strongly coupled heterotic string over the perturbative heterotic string) and as we will discuss further. Many models with large extra dimensions and a reduced string scale do not help much in reducing F_a, as we will see. But by reducing the mass M in (2.18), they do relax the requirements for the instanton action.

Those requirements are also relaxed if supersymmetry survives to a scale lower than M. In this case, instead of generating a potential proportional to the real part of $M^4 \exp(-S_{\text{inst}} + i(a + \psi))$, the instanton may generate a contribution to the superpotential proportional to $W_0 = M^3 \exp(-S_{\text{inst}} + i(a + \psi))$. When we evaluate the ordinary potential $V = |DW/D\Phi|^2 - 3G_N|W|^2$, if we simply set $W = W_0$, the a-dependence will cancel. An a-dependence of V can arise, for example, from an interference of the one-instanton term with a contribution to W from some other source. We suppose that this other contribution breaks supersymmetry at some scale μ, with $DW/D\Phi \sim \mu^2$. Then the ordinary potential contains terms that are roughly $M^2\mu^2 \exp(-S_{\text{inst}}) \cos(a + \psi)$, and (2.18) is replaced by

$$\exp(-S_{\text{inst}}) < 10^{-10} \frac{F_\pi^2 m_\pi^2}{M^2 \mu^2}. \tag{2.20}$$

If μ is even a few orders of magnitude below M, or M below the reduced Planck mass, this makes life easier for many models. Actually, supersymmetry can lead to more suppression of high scale instanton effects than we have just described if the instantons have fermion zero modes beyond those required by supersymmetry and contribute not to the superpotential but to higher order chiral operators, as considered in [51].

One important point is that the instanton that gives the dominant contribution to the axion mass may not entirely break the shift symmetry of the axion. It may reduce this symmetry to an n-fold discrete symmetry, for some integer n. For example, for an axion that actually solves the strong CP problem, the dominant instanton is an ordinary QCD instanton, and the integer n coincides with the integer r of the axion coupling in eq. (2.6). The shift symmetries that are not explicitly broken by the instanton are spontaneously broken by the axion expectation value. That will result, in this approximation, in having n degenerate vacua differing by the value of the axion field. However, in the context of string theory, one expects that this degeneracy will always be lifted at a lower energy by some other, subdominant, type of instanton. The full collection of string theory instantons (including worldsheet instantons, gauge and gravitational instantons, wrapped branes, etc.) is expected to fully break the axion symmetries. This statement roughly means that the theory has the full set of branes and other topological defects allowed by the periodicity of the axions, just as [52] it has the full set of magnetic charges allowed by Dirac quantization and the values of electric charges.

Apart from its coupling to the QCD instanton density, the axion may have additional couplings to Standard Model fields. The PQ shift symmetry allows the axion to have arbitrary derivative couplings to quarks and leptons. In addition, it may have couplings just analogous to (2.6) to the "instanton densities" of other gauge fields and gravity. The coupling to electromagnetism is of particular importance, since it is the basis for axion searches [17 – 19]. For example, in an SU(5) grand unified theory, the electromagnetic field

$f_{\mu\nu}$ appears in the underlying SU(5) gauge theory via an ansatz

$$F_{\mu\nu} = f_{\mu\nu} \begin{pmatrix} -1/3 & & & & \\ & -1/3 & & & \\ & & -1/3 & & \\ & & & 1 & \\ & & & & 0 \end{pmatrix}. \qquad (2.21)$$

If we extend the trace in (2.8) to a trace over the fundamental representation of SU(5), we get $\operatorname{tr} F_{\mu\nu} F_{\alpha\beta} = (4/3) f_{\mu\nu} f_{\alpha\beta} + \cdots$, so in an SU(5) grand unified theory, the coupling of the axion to electromagnetism is

$$\frac{4r}{3} \frac{1}{32\pi^2 F_a} \int d^4x \, \widetilde{a} \epsilon^{\mu\nu\alpha\beta} f_{\mu\nu} f_{\alpha\beta}. \qquad (2.22)$$

It is convenient to rewrite this in terms of a canonically normalized electromagnetic field. In our convention, as in (2.4), the gauge coupling appears as a constant multiplying the action, and the electromagnetic action is $\int d^4x f_{\mu\nu} f^{\mu\nu}/4e^2$, with e the charge of the electron. In terms of the conventionally normalized electromagnetic tensor $F_{\mu\nu}^{\mathrm{em}} = f_{\mu\nu}/e$ with kinetic energy $(F_{\mu\nu}^{\mathrm{em}})^2/4$, the coupling becomes

$$\frac{4r}{3} \frac{e^2}{32\pi^2 F_a} \int d^4x \, \widetilde{a} \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu}^{\mathrm{em}} F_{\alpha\beta}^{\mathrm{em}} = -\frac{4r\alpha}{3\pi F_a} \int d^4x \, \widetilde{a} \vec{E} \cdot \vec{B}. \qquad (2.23)$$

Here $E_i = F_{0i}^{\mathrm{em}}$ and $B_i = \frac{1}{2}\epsilon_{ijk} F_{jk}^{\mathrm{em}}$ are the usual electric and magnetic fields, and $\alpha = e^2/4\pi\hbar c$. This formula needs to be corrected, however, as we explain in section 9, because of mixing between the axion and the π^0 meson.

In deriving the specific value of the coupling in (2.23), we considered SU(5)-like GUT's, but a similar $\widetilde{a}\vec{E}\cdot\vec{B}$ coupling (possibly with a slightly different coefficient) arises in virtually all string-based models that solve the strong CP problem via an axion. For example, in heterotic string models, the axion couplings to strong, weak, and electromagnetic instanton densities only depend on the "levels" of the current algebra (or the embedding in E_8), so with a level 1 embedding of the Standard Model, the formula (2.23) will hold (with $r = 1$), even if there is no four-dimensional or even ten-dimensional field theoretic unification. (For gauge coupling constants, the analogous point is made in [53].) The same formula holds, most commonly with $r = 1$, in other string-derived models that incorporate the usual GUT relations for gauge coupling unification, even if there is no four-dimensional GUT. In models that are not unified, the couplings may be somewhat different.

General approach. Each string or M-theory model we consider is characterized by an asymptotic expansion parameter — g_s in the case of string theory, $1/R$ in the case of M-theory on a G_2 manifold of radius R, and so on. The asymptotic expansion is valid when the relevant parameter is small, and we expect it to give the correct order of magnitude when the parameter is of order one. To assess how small F_a can be, we calculate in the region where the expansion parameter is small and then extrapolate to the region where g_s, $1/R$, etc., is of order 1. The region with (say) $g_s \gg 1$ is not accessible in this way, but hopefully

is accessible via a dual asymptotic expansion valid for large g_s. Hence, to the extent that the asymptotic regions that we consider are representative, we should get a good overview of axion phenomenology, though in the interior of moduli space our computations are only valid qualitatively.

Conventions. In our string theory computations, we will adopt a few conventions that are aimed to minimize the factors of 2π in this paper. We define the string scale by[2] $\ell_s = 2\pi\sqrt{\alpha'}$, and we normalize p-form fields to have integer periods. Thus, our NS B-field is related to the usual one B_{conv} by $B_{\text{conv}} = \ell_s^2 B$. As a result, the coupling of the B-field to the string worldsheet Σ, usually $(i/2\pi\alpha')\int_\Sigma B_{\text{conv}}$, becomes $2\pi i \int_\Sigma B$. The contribution to the action proportional to the area of the worldsheet becomes $(2\pi/\ell_s^2)\int d^2\sigma\sqrt{\det G}$.

We follow [54], eq. (13.3.22), in normalizing the Type II dilaton ϕ_{II} and Type II string coupling $g_{II} = \exp(\phi_{II})$ so that the ten-dimensional gravitational coupling κ (appearing in the Einstein action $\int d^{10}x\sqrt{g}R/2\kappa^2$) is

$$\kappa^2 = \frac{1}{2}(2\pi)^7 g_{II}^2(\alpha')^4 = \frac{g_{II}^2 \ell_s^8}{4\pi}. \tag{2.24}$$

This convention ensures that $SL(2,\mathbb{Z})$ duality acts by $g_{II} \to 1/g_{II}$, with no numerical factor. We will use the same convention (2.24) for Type I and heterotic superstring theories (with, of course, g_{II} replaced by the corresponding string couplings $g_I = \exp(\phi_I)$ and $g_h = \exp(\phi_h)$). This ensures that under heterotic-Type I duality, we have $g_I = 1/g_h$ with no numerical constant. On the other hand, one must be careful of a factor of two in duality between Type I on a torus $\mathbb{T}^n$ and Type II on the orientifold $\mathbb{T}^n/\mathbb{Z}_2$. Note also the definition $\kappa_{10}^2 = \kappa^2/g^2$ in [54], used for all of the string theories,

With our definition $\ell_s = 2\pi\sqrt{\alpha'}$, the tension of a Dp-brane in Type II superstring theory becomes (following [54], eq. (13.2.3))

$$\tau_p = \frac{2\pi}{g_{II}\ell_s^{p+1}}. \tag{2.25}$$

Similarly, to minimize factors of 2π in eleven-dimensional supergravity, we introduce an M-theory length ℓ_{11} by $\ell_{11} = 2\pi/M_{11}$, where M_{11} is defined in terms of Type IIA superstring parameters by eq. (14.4.6) of [54]. In these units, the membrane and fivebrane tensions (in view of [54], eqs. (14.4.11) and (14.4.18)) are

$$\tau_{M2} = \frac{2\pi}{\ell_{11}^3}, \quad \tau_{M5} = \frac{2\pi}{\ell_{11}^6}. \tag{2.26}$$

Moreover, from eq. (14.4.5) of that reference, we have $2\kappa_{11}^2 = \ell_{11}^9/2\pi$. Moreover, we normalize the three-form field C to have integer periods; this means that our C is related

---

[2]Apart from the fact that it eliminates many factors of 2π from the formulas, a piece of evidence that this is a good definition is the following. In compactification on a circle, the self-dual *radius* is $R = \sqrt{\alpha'}$, and the self-dual *circumference* is hence $L = 2\pi\sqrt{\alpha'} = \ell_s$. The circumference — the length of a nontrivial closed geodesic — is an intrinsic measure of the size of the circle, while the radius is only a natural notion if the circle is embedded in a plane.

to the analogous field C' used in [54] by $C = C'/\ell_{11}^3$. The action of eleven-dimensional supergravity, from eq. (12.1.1), becomes

$$S_{11} = 2\pi \int \left(\frac{1}{\ell_{11}^9} d^{11}x \sqrt{-g} R - \frac{1}{2\ell_{11}^3} G \wedge \star G - \frac{1}{6} C \wedge G \wedge G \right). \qquad (2.27)$$

The Hodge star operator $\star$ is defined so that if W is a p-form, then

$$W \wedge \star W = \frac{\sqrt{-g}}{p!} W_{\mu_1 \mu_2 \cdots \mu_p} W^{\mu_1 \mu_2 \cdots \mu_p}.$$

We write tr for the trace in the fundamental representation of $SU(N)$, tr for the trace in the fundamental representation of $SO(M)$, and Tr for the trace in the adjoint representation of E_8 or $SO(32)$. If $SU(N)$ is embedded in $SO(2N)$, then tr $= 2\,$tr. For $SO(32)$, one has tr $=$ Tr $/30$, and for E_8 one defines tr $=$ Tr $/30$. When we pick a basis of the Lie algebra, we require $\operatorname{tr} t^a t^b = \delta^{ab}$ and hence (for $SU(N)$) $\operatorname{tr} t^a t^b = \delta^{ab}/2$.

In M-theory on a manifold with a boundary M^{10}, E_8 gauge fields appear on the boundary. The gauge action is

$$S_{YM} = -\frac{1}{4(2\pi)\ell_{11}^6} \int_{M^{10}} \operatorname{tr} F \wedge \star F. \qquad (2.28)$$

3. Axions in weakly coupled heterotic string theory

The relevant part of the ten-dimensional low energy Lagrangian of the heterotic string is, from eq. (12.1.39) of [54],

$$\begin{aligned} L &= \frac{1}{2\kappa_{10}^2} \sqrt{-g} R - \frac{1}{4\kappa_{10}^2} H \wedge \star H - \frac{\alpha'}{8\kappa_{10}^2} \operatorname{tr} F \wedge \star F \\ &= \frac{2\pi}{g_s^2 \ell_s^8} \sqrt{-g} R - \frac{2\pi}{g_s^2 \ell_s^4} \frac{1}{2} H \wedge \star H - \frac{1}{4(2\pi) g_s^2 \ell_s^6} \operatorname{tr} F \wedge \star F. \end{aligned} \qquad (3.1)$$

Here R is the Ricci scalar, H the field strength of the two-form field B, and F the $E_8 \times E_8$ or $SO(32)$ curvature; we have used conventions explained at the end of the section 2 together with the relation $\kappa_{10}^2/g_{10}^2 = \alpha'/4$ from eq. (12.3.36) of [54].

To reduce to four dimensions, we compactify on a six-manifold Z (not necessarily Calabi-Yau) with volume V_Z. The four-dimensional spacetime (which might be Minkowski spacetime) we call M. The relevant terms in the four-dimensional effective action include

$$S = \frac{M_P^2}{2} \int d^4x (-g)^{1/2} R - \frac{1}{4g_{YM}^2} \int d^4x \sqrt{-g}\, \operatorname{tr} F_{\mu\nu} F^{\mu\nu} - \frac{2\pi V_Z}{g_s^2 l_s^4} \int \left(\frac{1}{2} H \wedge \star H \right), \quad (3.2)$$

where the four-dimensional reduced Planck mass M_P and Yang-Mills coupling g_{YM} are

$$M_P^2 = 4\pi \frac{V_Z}{g_s^2 \ell_s^8} \qquad (3.3)$$

and

$$g_{YM}^2 = 4\pi \frac{g_s^2 \ell_s^6}{V_Z}. \qquad (3.4)$$

So $\alpha_{YM} = g_{YM}^2/4\pi$ is given by

$$\alpha_{YM} = \frac{g_s^2 \ell_s^6}{V_Z}. \tag{3.5}$$

If we assume the usual level one embedding of the Standard Model gauge fields in $E_8 \times E_8$, then we can identify α_{YM} with α_{GUT}, the unified gauge coupling at the string (or "GUT") scale. More generally, if we make a level k embedding of the Standard Model in the heterotic string (for discussion, see [55]), then tr $F_{\mu\nu}F^{\mu\nu}$, when evaluated for Standard Model gauge fields, has an extra factor of k, which reduces α_{GUT} by a factor of k. Thus we have in general

$$\alpha_{\mathrm{GUT}} = \frac{\alpha_{YM}}{k} = \frac{g_s^2 \ell_s^6}{k V_Z}. \tag{3.6}$$

In general, it may be that not all factors in the Standard Model gauge theory have the same k. To determine the axion mass and frequency, we really need the coupling of the axion to QCD. So we let k denote the level of the color SU(3) embedding in $E_8 \times E_8$. In all models considered in this paper, we define α_C to be the strong coupling constant α_s evaluated at the string scale or compactification scale – whatever is the scale at which four-dimensional field theory breaks down — whether or not there is a unified embedding of the Standard Model. Of course, for the coupling (2.23) of the axion to electromagnetism, we must interpret k to be the level of the current algebra embedding of the electromagnetic U(1).

The string scale $M_s = 1/\ell_s$ can be evaluated from the above formulas to be $M_s = (k\alpha_G/4\pi)^{1/2} M_P$. With the usual phenomenological estimate $\alpha_C \sim 1/25$, we get the usual perturbative heterotic string scale $M_s \sim M_P \sqrt{k}/18$.

Axions arise from the B-field of the heterotic string. The components $B_{\mu\nu}$ with μ and ν tangent to Minkowski spacetime (and constant on Z) can be dualized to make an axion. Since the existence and properties of this axion do not depend very much on the choice of Z, it has been traditionally called the model-independent axion of the heterotic string. Though the name seems somewhat anachronistic in an age in which there are many other string-based models, we will see that something rather like this mode does exist in many other asymptotic limits of string theory. Zero modes of $B_{\mu\nu}$ with μ and ν tangent to the compact manifold Z also have axion-like couplings and are traditionally known as model-dependent axions.

Model-independent axion. We consider first the model-independent axion. The Bianchi identity for the gauge-invariant field strength of H is

$$dH = \frac{1}{16\pi^2} \left(\mathrm{tr}\, R \wedge R - \mathrm{tr}\, F \wedge F \right). \tag{3.7}$$

(For example, the normalization can be extracted from eq. (12.1.40) of [54], bearing in mind that our H is ℓ_s^2 times the H-field used there.) Now focus on the modes that are constant on Z with all indices tangent to four-dimensional spacetime. The four-dimensional component of the B-field can be dualized by introducing a field a that is a Lagrange multiplier for the Bianchi identity, the coupling being $\int a \left(dH + \frac{1}{16\pi^2} \left(\mathrm{tr}\, F \wedge F - \mathrm{tr}\, R \wedge R \right) \right)$. Including also

the B-field kinetic energy from (3.2), the action is

$$-\frac{2\pi V_Z}{g_s^2 l_s^4} \int d^4x \, \frac{1}{2} H \wedge \star H + \int a \left(dH + \frac{1}{16\pi^2} \left(\operatorname{tr} F \wedge F - \operatorname{tr} R \wedge R \right) \right) \qquad (3.8)$$

Here H is an independent field variable (which can be expressed in terms of B if one integrates first over a to impose the Bianchi identity). As H has integer periods, a should have period 2π.[3] Instead, we integrate out H to get an effective action for a:

$$S(a) = \frac{g_s^2 \ell_s^4}{2\pi V_Z} \int d^4x \left(-\frac{1}{2} \partial_\mu a \partial^\mu a \right) + \int a \frac{1}{16\pi^2} \left(\operatorname{tr} F \wedge F - \operatorname{tr} R \wedge R \right). \qquad (3.9)$$

Since $\operatorname{tr} F \wedge F = 2k \operatorname{tr} F \wedge F$, we see that for this particular axion, the integer r characterizing the axionic coupling is equal to the current algebra level k. For future reference, let us note that the effect of the dualization from H to a is that F_a^2 is simply the inverse of the coefficient of $(1/2)H \wedge \star H$ in the four-dimensional kinetic energy of H.

So we can read off the axion decay constant:

$$F_a = \frac{g_s^2 \ell_s^2}{\sqrt{2\pi V_Z}} = \frac{k\alpha_G}{2\pi} \frac{M_P}{\sqrt{2}}. \qquad (3.10)$$

The axion couplings are proportional, according to (2.8), to F_a/k, which is

$$\frac{F_a}{k} = \frac{\alpha_C M_P}{2\pi\sqrt{2}}. \qquad (3.11)$$

If we take $\alpha_C = 1/25$, this gives $F_a/k \approx 1.1 \times 10^{16}\,\mathrm{GeV}$, as in [36]. If the model is not unified or the spectrum of particles contributing to the renormalization group running is not the minimal one, then α_C may have a somewhat different value, leading to a somewhat different value of F_a/k.

In order for the axionic mode whose coupling we have just evaluated to solve the strong CP problem, low energy QCD instantons must be the dominant mechanism that violates the associated PQ symmetry. There are a few other candidates to worry about. Gauge instantons at the string scale have action

$$\frac{2\pi}{\alpha_{YM}} = \frac{2\pi}{k\alpha_C} \sim \frac{157}{k}, \qquad (3.12)$$

by analogy with (2.5). For $k = 1$, this is somewhat below the value 200 that we need according to (2.19) if there is no suppression due to low energy supersymmetry. But it might be satisfactory if because of supersymmetry a formula such as (2.20) is more appropriate, or if for some reason α_{YM} is a bit smaller than $1/25$. For $k > 1$, it would be much harder to suppress the explicit PQ violation adequately.

Because of the $a \operatorname{tr} R \wedge R$ coupling, the PQ symmetry might also be violated by gravitational instantons, if they are relevant to string theory in asymptotically flat spacetime, but we do not know how to estimate their effects. Finally, it might happen that some other

---

[3]One can carry out the duality more precisely, in a way that is valid on a general four-manifold, by adapting the procedure used in [56] for two-dimensional T-duality.

part of the $E_8 \times E_8$ or SO(32) gauge group of the heterotic string becomes strong at an energy above the QCD scale. Then instantons of that group might be the dominant source of PQ symmetry violation.

If one of these sources of PQ violation gives an excessive mass to the mode a, then we need another axion to solve the strong CP problem. Interestingly, there are other candidates, as we now discuss.

Model-dependent axions. Model-dependent heterotic string axions arise from zero modes of the B-field on the compact manifold Z. Let there be $n = \dim H^2(Z, \mathbb{R})$ such zero modes $\beta_1, \ldots, \beta_n$. We normalize them so that

$$\int_{C_j} \beta_i = \delta_{ij}, \tag{3.13}$$

where the C_j are two-cycles representing a basis of $H_2(Z, \mathbb{Z})$ modulo torsion. Then we make an ansatz

$$B = \frac{1}{2\pi} \sum_i \beta_i b_i, \tag{3.14}$$

where b_i are four-dimensional fields. The factor of $1/2\pi$ is included so that the fields b_i have periods 2π, as is conventional for axions. Set

$$\gamma_{ij} = \int_Z \beta_i \wedge *\beta_j. \tag{3.15}$$

By dimensional reduction from the B-field kinetic energy in (3.1), the kinetic energy of the b_i fields in four dimensions comes out to be

$$S_{\mathrm{kin}} = -\frac{1}{2\pi g_s^2 \ell_s^4} \int d^4 x \frac{\gamma_{ij}}{2} \partial_\mu b_i \partial^\mu b_j. \tag{3.16}$$

Dimensional reduction of a *local* action in ten dimensions leads to a four-dimensional effective action in which the modes b_i only have derivative couplings. This results from the underlying gauge invariance $\delta B = d\Lambda$ in ten dimensions, and is why these modes have approximate PQ shift symmetries.

As first observed in [31], these modes acquire axionic couplings from the one-loop couplings that enter the Green-Schwarz anomaly cancellation mechanism. The relevant couplings are

$$\frac{-1}{4(2\pi)^3 4!} \int B \left\{ -\frac{\mathrm{Tr}\, F \wedge F \, \mathrm{tr}\, R \wedge R}{30} + \frac{\mathrm{Tr}\, F^4}{3} - \frac{(\mathrm{Tr}\, F \wedge F)^2}{900} \right\}. \tag{3.17}$$

(We have omitted some purely gravitational terms, which lead to a coupling of the modes b_i to $\mathrm{tr}\, R \wedge R$ in four dimensions.) To proceed further, we consider the $E_8 \times E_8$ heterotic string, embedding the Standard Model in the first E_8, and write tr_1 and tr_2 for traces in the first or second E_8. The qualitative conclusions are not different for the SO(32) heterotic string or if the Standard Model embedding in $E_8 \times E_8$ is more complicated. The couplings in four dimensions of the axion modes to $\mathrm{tr}_1 F \wedge F$ come out to be

$$-\frac{1}{2\pi^2 4!} \sum_i \int_Z \beta_i \left\{ -\frac{\mathrm{tr}\, R \wedge R}{2} + 2\, \mathrm{tr}_1 F \wedge F - \mathrm{tr}_2 F \wedge F \right\} \int_M b_i \frac{\mathrm{tr}_1 F \wedge F}{16\pi^2}. \tag{3.18}$$

– 15 –

Using the Bianchi identity (3.7), one can alternatively write these couplings as

$$-\sum_i \int_Z \beta_i \wedge \frac{1}{16\pi^2} \left(\text{tr}_1 F \wedge F - \frac{1}{2} \text{tr} R \wedge R \right) \int_M b_i \frac{\text{tr}_1 F \wedge F}{16\pi^2}. \tag{3.19}$$

For a reasonably isotropic Z, a typical matrix element of the metric γ defined in (3.15) is of order $V_Z^{1/3}$, and hence a typical linear combination b of the b_i has $F_b \sim V_Z^{1/3}/2\pi g_s^2 \ell_s^4$. If we use (3.3) and (3.6) to eliminate V_Z and ℓ_s in favor of M_P and α_C, we get

$$F_b = \frac{\alpha_C^{1/3} M_P}{2\pi\sqrt{2}k^{1/3}g_s^{2/3}} \gtrsim \frac{\alpha_C^{1/3} M_P}{2\pi\sqrt{2}k^{1/3}}. \tag{3.20}$$

We assume in the last step that $g_s \lesssim 1$, for validity of the computation. For $k \sim 1$, $\alpha_C \sim 1/25$, this gives a typical result $F_b \sim 10^{17}$ GeV.

To orient ourselves to how one might reduce F_b, let us consider the case that $b_2(Z) = 1$, so that there is a single harmonic two-form β and associated axion b. More specifically, we take $Z = C \times Y$ where C is a Riemann surface and Y is a four-manifold, which must be spin as the theory contains fermions. (Of course, in supersymmetric compactifications, Z is generally not such a product and is considerably more complicated.) For the mode b, the integer that appears in the $b \, \text{tr} \, F \wedge F$ coupling to four-dimensional gauge fields is *not* the same as the level k of the current algebra embedding of the Standard Model. Rather, from (3.19), this integer is[4]

$$k' = \frac{1}{16\pi^2} \int_Y \left(\text{tr}_1 F \wedge F - \frac{1}{2} \text{tr} R \wedge R \right). \tag{3.21}$$

We write V_C and V_Y for the volumes of C and Y. Then as $\int_C \beta = 1$, we have

$$\int_Z \beta \wedge \star\beta = V_C^{-1} V_Y = \frac{V_Z}{V_C^2}. \tag{3.22}$$

Taking the b kinetic energy from (3.16) and M_P from (3.3), we get

$$F_b = \frac{\ell_s^2}{2\pi V_C} \frac{M_P}{\sqrt{2}}. \tag{3.23}$$

We may as well assume that $V_C \gtrsim \ell_s^2$ (otherwise we should make a T-duality to a better description in which all dimensions in Z are sub-stringy in scale). So we get an approximate upper bound on F_b,

$$F_b \lesssim \frac{M_P}{2\pi\sqrt{2}} = 2.7 \times 10^{17} \text{ GeV}, \tag{3.24}$$

which is larger than the value for the model-independent axion and slightly larger than the generic estimate (3.20).

---

[4] Despite the factor of $1/2$ multiplying $\text{tr} \, R \wedge R$, this expression is an integer for Y a spin manifold, because the signature of a four-dimensional spin manifold is even and in fact divisible by 16.

To *lower* F_b, we can take V_C large, but there is a limit to how far we can go. From (3.5) and $V_Z = V_C V_Y$, we have

$$\frac{V_C}{\ell_s^2} = \frac{g_s^2}{k\alpha_C} \frac{\ell_s^4}{V_Y}. \tag{3.25}$$

For the model to have a qualitatively correct description as a weakly coupled heterotic string, we require $g_s \lesssim 1$. To have a sensible description in terms of compactification on a manifold $C \times Y$, we require $V_Y \gtrsim \ell_s^4$. If either of these conditions fails, one would want to make a duality transformation to a better description (in the second case, for instance, this would be a T-duality transformation) and proceed from there. So the validity of our approximations requires

$$\frac{V_C}{\ell_s^2} \lesssim \frac{1}{k\alpha_C}. \tag{3.26}$$

Inserting this in (3.23), we find that

$$\frac{F_b}{k'} \gtrsim \frac{k}{k'} \frac{\alpha_C M_P}{2\pi\sqrt{2}}. \tag{3.27}$$

For $k = k'$, this lower bound agrees precisely with the actual value that we computed for the model-independent axion. To minimize F_b/k', we can take $k = 1$ and try to take k' large. Let us see how far we can get if we assume a supersymmetric compactification. A Calabi-Yau three-fold Z cannot quite be a product $C \times Y$. But it can be a fibration, with fibers $Y = \mathrm{K3}$, over $C = \mathbb{CP}^1$. This is good enough to justify the above formulas. We let N_1 and N_2 denote the instanton numbers in the two factors of the $E_8 \times E_8$ gauge bundle over Y. Supersymmetry requires $N_1, N_2 \geq 0$, and the Bianchi identity for H implies that $N_1 + N_2 = 24$ (we use the fact that $24 = (1/16\pi^2) \int_{\mathrm{K3}} \mathrm{tr}\, R \wedge R$ is the Euler characteristic of K3). Moreover, from (3.21), we have $k' = N_1 - 12$, so $|k'| \leq 12$.

So we can get F_b/k' as small as about 10^{15} GeV with this kind of model, by putting all instantons in the same E_8 to get $k' = 12$. (In fact, this is the most traditional type of heterotic string model.) To get F_b as small as this, we need to make V_C as large as possible. There is another virtue in doing so. Because of the coupling $2\pi i \int_\Sigma B$ to the string worldsheet Σ, a worldsheet instanton described by a string wrapping around C explicitly breaks the PQ symmetry of this particular axion. The action of such an instanton is, in our conventions, $I = 2\pi V_C/\ell_s^2$. To suppress explicit PQ violation, we should make V_C large. When the inequality in (3.27) is saturated, we have $I = 2\pi/k\alpha_C$, the same value as for the model-independent axion. Again, for $k = 1$ and with the help of some suppression by low energy supersymmetry, the PQ symmetry might be good enough to solve the strong CP problem.

We introduced this example as a convenient special case, but actually, to get F_b small by making a two-cycle C large, a fibration over C is the natural case to consider. Another approach to making F_b small is to leave V_C/ℓ_s^2 of order 1 and try to make k' very large. An obvious potential problem with this comes from the formula (3.21) for k'. If $V_Y \sim \ell_s^4$ (since larger V_Y makes F_b larger) and k' is large, then F or R is large pointwise and the approximations may be invalid. We consider some examples seeking to make k' large in section 5. But first, we repeat our analysis up to this point for the strongly coupled $E_8 \times E_8$ heterotic string.

4. Heterotic M-theory

To describe the strongly coupled $E_8 \times E_8$ heterotic string, we consider compactification of M-theory on $Z \times I$, where Z is a six-manifold and I is an interval. At each end of the interval I lives an E_8 gauge multiplet. Our M-theory normalizations have been defined in section 2.

If the interval I is relatively short, the metric on $Z \times I$ is approximately a product metric. In this case, we write $\pi\rho$ for the length of I, V_Z for the volume of Z, and $V_7 = V_Z \pi\rho$ for the volume of $Z \times I$. (Lengths and volumes, of course, are now computed using the M-theory metric.) When I becomes long, the metric on $Z \times I$ ceases to be a product metric, and if the instanton numbers are not equal at the two ends of I, then the volume of Z is larger at one end than the other [57]. We have the option of embedding the Standard Model into either of the two ends. In the usual approach to phenomenology with the strongly coupled heterotic string, the Standard Model lives at the end where volume of Z is larger. For now, we assume that this is the case. (The other possibility is treated in section 4.3.) We let V_Z be the volume of Z at the end where the Standard Model lives. For the volume of $Z \times I$, we write $V_7 = V_Z \pi\rho\epsilon$, where ϵ takes account of the shrinking of Z and equals 1 if ρ is small and $1/2$ if the volume of Z varies linearly from V_Z at one end to zero at the other.

By dimensional reduction from (2.27) and (2.28), we find

$$M_P^2 = \frac{4\pi^2 \rho V_Z \epsilon}{\ell_{11}^9}$$

$$\alpha_{YM} = \frac{\ell_{11}^6}{V_Z} \tag{4.1}$$

The value written in (4.1) for $\alpha_{YM} = g_{YM}^2/4\pi = k\alpha_C$ is valid at either end of I as long as one uses the appropriate value of V_Z at that end; since we have assumed that V_Z is the larger of the two volumes, the formula has really been written for the end of I at which the gauge coupling is weaker. If the Standard Model lives at the end with smaller volume of Z, the effects of warping can be significant; we will discuss this case in a later subsection.

4.1 Model-independent axion

We will now compute the coupling parameter F_a of the model-independent axion. The model-independent axion comes from a mode of $G = dC$ with one index tangent to I and the other three tangent to the four-dimensional spacetime M:

$$G = \frac{dx^{11}}{\pi\rho} H. \tag{4.2}$$

With this normalization, H, like G, has integer periods. The kinetic term for the four-dimensional field H comes from Kaluza-Klein reduction of (2.27):

$$-\frac{2V_Z \epsilon}{\rho \ell_{11}^3} \int_M \frac{1}{2} H \wedge \star H. \tag{4.3}$$

As in the discussion of (3.13), H can be dualized to an axion a, with F_a equal to the inverse of the H-field kinetic energy in (4.2):

$$F_a = \sqrt{\frac{\rho \ell_{11}^3}{2 V_Z \epsilon}} = \frac{k \alpha_C}{2\pi \epsilon} \frac{M_P}{\sqrt{2}}. \tag{4.4}$$

For small ρ, $\epsilon = 1$ and the result for F_a agrees with the analogous formula for the weakly coupled heterotic string; for large ρ, F_a may be larger by a factor of 2 because of the factor $1/\epsilon$.

This computation also gives us our first illustration of the fact that in many classes of model, it is hard to significantly reduce F_a by going to a model with large extra dimensions. Heterotic M-theory gives a model with a large fifth dimension if we simply assume that the instanton numbers are equal at the two ends of I, in which case ρ can become very large, limited only by experimental tests of Newton's law of gravity. The above computation gave a result for the coupling parameter F_a/k that can be expressed in terms of the observables α_C and M_P, independent of ρ. (When the instanton numbers are equal at the two ends, the metric on $Z \times I$ is nearly a product and we can set ϵ to 1.) Of course, the reason that this happened is that the relevant axionic mode is a bulk mode, like the graviton.

The discussion of effects other than low energy QCD instantons that violate the shift symmetry of the model-independent axion would be similar to what it was for weak coupling. Obvious candidates are string scale instantons with action $2\pi/k\alpha_C$, gravitational instantons, and instantons in the second E_8.

4.2 Model-dependent axions

The model-dependent axions appear in an ansatz for the C-field:

$$C = \sum_i \beta_i \frac{b_i}{2\pi} \wedge \frac{dx^{11}}{\pi \rho}. \tag{4.5}$$

The β_i are defined as in (3.13), and as before the b_i are periodic scalars with period 2π and approximate shift symmetries.

The couplings of these modes to $\operatorname{tr} F \wedge F$ and $\operatorname{tr} R \wedge R$ in four dimensions are the same as they were for weak string coupling. In fact, those couplings, being integers, are independent of g_s. From the standpoint of heterotic M-theory, these couplings are evaluated by reducing to four dimensions the term $(2\pi/6) \int C \wedge G \wedge G$ in the M-theory effective action, and using the boundary condition [58] that $G|_{x^{11}=0} = (\operatorname{tr}_1 F \wedge F - \frac{1}{2} \operatorname{tr} R \wedge R)/16\pi^2$, $G|_{x^{11}=\pi\rho} = -(\operatorname{tr}_2 F \wedge F - \frac{1}{2} \operatorname{tr} R \wedge R)/16\pi^2)$. Upon making the reduction, one finds, just as for the weakly coupled heterotic string, that the relevant axionic couplings are

$$\sum_i \int_Z \beta_i \wedge \frac{1}{16\pi^2} \left(\operatorname{tr}_1 F \wedge F - \frac{1}{2} \operatorname{tr} R \wedge R \right) \int_Z b_i \operatorname{tr} F \wedge F$$

$$= -\sum_i \int_Z \beta_i \wedge \frac{1}{16\pi^2} \left(\operatorname{tr}_2 F \wedge F - \frac{1}{2} \operatorname{tr} R \wedge R \right) \int_Z b_i \operatorname{tr} F \wedge F. \tag{4.6}$$

The two expressions are equal because of the Bianchi identity, which implies that $[\mathrm{tr}_1 \, F \wedge F] + [\mathrm{tr}_2 \, F \wedge F] - [\mathrm{tr} \, R \wedge R] = 0$ (we write $[\alpha]$ for the cohomology class of a closed differential form α).

The Bianchi identity tells us that $[\mathrm{tr}_1 \, F \wedge F - \frac{1}{2} \mathrm{tr} \, R \wedge R] = 0$ if and only if $[\mathrm{tr}_1 \, F \wedge F] = [\mathrm{tr}_2 \, F \wedge F]$, that is, if and only if the two E_8 bundles over Z are topologically equivalent. Precisely when this is so, it is possible in heterotic M-theory for the eleventh dimension to be extremely long, giving a simple example of a model with a large extra dimension. But in this model, the phrase "model-dependent axions" is a misnomer, as the modes in question do not have axionic couplings. When $[\mathrm{tr}_1 \, F \wedge F] \neq [\mathrm{tr}_2 \, F \wedge F]$, the b_i do have axionic couplings, and the length of the eleventh dimension has an upper bound, which was found in [57] and which we will discuss later.

By dimensional reduction, ignoring the variation of the geometry with x^{11}, the kinetic energy of the fields b_i is

$$\frac{\epsilon}{2\pi^2 \ell_{11}^3 \rho} \int_M \frac{1}{2} \gamma_{ij}^M \, \partial b_i \wedge \star \partial b_j, \tag{4.7}$$

where now

$$\gamma_{ij}^M = \int_Z \beta_i \wedge \star \beta_j. \tag{4.8}$$

The integral is evaluated at the end where the volume of Z is greater.

If Z is fairly isotropic, matrix elements of γ_M are of order $V_Z^{1/3}$ and so the kinetic energy for a generic linear combination of the b_i is proportional to $F_b^2 \sim \epsilon V_Z^{1/3} / 2\pi^2 \ell_{11}^3 \rho$. As will be clear from our evaluations below, this leads to F_b being above the GUT scale (except in the case that the b_i lack the couplings of axions). As in the weakly coupled case, we can try to make F_b small for one mode by taking a two-cycle to be large. We can also try to make F_b small by making ρ large. We will consider first the case of a large two-cycle, and show that it is not really independent as it forces us to try to make ρ large.

So we again consider the case that Z is fibered over a Riemann surface C, with fiber Y. The volumes then factorize: $V_Z = V_C Y_Y$. Moreover, for the axionic mode that is a pullback from C, the relevant component of the metric is

$$\int_Z \beta \wedge \star \beta = \frac{V_Z}{V_C^2}. \tag{4.9}$$

For this axion, we get then

$$F_b = \frac{M_P \ell_{11}^3}{2\pi^2 \sqrt{2} V_C \rho}. \tag{4.10}$$

We can try to make F_b small by making V_C large, but as in the weakly coupled case, there is a limit to how far one can go in that direction. Since $k\alpha_C = \ell_{11}^6 / V_Z = \ell_{11}^6 / V_Y V_C$, we have $1/V_C = k\alpha_C V_Y / \ell_{11}^6$. The M-theory description only makes sense if $V_Y \gtrsim \ell_{11}^4$, so $1/V_C \gtrsim k\alpha_C / \ell_{11}^2$. Hence

$$F_b \gtrsim \frac{k\alpha_C}{2\pi} \frac{M_P}{\sqrt{2}} \frac{\ell_{11}}{\pi \rho}. \tag{4.11}$$

So we cannot make much progress in reducing F_b by going to large V_C, unless ρ is also large.

Let us then discuss what happens upon taking ρ large. If $[\text{tr}_1 F \wedge F] = [\text{tr}_2 F \wedge F]$, then in heterotic M-theory, $\ell_{11}/\pi\rho$ can be many orders of magnitude less than 1, and F_b can be in the range allowed by the usual cosmological arguments, $10^9 \, \text{GeV} \lesssim F_b \lesssim 10^{12} \, \text{GeV}$, or much lower. However, precisely in this case, as we have already discussed, the modes b_i lacks axionic coupling to $\text{tr} \, F \wedge F$.

So if we want to solve the strong CP problem using some of these modes as axions, we have to assume that $[\text{tr}_1 F \wedge F] \neq [\text{tr}_2 F \wedge F]$. That being so, there is an upper bound on $\pi\rho/\ell_{11}$, which [57] cannot be larger than roughly $\alpha_C^{-2/3}$ times a number of order one. (Moreover, reasonable GUT phenomenology can arise [57, 44] when $\pi\rho$ is near the upper bound.) Because the axion kinetic energies (4.7) or (4.10) are proportional to $1/\rho$, this enables us to suppress the F_{b_i} by about an order of magnitude relative to the familiar value $\alpha_C M_P/2\pi\sqrt{2} \sim 1.1 \times 10^{16} \, \text{GeV}$.

The upper bound on ρ comes from the way the metric on Z varies as a function of the eleventh dimension. For example, if the volume of Z is going to zero at one endpoint of I (which is not necessarily so as the extended Kahler cone has other boundaries), then the upper bound on the possible value of $\pi\rho$ was very roughly, within perhaps a factor of 2, estimated in [57] to be

$$\pi\rho_{\text{max}} = \frac{V_Z}{\ell_{11}^3 \left| \int_Z \omega \wedge \frac{\text{tr} \, F \wedge F - \frac{1}{2} \text{tr} \, R \wedge R}{16\pi^2} \right|}, \tag{4.12}$$

where ω is the Kahler class of Z. Under an overall scaling of Z, the integral in the denominator scales as $V_Z^{1/3}$, so $\pi\rho_{\text{max}}/\ell_{11}$ scales as $V_Z^{2/3}/\ell_{11}^4 \sim \alpha_C^{-2/3}$. In the example that Z is a K3 fibration over a large two-cycle C, if we assume that the dominant contribution to the integral in the denominator is $\int_Y (\text{tr}_1 F \wedge F - \frac{1}{2} \text{tr} \, R \wedge R)/16\pi^2 \cdot \int_C \omega$, then we can estimate the integral as $V_C|N_1 - 12| = V_C|k'|$, so

$$\pi\rho_{\text{max}} V_C \lesssim V_Z/\ell_{11}^3 |k'|. \tag{4.13}$$

When this is inserted in (4.10), we get $F_b/|k'| \gtrsim k\alpha_C M_P/2\pi\sqrt{2} \sim k \cdot 1.1 \times 10^{16} \, \text{GeV}$, a familiar value from the heterotic string. What has happened is simply that since ρ and V_C vary reciprocally, we cannot make them both large.

Explicit violation of the PQ symmetries. The PQ symmetry of these modes is violated explicitly by membrane instantons wrapped on $D \times I$, where D is a two-cycle in Z. The volume of such a membrane is $V_D = \pi\rho\widetilde{\epsilon}$ (where $\widetilde{\epsilon}$ is a factor similar to ϵ and accounts for how the area of D varies in the eleventh dimension). As the M2-brane tension is $2\pi/\ell_{11}^3$, the action is $S = 2\pi^2\rho V_D/\ell_{11}^3$. For a fairly isotropic Z, $V_D/\ell_{11}^2 \sim (V_Z/\ell_{11}^6)^{1/3} \sim \alpha_C^{-1/3}$, and $\pi\rho_{\text{max}}/\ell_{11} \sim \alpha_C^{-2/3}$, so $S \sim 2\pi/\alpha_C$. Hence, explicit violation of the PQ symmetries might be small enough to solve the strong CP problem.

For the case of a fibration $Z \to C$ with a large two-cycle C, our inequality (4.13) together with (4.1) leads to $S = 2\pi\widetilde{\epsilon}/\alpha_C|k'|$. So we might need $|k'| = 1$ to have a sufficiently good PQ symmetry for this mode.

4.3 Warped compactification

In heterotic M-theory, the metric on $M \times Z \times I$ is actually a warped product, not an ordinary product, but we have not taken this into account so far. Here, we study the question of whether warping could play an important effect in heterotic M-theory. As we will see, the warping does not change the qualitative conclusions in the usual approach to phenomenology in which the Standard Model is enbedded at the end of I at which the volume is larger. However, it can be quite important in the opposite case.

In supersymmetric compactification of heterotic M-theory, the warped metric takes the simple form [57, 59, 60]

$$ds^2 = e^{-f(x_{11})}\eta_{\mu\nu}dx^\mu dx^\nu + e^{f(x_{11})}\left(g_{mn}dy^m dy^n + dx^{11}dx^{11}\right). \tag{4.14}$$

Here y^m are local coordinates on Z and x^{11} is parameterizes the heterotic M-theory interval. The warp factor is

$$e^{f(x_{11})} = (1 + x^{11}Q)^{2/3}. \tag{4.15}$$

(Without assuming low energy supersymmetry, more general fluxes are possible, but the supersymmetric case seems general enough to illustrate our point.) The parameter Q, which we will loosely call "instanton number," is given by

$$Q = \frac{\ell_{11}^3}{2V_Z}\int_Z \omega \wedge \frac{1}{16\pi^2}\left(\mathrm{tr}_1 F \wedge F - \frac{1}{2}\mathrm{tr}\, R \wedge R\right), \tag{4.16}$$

where ω is the Kahler form of Z. Note that, if the instanton number at $x^{11} = 0$ is larger than at $x^{11} = \pi\rho$, then the integral (4.16) is negative. This follows from the supersymmetry relation $\omega_{ij}F^{ij} = 0$. The integral is roughly $V_Z^{1/3} = \ell_{11}^2(k\alpha_C)^{-1/3}$, so we express it in terms of a dimensionless number q of order one

$$Q = \frac{(k\alpha_C)^{2/3}}{2\ell_{11}}q. \tag{4.17}$$

If the instanton number at the $x^{11} = 0$ boundary is greater than the instanton number at the other boundary, Q is negative and Z shrinks along x^{11}. In this case, there is a critical coordinate distance

$$\pi\rho_{\max} = \frac{1}{|Q|} \tag{4.18}$$

at which the warp factor (4.15) becomes zero and our supergravity approximation breaks down. Hence, in the following we assume that $\rho < \rho_{\max}$. In the opposite situation with Q positive, Z expands along the interval so the length of the interval could be in principle arbitrarily large.

Dimensional reduction of the gravity action (2.27) to four dimensions gives

$$S = \frac{2\pi V_Z}{\ell_{11}^9}\int_0^{\pi\rho} dx^{11} e^{5/2f(x^{11})}\int d^4x\sqrt{-g}R, \tag{4.19}$$

whence the four-dimensional Planck mass is

$$M_P^2 = \frac{3\pi V_Z}{2\ell_{11}^9 Q}\left[(1 + \pi\rho Q)^{8/3} - 1\right]. \tag{4.20}$$

– 22 –

The gauge fields of the two E_8's live on the two ends of the interval. Reducing the Yang-Mills action (2.28)

$$S_{YM} = -\frac{1}{4(2\pi)\ell_{11}^6} \int_{M^{10}} \text{tr}\, F \wedge \star F \tag{4.21}$$

to four dimensions gives the gauge coupling as a function of the boundary location

$$\frac{1}{\alpha_{YM}(x^{11})} = \frac{V_0}{\ell_{11}^6}(1 + x^{11}Q)^2. \tag{4.22}$$

In the linear approximation this agrees with the result of [57]

$$\frac{d}{dx^{11}}\left(\frac{1}{\alpha_{YM}}\right) = \ell_{11}^{-3} \int_Z \omega \wedge \left(\text{tr}_1 F \wedge F - \frac{1}{2}\text{tr}\, R \wedge R\right). \tag{4.23}$$

Standard approach to phenomenology. Now let us re-examine the case that the Standard Model is embedded in the larger of the two boundaries. This will give results qualitatively similar to those we found above, but somewhat more precise. With the Standard Model embedded at the larger end, the volume is decreasing away from the Standard Model boundary. At a coordinate distance $\pi\rho_{\text{crit}} = 1/|Q|$ the volume becomes zero in our approximation. This gives an upper bound on M_P [57]. Substituting $1/|Q|$ for $\pi\rho$ into the expression for the Planck mass we find

$$M_{P,\text{max}}^2 = \frac{1}{\ell_{11}^2}\frac{3\pi}{q(k\alpha_C)^{5/3}}, \tag{4.24}$$

which is a factor of $3/2$ larger [59] than the limit on M_P^2 obtained in the linear approximation [57].

The warping of the metric (4.14) affects the internal wavefunction of the axion as well. The warping is different for the model-independent and for the model-dependent axions. For model-independent axion we modify the ansatz (4.2) by a nontrivial warp factor

$$C = \alpha e^{g(x^{11})} B(x^\mu) \wedge dx^{11}, \tag{4.25}$$

where α is a normalization constant and $g(x^{11})$ captures the dependence of C on x^{11}. To find g we require that the three-form equation of motion $d \star dC = 0$ in eleven dimensions reduces to the four-dimensional equation of motion $d_4 \star_4 d_4 B = 0$ for the two-form field B. We substitute (4.25) into the equation of motion $d \star dC = 0$ which, after taking the first exterior differential and the Hodge dual becomes

$$d\left(e^{g+\frac{7}{2}f}d\,\text{vol}_Z \wedge (\star_4 d_4 B)\right) = 0, \tag{4.26}$$

where $d\,\text{vol}_Z$ is the volume form on Z. Expanding the second exterior differential into its four-dimensional and seven-dimensional parts $d = d_4 + d_7$, (4.26) becomes

$$d_4 \star_4 d_4 B + d_7 \left(g + \frac{7}{2}f\right) \star_4 d_4 B = 0. \tag{4.27}$$

– 23 –

In writing this we suppressed an overall factor of $e^{g+7f/2}d\,\mathrm{vol}_Z$. The equation (4.27) reduces to the four-dimensional equation of motion for the massless two-form B if $g = -7f/2 + c$. Here c is a constant that we absorb into the normalization constant α of the C-field. Hence the ansatz for the model-independent axion is

$$C = \alpha(1 + x^{11}Q)^{-7/3}B \wedge dx^{11}. \tag{4.28}$$

We fix the normalization constant α by requiring $G = dC$ to have integer periods. We assume that B is normalized so that $H = dB$ has integer periods, so that the axion has period 2π. With this normalization, G will have integer periods if $\alpha(1 + x^{11}Q)^{-7/3}dx^{11}$ integrates to one on the M-theory interval I. This condition fixes α to be

$$\alpha^{-1} = \frac{3}{4Q}\left[1 - (1 + \pi\rho Q)^{-4/3}\right]. \tag{4.29}$$

Dimensional reduction of the G-field kinetic energy (2.27) gives the kinetic term for the four-dimensional B-field

$$\frac{2\pi\alpha V_Z}{\ell_{11}^3}\int\left(-\frac{1}{2}H \wedge \star H\right). \tag{4.30}$$

In four dimensions, B is dual to an axion with axion decay constant equal to the inverse of the coefficient of the H-field kinetic energy

$$F_b^2 = \frac{1}{\ell_{11}^2}\frac{3(k\alpha_C)^{1/3}}{4\pi q}\left[1 - (1 + \pi\rho Q)^{-4/3}\right]. \tag{4.31}$$

From (4.31) we read off the dependence of F_b on the length of the M-theory interval. We assume that Q is negative, so that Z gets smaller away from the Standard Model boundary. When the interval is short, this agrees with the formula (4.4) that does not take into account warping. As the interval gets longer, both F_b and M_P grow compared to ℓ_{11}^{-1}. At $\pi\rho_{\max} = 1/|Q|$, M_P reaches its maximum value (4.24) while F_b becomes formally arbitrarily large. Intuitively, this is because the wavefunction of the model-independent axion is concentrated towards the $x^{11} = \pi\rho$ end of the interval, as is clear from (4.28), so the axion couples weakly to the Standard Model fields which are supported on the other end of I. However, before F_b becomes super-Planckian, the size of Z at the $x^{11} = \pi\rho$ end of the interval becomes sub-Planckian and our low energy supergravity description breaks down. This regime should be studied in a well behaved dual description, such as F-theory [61–63]. We estimate the maximum value of $\pi\rho$ for which we can trust our formula for F_b by requiring that the volume of Z is at least one in eleven-dimensional Planck units. From the warped metric (4.14) we read off the dependence of the volume of Z on x^{11} to be $V = \ell_{11}^6(1 + x^{11}Q)^2/(k\alpha_C)$. Requiring this to be at least ℓ_{11}^6 bounds the length of I: $(1 + \pi\rho_{\max}Q) \lesssim (k\alpha_C)^{1/2}$. Substituting this into (4.31) we estimate an upper bound on the decay constant of the model-independent axion

$$F_b \lesssim \frac{1}{\ell_{11}}\sqrt{\frac{3}{4\pi q(k\alpha_C)^{1/3}}} = \frac{k\alpha_C}{2\pi\sqrt{q}}M_P. \tag{4.32}$$

Up to a factor of order one, this agrees with the formula (4.4) that does not take into account warping. Hence, the warping does not modify the scale of the axion coupling parameter significantly, when the Standard Model is embedded into the larger of the two boundaries of I.

Let us now find the warp correction to the axion decay constant of the model-dependent axions. We start by modifying the ansatz (4.5) for the three-form field with a warp factor

$$C = \alpha \sum_i e^{g(x^{11})} \frac{b_i}{2\pi} \beta_i \wedge dx^{11}. \tag{4.33}$$

Here α is an normalization constant that we will determine later. To find the warp factor $g(x^{11})$, we substitute (4.33) into the equation of motion $d \star dC = 0$ which, after taking the first exterior derivative and the Hodge star operator becomes

$$\sum_i d\left(\star_4 d_4 b_i \wedge \star_6 \beta_i\, e^{g-\frac{1}{2}f}\right) = 0. \tag{4.34}$$

The symbols $\star_4, \star_6$ denote the four and six-dimensional Hodge operators. We simplified the equation using the fact that β_i are closed. (4.34) reduces to a set of four-dimensional equations of motion for massless scalars b_i if $g = f/2$ plus a constant that we can absorb into the normalization constant α. So the model-dependent axions come from the ansatz

$$C = \alpha \sum_i (1 + x^{11}Q)^{1/3} \frac{b_i}{2\pi} \beta_i \wedge dx^{11}. \tag{4.35}$$

The normalization constant α is fixed by requiring that each three-form $\alpha(1+x^{11}Q)^{1/3}\beta_i \wedge dx^{11}$ has unit flux through $C_i \times I$. Since $\int_{C_i} \beta_i = 1$, this gives

$$\alpha^{-1} = \frac{3}{4Q}[(1 + \pi\rho Q)^{4/3} - 1]. \tag{4.36}$$

Dimensional reduction of the C-field kinetic energy (2.27) leads to the kinetic terms of the axions b_i

$$\frac{\alpha}{2\pi\ell_{11}^3} \int_M -\frac{1}{2}\gamma_{ij}^M \partial b_i \wedge \star\partial b_j \tag{4.37}$$

with γ_{ij}^M given by (4.8).

In the standard approach to phenomenology, the volume of Z decreases away from the end of the interval with Standard Model fields. We saw already in the linear approximation (4.10) that F_b decreases as we increase the length of the interval I. Hence, to find the smallest possible F_b we take ρ large, $\pi\rho_{\max} = |Q|^{-1}$. For an axion coming from a generic cycle C, we estimate the integral $\int_Z \beta \wedge \star\beta \sim V_Z^{1/3} = \ell_{11}^2(k\alpha_C)^{-1/3}$, so the axion decay constant in terms of the Planck mass (4.24) is

$$F_b \gtrsim \frac{qk\alpha_C}{3\pi} M_P, \tag{4.38}$$

which is close to the familiar value $1.1 \times 10^{16}\text{GeV}$. As we discussed before, we can try to make F_b small by taking size of the cycle C large. So we take Z to be a fibration over a large Riemann surface C with a fiber Y. A calculation identical to the one performed in section 4.2 shows that taking C large does not help in lowering F_b because V_C and $\rho_{\max}$ vary reciprocally, so we cannot make them both large.

Non-standard approach to phenomenology. The warping does not modify the scale of the axion coupling parameter significantly, when the Standard Model is embedded into the larger of the two boundaries of I. On the other hand we expect the effects of the warping to be pronounced if the size of Z is increasing away from the Standard Model boundary. In this case, the length of the interval can be arbitrarily large. As we increase it, the volume of the seven-dimensional compactification manifold gets larger so the four-dimensional Planck scale grows compared to the M-theory scale ℓ_{11}^{-1}. This is clear from the formula (4.20) for M_P with Q taken to be positive.

The wavefunction of the model-independent axion (4.28) is localized near the SM end of the interval, so the axion decay constant stays close to the M-theory scale ℓ_{11}^{-1}. Indeed, in the limit of large interval $\pi\rho \gg Q^{-1}$, the axion decay constant (4.31) approaches a constant multiple of ℓ_{11}^{-1}

$$F_b^2 = \frac{1}{\ell_{11}^2} \frac{3(k\alpha_C)^{1/3}}{4\pi q}, \qquad (4.39)$$

while the Planck mass (4.20) grows with $\pi\rho$ as

$$M_P^2 = \frac{1}{\ell_{11}^2} \frac{3\pi(\pi\rho Q)^{8/3}}{q(k\alpha_C)^{5/3}}. \qquad (4.40)$$

Hence, increasing ρ lowers the axion decay constant compared to M_P

$$F_b = \frac{k\alpha_C}{2\pi} \frac{M_P}{(\pi\rho Q)^{4/3}}. \qquad (4.41)$$

Solving (4.39) and (4.40) for M_{11} and $\pi\rho$, we get

$$M_{11} = F_b \sqrt{\frac{4\pi q}{3(k\alpha_C)^{1/3}}},$$
$$\pi\rho = \ell_{11} \left(\frac{M_P}{F_b}\right)^{3/4} \frac{2^{1/4}(k\alpha_C)^{1/12}}{\pi^{3/4} q}. \qquad (4.42)$$

For phenomenologically preferred axion decay constant $10^9 \text{GeV} < F_b < 10^{12} \text{GeV}$, the allowed range of the M-theory parameters is

$$3.5 \times 10^9 \text{GeV} \lesssim M_{11} \lesssim 3.5 \times 10^{12} \text{GeV},$$
$$4.2 \times 10^6 \ell_{11} \gtrsim \rho \gtrsim 2.3 \times 10^4 \ell_{11}. \qquad (4.43)$$

Let us now consider a generic model-dependent axion with $\int_Z \beta \wedge \star\beta \sim V_Z^{1/3}$. We get the axion decay constant by substituting for the normalization constant $\alpha = \frac{4Q}{3}(\pi\rho Q)^{-4/3}$ from (4.36) into (4.37)

$$F_b = M_P \frac{4}{3\pi q(k\alpha_C)^{1/3}} \left(\frac{\ell_{11}}{\pi\rho}\right)^2. \qquad (4.44)$$

To simplify (4.44) we used that $Q = q(k\alpha_C)^{2/3}/2\ell_{11}$, $V_Z = \ell_{11}^6/(k\alpha_C)$ and the expression (4.40) for M_P. It is clear from the formula for F_b that as the length of I gets larger, the axion decay scale is parametrically lowered compared to M_P. We can estimate how

much is it necessary to elongate the interval to bring F_b down to the cosmologically favored region. From (4.44) and (4.40) we express $\pi\rho$ and $M_{11} = \ell_{11}^{-1}$ as

$$M_{11} = F_b \left(\frac{M_P}{F_b}\right)^{1/3} \left(\frac{3\pi k \alpha_C}{q}\right)^{1/6}$$

$$\pi\rho = \ell_{11} \left(\frac{M_P}{F_b}\right)^{1/2} \frac{2}{\sqrt{3\pi q (k\alpha_C)^{1/3}}}. \tag{4.45}$$

If we take $k, q = 1$, then the axion decay constant takes the phenomenologically preferred values $10^9 \text{GeV} < F_b < 10^{12} \text{GeV}$ for the M-theory scale and the length of the interval lying in the range

$$1.1 \times 10^{12} \text{GeV} \lesssim M_{11} \lesssim 1.1 \times 10^{14} \text{GeV}$$

$$5.4 \times 10^4 \ell_{11} \gtrsim \pi\rho \gtrsim 1.7 \times 10^3 \ell_{11}. \tag{4.46}$$

Generalization to other warped compactifications. In the context of heterotic M-theory, we have seen how significant warping can lower the axion decay scale. A similar effect can occur in other warped compactifications, for example in compactifications of the type II superstring. If the axion is supported near a region with significant warping, its decay constant will be lowered compared to the Planck scale just like in the case of heterotic M-theory with non-standard approach to phenomenology.

5. Anomalous $U(1)$ symmetries in string theory

To discuss the role of anomalous $U(1)$ symmetries for axion physics, we begin by considering compactifications of the heterotic string on a smooth six-manifold. In many such compactifications, the low energy gauge group, understood as the subgroup of $E_8 \times E_8$ or $SO(32)$ that commutes with the gauge field expectation value on the compact manifold, contains an anomalous abelian gauge symmetry $U(1)_B$ (or several such symmetries). For example, this is very common in supersymmetric $(0,2)$ models in which the gauge fields on the compact manifold have $U(1)$ factors. The anomaly appears in the charges of the massless fermions. It is canceled by a Green-Schwarz mechanism involving one of the axion multiplets. For brevity, we will consider the case that the axion in question is the model-independent one. The relevant fields participating in the four-dimensional Green-Schwarz mechanism are in that case the vector field V_B of the anomalous $U(1)_B$ and the axion-dilaton field $S = 1/g_B^2 + ia/8\pi^2$, where g_B is the four-dimensional gauge coupling of $U(1)_B$. The Kahler potential for these fields is [66, 67, 65]

$$K = -M_P^2 \ln(S + \overline{S} - cV_X), \tag{5.1}$$

where $c = \text{tr}\, B/6$ is a multiple of the chiral trace of the generator of $U(1)_B$ over the massless fermions. The chiral trace counts right-handed fields with an extra minus sign compared to left-handed ones. Under a gauge variation $V_B \to V_B + i(\Lambda_B - \overline{\Lambda}_B)$, the axion-dilaton superfield transforms as $S \to S + ic\Lambda_B$. Hence the axion has shift symmetry

$$a \to a + c\,\theta_B, \tag{5.2}$$

where $\theta_B = \mathrm{Re}\,\Lambda_B|_{\theta=\bar{\theta}=0}$. This makes possible the anomaly cancellation. The axion has anomalous couplings $a\,\mathrm{tr}\,F_i \wedge F_i$ to nonabelian gauge fields; the variation of these couplings under (5.2) exactly cancels the $U(1)_B$ anomaly.

We can now study the four-dimensional effective theory of the axion and the $U(1)_B$ gauge fields. The effective action contains the terms

$$-\frac{1}{2}F_a^2\left(\partial_\mu a + B_\mu \frac{\mathrm{tr}\,B}{12}\right)^2 + \zeta^2 D_B + a\frac{1}{16\pi^2}\,\mathrm{tr}\,(F \wedge F)_{\mathrm{QCD}}, \qquad (5.3)$$

where the first two terms come form expanding the Kalher term (5.1) and $\mathrm{tr}\,(F \wedge F)_{\mathrm{QCD}}$ is a multiple of the QCD instanton density. $F_a = k\alpha_C M_P/(2\pi\sqrt{2})$ is the familiar axion decay constant of the model-independent axion and ζ is the Fayet-Iliopoulos term

$$\zeta^2 = k\alpha_C M_P^2 \frac{\mathrm{tr}\,B}{48\pi}. \qquad (5.4)$$

From the effective action (5.3), we see that the gauge boson eats the axion and acquires a mass via the Higgs mechanism:

$$M_B = \frac{k\alpha_C}{\ell_s}\frac{\mathrm{tr}\,B}{6\sqrt{2}}. \qquad (5.5)$$

Below the scale M_B, the anomalous gauge boson decouples, leaving behind, in sigma model and spacetime perturbation theory, an anomalous global $U(1)_B$ symmetry [31]. This symmetry is broken explicitly by instantons.

It has been recognized in [70] that the surviving global symmetry can be used to solve the strong CP-problem with F_a well below the string scale. In reducing to four dimensions, the four-dimensional massless spectrum frequently contains scalars, i.e. from the gauge bundle moduli, that are charged under the anomalous symmetry and are neutral under other gauge groups. For illustration, we assume that there is one such scalar field ϕ with charge q under $U(1)_B$. If ϕ acquires a VEV $\langle\phi\rangle$, this spontaneously breaks the global $U(1)_B$. We will assume that this VEV results from some dynamics at an energy scale below the string scale, so that $|\langle\phi\rangle| < M_B$. The $U(1)_B$ gets realized nonlinearly by shifting the phase of $\phi = |\phi|e^{ib}$

$$b \to b + q\theta_B. \qquad (5.6)$$

The kinetic energy of b follows from the ϕ kinetic term $-|D_\mu \phi|^2$

$$-|\phi|^2(\partial_\mu b - qA_\mu). \qquad (5.7)$$

Hence, b is a PQ axion with decay constant $F_b = \sqrt{2}|\phi|$. Its coupling to the QCD instanton density is determined by the underlying $U(1)_B$ anomaly to be

$$b\frac{c}{16\pi^2 q}\,\mathrm{tr}\,(F \wedge F)_{\mathrm{QCD}} \qquad (5.8)$$

with axionic coupling $k = c/q$. The couplings of this axion to matter are determined by the ratio F_b/k. This is

$$\frac{F_b}{k} = \frac{12\sqrt{2}q}{\mathrm{Tr}\,B}|\phi|, \qquad (5.9)$$

which is roughly $|\phi.|$ Hence, if ϕ can be stabilized with expectation value much less than the string scale, the axion might have decay constant in the favored range $10^9 \text{GeV} < F_b < 10^{12} \text{GeV}$.

In supersymmetric compactifications of the heterotic string, it seems difficult to stabilize $|\phi|$ at small expectation values because of the D-term constraint

$$D_B = q|\phi|^2 - \zeta^2 = 0 \tag{5.10}$$

that has to be satisfied to preserve supersymmetry. This forces ϕ to acquire a nonzero VEV $|\phi|^2 = \zeta^2/q$. Hence the b has axion decay constant

$$F_b = \sqrt{2}|\phi| = M_P \sqrt{\frac{k\alpha_C \, \text{Tr} \, B}{24\pi q}}. \tag{5.11}$$

From (5.4), (5.11) we see that $F_b \sim F_a$ so the optimistic hypothesis that $F_b \ll F_a$ does not apply. Since the axion decays scales of a, b are comparable, to get a quantitative description, we need to keep in the effective action both axions a, b and the anomalous gauge field B_μ:

$$-\frac{1}{2}F_a^2(\partial_\mu a - q_a B_\mu)^2 - \frac{1}{2}F_b^2(\partial_\mu b - q_b B_\mu)^2 - \frac{1}{4g_B^2}\text{tr}\, F_{B,\mu\nu}F_B^{\mu\nu} + a\frac{1}{16\pi^2}\text{tr}(F_\wedge F)_{\text{QCD}} \tag{5.12}$$

Here $q_a = \text{Tr} \, B/12, q_b = q$. Diagonalizing (5.12), it can be shown that one linear combination of a, b gets eaten by the $\text{U}(1)_B$ gauge field, giving it a string scale mass. The other linear combination survives to low energies as a Peccei-Quinn axion. Since $F_a, F_b \sim M_s$, the PQ axion has roughly string scale decay constant [66]. Having several fields ϕ_α charged under $\text{U}(1)_B$ does not appear likely to change this conclusion.

Anomalous $\text{U}(1)$'s in type II string theory. It may be possible to make models with low axion decay constant in Type II string with intersecting D-branes. Here, the axion that cancels the $\text{U}(1)_B$ anomaly is a twisted RR axion [73, 72]. At tree level, the FI-term is determined in Type IIA string by complex structure moduli and in Type IIB string by Kahler moduli [76]. It has been argued that the one loop contribution to FI-terms (that in heterotic string theory generates string scale FI-term) is absent in Type II-orientifolds by reinterpreting the FI terms as closed string tadpoles [71, 74]. Hence, there might be a value of moduli for which the FI-term of the anomalous $\text{U}(1)$ is well below the string scale.

6. M-theory on a manifold of G_2 holonomy

In the present section, we consider M-theory on a seven-manifold X of G_2 holonomy, which we call X. If one is not concerned about maintaining $\mathcal{N} = 1$ supersymmetry in four dimensions, much of the discussion applies to M-theory compactification on more general seven-manifolds. We suppose that X contains a three-manifold Q of orbifold singularities, leading to four-dimensional gauge fields. For example, Q may be a locus of $\mathbf{A}_2$ singularities, leading to color SU(3) gauge fields, or $\mathbf{A}_4$ singularities, leading to a theory somewhat similar

to SU(5) GUT's in four dimensions. For the qualitative investigation of the axion coupling parameters, such details are inessential.

Section 4 was also devoted to M-theory on a seven-manifold, of the form $Z \times I$. The main difference is the strategy for getting Standard Model gauge fields; in section 4, these were supposed to arise at the boundary of the world, while in our present discussion, we will get gauge fields from orbifold singularities.

From (2.27), it follows that in reduction on $M \times X$, the four-dimensional Planck scale is given by

$$M_P^2 = \frac{4\pi V_X}{\ell_{11}^9}. \tag{6.1}$$

The action for SU(3) (or SU(5)) gauge fields along $M \times Q$ is[5]

$$\frac{1}{8\pi\ell_{11}^3} \int_{M \times Q} d^7x \sqrt{-g}\, \text{tr}\, F_{\mu\nu}F^{\mu\nu}. \tag{6.2}$$

Upon reduction to four dimensions and recalling the convention $\text{tr}\, t^a t^b = \frac{1}{2}\delta^{ab}$ for Lie algebra generators, this becomes

$$\frac{V_Q}{16\pi\ell_{11}^3} \int_M d^4x \sqrt{-g} F_{\mu\nu}^a F^{\mu\nu\,a}. \tag{6.3}$$

So the color SU(3) gauge coupling at the compactification scale obeys $g_C^2 = 4\pi\ell_{11}^3/V_Q$ or equivalently

$$\alpha_C = \frac{\ell_{11}^3}{V_Q}. \tag{6.4}$$

Axions arise from zero modes of the three-form field C of M-theory. If γ_i, $i = 1, \ldots, b_3(X)$ are the harmonic three-forms on X, normalized so that

$$\int_{D_i} \gamma_j = \delta_{ij}, \tag{6.5}$$

for a suitable basis of three-cycles D_i, then we make an ansatz

$$C = \frac{1}{2\pi} \sum_i c_i \gamma_i, \tag{6.6}$$

where c_i are massless fields on M. The kinetic energy for the c_i is obtained by dimensional reduction from (2.27). It is

$$\frac{1}{2\pi\ell_{11}^3} \int_M \frac{1}{2} \widetilde{\gamma}_{ij}^M dc_i \wedge \star dc_j, \tag{6.7}$$

where

$$\widetilde{\gamma}_{ij}^M = \int_X \gamma_i \wedge \star \gamma_j. \tag{6.8}$$

---

[5]To get this formula, start with eq. (13.3.25) of [54], which shows that the kinetic energy for gauge fields on a Type IIA $D6$-brane is $(1/8\pi g_s\ell_s^3) \int d^7x\sqrt{-g}\, \text{tr}\, F_{\mu\nu}F^{\mu\nu}$. Then convert to M-theory parameters via the relation $g_s\ell_s^3 = \ell_{11}^3$, which follows from the equality of $D2$-brane and $M2$-brane tensions using our conventions in section 2.

The c_i have approximate shift symmetries because of the underlying gauge-invariance of the C-field. They have axion-like couplings because the M-theory effective action at an orbifold singularity includes a coupling

$$2\pi \int_{M \times Q} C \wedge \frac{1}{8\pi^2} \operatorname{tr} F \wedge F. \tag{6.9}$$

The existence and coefficient of this coupling follow from the analogous coupling on the world-volume of a $D6$-brane in Type IIA superstring theory; note that the coupling in (6.9) is quantized for topological reasons, so it is completely determined in M-theory by what it is in Type IIA. (6.9) reduces in four dimensions to

$$\sum_i k_i \int_M c_i \frac{1}{8\pi^2} \operatorname{tr} F \wedge F, \tag{6.10}$$

with

$$k_i = \int_Q \gamma_i. \tag{6.11}$$

Thus, the modes c_i indeed have the couplings of axions.

Let us define a "radius" R of X by $V_X = R^7$. Under an overall scaling of the metric of X, the integral in (6.7) scales like R. Thus if X is reasonably isotropic, a "generic" linear combination of the c_i has

$$F_c^2 = \frac{xR}{2\pi \ell_{11}^3}, \tag{6.12}$$

where x is formally of order 1. Let us first assume that Q is a generic three-cycle, say with R_Q of order R. If we set $R_Q = R$, then (6.12) and (6.4) give us $F_c = x\alpha_C M_P/2\pi\sqrt{2}$, a familiar sort of formula. With R_Q of order R, compactification on a manifold of G_2 holonomy can give reasonable GUT-like phenomenology and a sensible relations between Newton's constant and the GUT scale; these issues have been explored in [77]. In this case, evidently, we expect the axion coupling parameter to be close to the GUT scale. The action for a membrane wrapped on Q is precisely $I = 2\pi R_Q^3/\ell_{11}^3 = 2\pi/\alpha_C$, and indeed such a membrane is equivalent to an instanton of the gauge theory. For $R_Q \sim R$, generic membrane instantons have actions comparable to this. Hence, with $\alpha_C \sim 1/25$, explicit PQ violation at high energies might be small enough, subject to the usual caveats, to lead to a solution of the strong CP problem.

Alternatively, we can ask what are the necessary parameters that give F_c in the range allowed by the usual cosmological arguments. We can solve (6.1) and (6.12) for ℓ_{11} and R in terms of F_c and M_P:

$$\ell_{11} = \frac{1}{F_c} \left(\frac{M_P}{F_c} \right)^{1/6} \frac{x^{7/12}}{2^{3/4}\pi^{2/3}}$$
$$R = \ell_{11} \left(\frac{M_P}{F_c} \right)^{1/3} \frac{1}{2^{1/2}\pi^{1/3}}. \tag{6.13}$$

For $10^9 \,\mathrm{GeV} \lesssim F_c \lesssim 10^{12} \,\mathrm{GeV}$, and taking $x = 1$, we have roughly

$$\frac{10}{F_c} \gtrsim \ell_{11} \gtrsim \frac{3}{F_c}$$
$$500\ell_{11} \gtrsim R \gtrsim 50\ell_{11}. \tag{6.14}$$

Even at the lower end of this range, a generic membrane instanton action $2\pi R^3/\ell_{11}^3$ is prohibitively large, so significant PQ violation will come only from gauge instantons — or from membranes wrapped on the vanishing cycles that we discuss later.

Similarly, we define a "radius" R_Q of Q such that $V_Q = R_Q^3$. From (6.4), $\alpha_C = \ell_{11}^3/R_Q^3$. To get a reasonable value of α_C, R_Q must be fairly close to ℓ_{11}. It follows, therefore, from (6.14) that if F_c is to be in the usual cosmological range, R_Q must be substantially less than R. So the three-cycle on which gauge fields are supported must be substantially smaller than a generic three-cycle in X.

It is believed to be possible for a manifold X of G_2 holonomy to develop, as its moduli are varied, a supersymmetric "vanishing cycles" — a three cycle that collapse to zero even as the metric on the rest of X has a limit. The example about which most is known is the case that the vanishing cycle is a three-sphere S^3 – or an orbifold quotient thereof, S^3/Γ, with Γ a finite group of symmetries. The local structure near the vanishing cycle is given by an explicitly known [78, 79] asymptotically conical metric of G_2 holonomy on the manifold $S^3 \times \mathbb{R}^4$ (or an orbifold quotient thereof). Its role in M-theory has been analyzed in some detail [80 – 82]

In our problem, to get F_c in the usual cosmological range, it is natural to assume that Q is such a vanishing cycle — and the vacuum is near a point in moduli space at which V_Q would go to zero. (We do not know why the vacuum would be near such a point, but we recall that for Type IIB superstring theory, mechanisms have been proposed [83] that can lead to a vacuum near a point with a vanishing cycle.) In fact, the example that the local structure is an orbifold quotient of $S^3 \times \mathbb{R}^4$ has some of the right properties for us. By dividing by a group $\Gamma' = \mathbb{Z}_3$ or $\Gamma' = \mathbb{Z}_5$ that acts only on $\mathbb{R}^4$ with an isolated fixed point at the origin, one can get SU(3) or SU(5) gauge fields supported on S^3. In the SU(5) case, if one also divides by a group Γ that acts freely on S^3, one introduces the possibility of spontaneously breaking SU(5) to the Standard Model. To include quarks and leptons, one would need to complicate the singularity structure [84]. Such models have some obvious potential problems; because the compactification scale is well below the usual GUT scale (to get F_c in the usual cosmological range) it will be hard to avoid rapid proton decay or to get the right low energy gauge couplings. A more complicated structure of vanishing cycles might be necessary.

Axion physics with a small S^3. In the case when X develops a vanishing S^3, we can be more precise in estimating F_c, as the explicit metric of the local G_2 holonomy manifold with vanishing S^3 is known [79]. Locally the manifold has the topology $S^3 \times \mathbb{R}^4/\Gamma$, where Γ is a finite subgroup of SU(2) with which we orbifold $\mathbb{R}^4$ to get nonabelian gauge symmetry. Asymptotically, the manifold is a cone over $S^3 \times S^3/\Gamma$ where the S^3 is homologous to the vanishing three-cycle Q and S^3/Γ is the quotient of the unit three-sphere in $\mathbb{R}^4$.

The axion decay constant is (6.7)

$$F_c^2 = \frac{1}{2\pi\ell_{11}^3} \int_X \gamma_Q \wedge \star\gamma_Q, \qquad (6.15)$$

so a more precise evaluation of F_c amounts to a calculation of the norm of γ_Q.

For F_c in the range allowed by standard cosmological arguments, the estimate (6.14) gave $R \gg R_Q$. Hence, to a high degree of accuracy, we can neglect the finite size of Q and treat X locally as a cone over $S^3 \times S^3 / \Gamma$. We work in the "upstairs picture" on a cone over $S^3 \times S^3$. To take into account quotienting by Γ, one divides F_c^2 at the end of the calculation by $|\Gamma|$.

To write down the metric explicitly, we introduce invariant one-forms σ_i, Σ_i, $i = 1, \ldots, 3$ on the two three-spheres. We normalize them so that the usual round metric on a three-sphere with radius one is $ds^2 = \sum_i \sigma_i^2$. With this normalization, σ_i obey the usual SU(2) relations $d\sigma_i = -\frac{1}{2}\epsilon_{ijk}\sigma_j \wedge \sigma_k$ and the volume form on S^3 is just $\sigma_1 \wedge \sigma_2 \wedge \sigma_3$. The metric on the base of the cone, $S^3 \times S^3$, is the squashed Einstein metric

$$
\begin{aligned}
ds^2 &= dr^2 + \frac{r^2}{9}\left(\sigma_i - \frac{\Sigma_i}{2}\right)^2 + \frac{r^2}{12}\Sigma_i^2, \\
&\equiv dr^2 + \nu_i^2 + e_i^2.
\end{aligned}
\tag{6.16}
$$

In the second line we introduced the orthonormal basis $\nu_i = r(\sigma_i - \Sigma_i/2)/3$, $e_i = r\Sigma_i/\sqrt{12}$. In writing the metric we chose Σ_i to be the one-forms of the non-contractible sphere Q. The vanishing cycle Q is Poincare dual to a harmonic three-form [85], eq. (2.44), which in the conical limit becomes

$$
\omega_3 = \frac{1}{r^3}\epsilon^{ijk}\nu_i \wedge \nu_j \wedge e_k + \frac{3}{r^3}e_1 \wedge e_2 \wedge e_3.
\tag{6.17}
$$

The Hodge dual of ω_3 is

$$
\star\omega_3 = \frac{dr}{r^3}\left(\epsilon_{ijk}\nu_i \wedge e_j \wedge e_k - 3\nu_1 \wedge \nu_2 \wedge \nu_3\right).
\tag{6.18}
$$

The flux of ω_3 through the non-contractible S^3 is

$$
\int_{S^3} \omega_3 = \frac{\pi^2}{4\sqrt{3}},
\tag{6.19}
$$

so the three-form with unit flux through Q is $\gamma_Q = 4\sqrt{3}\omega_3/\pi^2$ and the axion comes from the ansatz $C = \gamma_Q a/2\pi$. We estimate the axion decay constant by substituting the explicit form of γ_Q into (6.15)

$$
\begin{aligned}
F_c^2 &= \frac{1}{2\pi\ell_{11}^3 N}\left(\frac{4\sqrt{3}}{\pi^2}\right)^2 \frac{12}{(6\sqrt{3})^3}\int_X dr \wedge d\sigma_1 \wedge d\sigma_2 \wedge d\sigma_3 \wedge d\Sigma_1 \wedge d\Sigma_2 \wedge d\Sigma_3, \\
&= \frac{R}{\ell_{11}}\frac{2^4 x}{3^{5/2}\pi N}M_{11}^2,
\end{aligned}
\tag{6.20}
$$

where x is of order one and $R = V_X^{1/7}$ is the linear size of X (at which we cut off the integral). We divided the integral by $|\Gamma| = N$ to take into account the orbifolding of X by Γ. For SU(5) gauge symmetry along $M \times Q$ we take $\Gamma = \mathbb{Z}_5$. With $|\Gamma| = 5$, (6.20) comes out to be roughly the same size as (6.12). There is an additional contribution to F_c from the region that compactifies the cone. This gives an additive contribution to F_c^2 of size given by (6.12), hence it does not change (6.20) significantly.

With the estimate (6.20) for F_c, we re-derive the values for M-theory scale M_{11} and the size R of X in terms of M_P and F_c

$$M_{11} = F_c \left(\frac{F_c}{M_P}\right)^{1/6} \frac{3^{35/24}\pi^{2/3}N^{7/12}}{2^{13/6}x^{7/12}},$$

$$R = \ell_{11}\left(\frac{M_P}{F_c}\right)^{1/3} \frac{2^{1/3}x^{1/6}}{3^{5/12}\pi^{1/3}N^{1/6}}. \tag{6.21}$$

For F_c in the range allowed by standard cosmological constraints, setting $x = 1, N = 5$, we find

$$\frac{F_c}{6} \lesssim M_{11} \lesssim \frac{F_c}{2},$$

$$550\ell_{11} \gtrsim R \gtrsim 55\ell_{11}. \tag{6.22}$$

We see that the more precise calculation of F_c does not change the result (6.14) significantly.

Anisotropic seven-manifolds. So far we have assumed that our manifold X of G_2 holonomy is reasonably isotropic, with gauge fields supported either on a generic cycle or on a vanishing cycle. It is also possible for X to be highly anisotropic.

A G_2 manifold can have two types of supersymmetric fibration: a fibration by three-tori over a four-dimensional base, or by K3 surfaces over a three-dimensional base. Apart from special constructions involving orbifolds (as opposed to more generic G_2 manifolds), these are the most obvious kinds of highly anisotropic G_2 manifolds. Much of our discussion, in any case, has nothing to do with supersymmetry and would carry over to other kinds of highly anistopropic manifold.

Some features of these two types of example can be treated together. We refer to the fiber as F and the base as B; we write d for the dimension of the fiber (so $d = 3$ or 4), and we express the volumes as $V_X = V_F V_B$, $V_F = R_F^d$, $V_B = R_B^{7-d}$. We only need to consider the case that $R_F \ll R_B$. If $R_F \sim R_B$, we are back in the case we have already considered of a more or less isotropic manifold. It is generically impossible in a supersymmetric fibration to have $R_F \gg R_B$, but in any case, when this occurs one should look for an alternative description with the roles of F and B exchanged.

The four-dimensional reduced Planck mass obviously becomes

$$M_P^2 = \frac{4\pi V_F V_B}{\ell_{11}^9}. \tag{6.23}$$

The generalization of the formula (6.4) for α_C depends on the geometry. Let us suppose that Q has a dimensions wrapped on F and $3 - a$ wrapped on B, and that F and B are each more or less anisotropic. In this case,

$$\alpha_C = \frac{\ell_{11}^3}{V_Q} = y\frac{\ell_{11}^3}{R_F^a R_B^{3-a}} \tag{6.24}$$

where y is a constant of order 1. Obviously, this formula could be substantially changed if Q wraps vanishing cycles in either F or B. This would lead to a hybrid of our present discussion of anisotropic X with the earlier discussion of vanishing cycles.

Similarly, the estimation of F_c depends on what kind of three-cycle we consider. If ω is a harmonic three-form on X with b indices tangent to F and $3 - b$ indices tangent, and we keep the topological class of ω fixed while letting R_F and R_B vary, then $\int_X \omega \wedge \star \omega \sim V_X/R_F^{2b} R_B^{6-2b}$. In general (as described most fully by the appropriate "spectral sequence"), a harmonic form is a sum of components with different values of b. For $R_F \ll R_B$, the component with the largest b will dominate. The analog of (6.12) for a generic axion c with a given largest value of b is

$$F_c^2 = \frac{x V_X}{2\pi \ell_{11}^3 R_F^{2b} R_B^{6-2b}}. \tag{6.25}$$

To make F_c small, we want b to be as small as possible. However, if $b < a$, then the mode c does not have an axionic coupling to gauge fields that are supported on $M \times Q$. The smallest relevant F_c therefore has $b = a$. For such a mode, we can combine (6.25), (6.24), and (6.23) to the familiar order of magnitude relation $F_c \sim \alpha_C M_P$.

Thus, taking X to be highly anisotropic is not an efficient way to reduce the order of magnitude of F_c. To get a small F_c from a fibration with $R_F \ll R_B$, we must assume that Q is wrapped on a vanishing cycle in either F or B.

The arguments that we have just given also apply to other examples (such as gauge fields supported on Type II D-branes) that we will consider later.

7. Intersecting D-brane models

In this section, we consider intersecting D-brane models both in type IIA and IIB string theory. We assume that gauge symmetry lives on $D(3+q)$-branes which are extended along the four noncompact dimensions and wrap a q-cycle Q in the compactification manifold. In Type IIA, one takes a stack of five $D6$-branes wrapped around $M \times Q$ where M is the Minkowski space and Q is a compact special Lagrangian three-cycle in the compact manifold X. In Type IIB, one takes instead $D3$, $D5$ or $D7$-branes wrapping a holomorphic cycle in X. Three D-branes wrapping a cycle Q support U(3) gauge theory whose nonabelian part could be the QCD gauge symmetry. Five D-branes would lead instead to a theory somewhat similar to an SU(5) GUT. In case that the q-cycle Q has nonzero $\pi_1(Q)$, we can break the GUT gauge group down to the Standard Model gauge group by turning on discrete Wilson lines.

The low energy effective supergravity action contains the gravitational term (2.24)

$$S_{GR} = \frac{2\pi}{g_s^2 \ell_s^8} \int \sqrt{-g} R. \tag{7.1}$$

Dimensionally reducing the Einstein action to four dimensions determines the Planck mass

$$M_P^2 = 4\pi \frac{V_X}{g_s^2 \ell_s^8}. \tag{7.2}$$

The low energy action of the RR q-form field C_q is, [54] eq. (13.3.5),

$$-\frac{2\pi}{\ell_s^{8-2q}} \int \frac{1}{2} F_{q+1} \wedge \star F_{q+1} + 2\pi \int C_q, \tag{7.3}$$

where the second term is the coupling of the q form to $D(q-1)$ branes. We normalized C_q so that the RR field strength F_{q+1} has integer periods. Thus, our RR-field is related to the usual one $C_{q,conv}$ by $C_{q,conv} = \ell_s^q C_q$.

The effective action of the gauge theory living on the D-branes is, [54] eq. (13.3.25)

$$S_{YM} = -\frac{1}{4(2\pi)g_s\ell_s^{4+q}} \int d^q x \sqrt{-g}\, \mathrm{tr}\, F_{\mu\nu}F^{\mu\nu}, \tag{7.4}$$

where the trace is in the fundamental $\mathbf{N}$ representation of SU(N). Reducing the gauge action on Q to four dimensions leads to the action

$$-\frac{V_Q}{8(2\pi)g_s\ell_s^q} \int d^4 x \sqrt{-g} F_{\mu\nu}^a F^{\mu\nu\,a}, \tag{7.5}$$

where we used the normalization of SU(N) generators $\mathrm{tr}\, t^a t^b = \frac{1}{2}\delta^{ab}$ that is conventional in the GUT literature. From (7.5) we read off the four-dimensional gauge coupling

$$\alpha_{\mathrm{GUT}} = \frac{g_s\ell_s^q}{V_Q}. \tag{7.6}$$

In type II string theory, the axions come from zero modes of the q-form gauge field C_q. The axions are the phases $2\pi \int_Q C_q$ necessary for the complete definition of the string theory path integral in the presence of $D(q-1)$ branes. They are angular variables with period 2π. Let Q_α, $\alpha = 1, \ldots, b_q(X)$ be an integral basis of the homology group $H_q(X, \mathbb{Z})$ modulo torsion. We take ω_α to be harmonic representatives of the basis of $H^q(X)$ dual to the basis Q_α, so that $\int_{Q_\alpha} \omega_\beta = \delta_{\alpha\beta}$.

The axions come from the ansatz

$$C_q = \frac{1}{2\pi} \sum_\alpha a_\alpha \omega_\alpha, \qquad \alpha = 1, \ldots, b_q(X). \tag{7.7}$$

We included a factor of $1/2\pi$ so that a_α have period 2π. Substituting this into the RR-field effective action (7.3) we get the kinetic energy of the axions

$$S = -\frac{1}{2} \sum_{\alpha,\beta} \gamma_{\alpha\beta} \partial_\mu a_\alpha \partial^\mu a_\beta, \tag{7.8}$$

where

$$\gamma_{\alpha\beta} = \frac{1}{2\pi\ell_s^{8-2q}} \int_X \omega_\alpha \wedge \star\omega_\beta. \tag{7.9}$$

The axions acquire axionic couplings from the D-brane Chern-Simons term, (13.3.18) of [54]

$$2\pi \int_{M\times Q} C_q \wedge \frac{1}{8\pi^2}\, \mathrm{tr}\, F \wedge F. \tag{7.10}$$

Dimensionally reducing this to four dimensions using the ansatz (7.7) leads to the couplings

$$\sum_\alpha r_\alpha \int a_\alpha \frac{\mathrm{tr}\, F \wedge F}{8\pi^2}, \tag{7.11}$$

where $r_\alpha = \int_Q \omega_\alpha$ are integers.

Let us first consider branes wrapping a q-cycle Q with $q > 0$. The case $q = 0$ with a stack of $D3$ branes localized at a point will be considered later. We define R to be linear size of X, so that $V_X = R^6$. In terms of R, the Planck mass is

$$M_P^2 = \frac{4\pi R^6}{g_s^2 \ell_s^8}. \tag{7.12}$$

For a generic axion, $\int_X \omega_Q \wedge \star \omega_Q = x R^{6-2q}$, where x is of order one, so

$$F_a = \sqrt{\frac{x R^{6-2q}}{2\pi \ell_s^{8-2q}}} = M_P \left(\frac{\ell_s}{R}\right)^q \sqrt{\frac{x g_s^2}{8\pi^2}}, \tag{7.13}$$

where x is a dimensionless number of order one. Hence F_a is in the phenomenologically preferred range if either $g_s \ll 1$ or $R \gg \ell_s$. The first possibility is excluded by considerations about gauge couplings. Indeed, if the gauge symmetry comes from D-branes wrapping a q-cycle Q of radius $R_Q = V_Q^{1/q}$, the four-dimensional gauge coupling is $\alpha_C = g_s \ell_s^q / R_Q^q$. To have a good perturbative description, R_Q should not be much smaller than ℓ_s so that α' corrections are suppressed (otherwise we can go into a T-dual description, in which $R_Q \gtrsim \ell_s$. Setting $R_Q \gtrsim \ell_s$ gives an upper bound on the string coupling $g_s \gtrsim \alpha_C$. Hence, we do not have a freedom to lower g_s to arbitrarily small values. The latter possibility leads to low scale axions if the compactification manifold has very large size in string units.

To estimate the parameters of the compactification that lead to phenomenologically preferred axion decay constants, we express R and $M_s = \ell_s^{-1}$ in terms of F_a, M_P from (7.12) and (7.13):

$$R = \ell_s \left(\frac{M_P}{F_a}\right)^{\frac{1}{q}} \left(\frac{x g_s^2}{8\pi^2}\right)^{\frac{1}{2q}},$$

$$M_s = F_a \left(\frac{F_a}{M_P}\right)^{\frac{3-q}{q}} \left(\frac{2\pi}{x}\right)^{\frac{1}{2}} \left(\frac{8\pi^2}{x g_s^2}\right)^{\frac{3-q}{2q}}. \tag{7.14}$$

Requiring that the axion decay constant falls into the range $10^9 \text{GeV} \leq F_a \leq 10^{12} \text{GeV}$ implies large compactification radius $R \gg \ell_s$ and a low string scale $M_s \ll M_{\text{GUT}}$.

In these compactifications, gauge symmetry lives on D-branes wrapping a cycle Q. The radius of Q is at most a few string lengths; otherwise, the string coupling necessary for getting the correct four-dimensional gauge coupling $g_s = \alpha_C (R/\ell_s)^q$ get nonperturbatively large and our target space effective description breaks down. Hence, Q is a 'vanishing' cycle in X with $R_Q \ll R$. A possible moduli stabilization mechanism that fixes the moduli of X in this regime has been recently discussed in [86, 39]. The actual results for the preferred compactification parameters depend on which q-form RR-field does the axion originate from and on the geometry of the compactification manifold. In section 7.1, we give a more precise treatment of compactifications with vanishing cycles in two concrete examples.

Asymmetric Calabi-Yau manifolds. As an alternative to assuming that Q is a vanishing cycle, one can ask whether taking an asymmetric Calabi-Yau manifold that is a fibration with small fiber F over a large base B could substantially lower the axion coupling parameter. Based on our experience with heterotic string in section 3 we expect that the decay constants of axions in asymmetric compactifications are still around the GUT scale.

To explore this question, we will consider supersymmetric fibrations with T^2, T^3 or $K3$ fibers, but our conclusions apply to more general fibrations (including nonsupersymmetric ones). The volume of the Calabi-Yau manifold is $V_X = V_B V_F$, hence the reduced Planck mass (7.2) is

$$M_P^2 = \frac{4\pi V_B V_F}{g_s^2 \ell_s^8}. \tag{7.15}$$

We let R_F be the 'radius' of the fibre, so that $V_F = R_F^d$ and R_B be the 'radius' of the base $V_B = R_B^{6-d}$. An axion coming from a zero mode of the q-form field C_q with b indices along the fibre and $q - b$ indices along the base has axion coupling parameter (7.9)

$$F_a^2 = \frac{x V_X}{2\pi \ell_s^{8-2q} R_B^{2q-2b} R_F^{2b}}, \tag{7.16}$$

where x is a dimensionless number of order one. We assume that the gauge symmetry comes from a $D(3 + q)$-branes wrapping a cycle with a dimensions wrapped around the fibre and $q - a$ dimensions wrapped along the base, so the gauge coupling is

$$\alpha_C = \frac{g_s \ell_s^q}{R_F^a R_B^{q-a}}. \tag{7.17}$$

The axionic coupling is nonzero only for axions with $b = a$. For these, we find

$$F_a = \frac{\sqrt{x} \alpha_C}{2\pi} \frac{M_P}{\sqrt{2}}, \tag{7.18}$$

which gives the familiar answer $\sim 10^{16} \text{GeV}$.

$D3$-brane models. The cycle that the D3-branes "wrap" is a point which results in a different behavior of the axion decay parameter compared to other string theories. Hence, there is no hierarchy between the size of the vanishing cycle and the size of the compactification manifold that could help lower the axion decay constant. The low energy gauge group on N $D3$-branes at a generic point in X is $U(N)$. The gauge coupling is fixed by the string coupling (7.4):

$$\alpha_C = g_s. \tag{7.19}$$

The axions are four-dimensional fields coming from reduction of the RR zero-form. A harmonic zero-form is just a constant, so we use the ansatz

$$C_0 = \frac{a}{2\pi}, \tag{7.20}$$

where a is a four-dimensional pseudo-scalar field. It follows from the D-brane Chern-Simons coupling (7.10) that the axion has $r = 1$ coupling to the QCD instanton density

$$\int a \frac{\text{tr} \, F \wedge F}{8\pi^2}. \tag{7.21}$$

The kinetic energy of the RR zero-form (7.3) is easily reduced to four dimensions, giving the axion kinetic energy

$$\frac{V_X}{2\pi \ell_s^8} \int d^4 x \left(-\frac{1}{2} \partial_\mu a \partial^\mu a \right), \tag{7.22}$$

whence the axion coupling constant is

$$F_a = \sqrt{\frac{V_X}{2\pi\ell_s^8}} = \frac{\alpha_C}{2\pi}\frac{M_P}{\sqrt{2}}. \tag{7.23}$$

If we take $\alpha_C \sim 1/25$, we get $F_a = 1.1 \times 10^{16}$GeV, which is the same as the axion coupling parameter of the model-independent axion in weakly coupled heterotic string theory. The shift symmetry of the axion is explicitly broken by $D(-1)$-brane instantons that are located on the $D3$-brane worldvolume. These instantons are equivalent the $SU(N)$ gauge theory instantons. Their action is $I = 2\pi/g_s = 2\pi/\alpha_C$. With $\alpha_C \sim 1/25$, their action is $I \sim 157$, so the explicit violation of the shift symmetry might be small enough for the axion to be a candidate for Peccey-Quinn axion.

7.1 Intersecting D-brane models with small cycles

In the previous subsection we found that the axion coupling parameter can be lowered in type II string theory, if the the gauge symmetry comes from D-branes wrapping a vanishing cycle. To lower F_a into the range 10^9GeV $< F_a < 10^{12}$GeV, the radius of the compactification manifold has to be much larger than the string length. This lowers the string scale relative to the Planck scale. We estimated this in (7.14).

To get a more precise estimate of the physical scales involved in getting F_a in the phenomenologically preferred range, we study in detail the compactification of Type II string on a CY manifold X that is developing a conifold singularity. The vanishing cycle at the tip is either an S^2 or an S^3, depending on whether the conifold is resolved or deformed. To get $SU(5)$ gauge symmetry, we wrap a stack of five $D5$ or $D6$-branes around the vanishing cycle. The D-branes warp the geometry in a region of size $\ell \sim (g_s N)^{1/4}\ell_s$, as is familiar from AdS/CFT correspondence. For $N = 5$, that is necessary for $SU(5)$ gauge symmetry, the warped region has size around the string length. Since we took the radius of the CY manifold much larger than the string length, we can neglect the effect of the warping on the axion coupling parameter $F_a \sim \int_X \omega \wedge \star\omega$.

The conifold is a cone over $T^{1,1}$, where $T^{1,1}$ is topologically $S^2 \times S^3$. It is an S^1 fibration over $S^2 \times S^2$. Its the metric is [87, 88]

$$ds^2 = dr^2 + r^2 ds_{T^{1,1}}. \tag{7.24}$$

To describe the metric of $T^{1,1}$, we parametrize the S^1 fiber with ψ, which ranges from 0 to 4π and the two S^2's with spherical coordinates $(\theta_i, \phi_i), i = 1, 2$. We introduce the following basis of one-forms [88]

$$\begin{aligned}
g^1 &= \frac{e^1 - e^3}{\sqrt{2}}, \quad g^2 = \frac{e^2 - e^4}{\sqrt{2}}, \\
g^3 &= \frac{e^1 + e^3}{\sqrt{2}}, \quad g^4 = \frac{e^2 + e^4}{\sqrt{2}}, \\
g^5 &= e^5,
\end{aligned} \tag{7.25}$$

where

$$
\begin{aligned}
e^1 &= -\sin\theta_1 d\phi_1, \quad e^2 = d\theta_1, \\
e^3 &= \cos\psi\sin\theta_2 d\phi_2 - \sin\psi d\theta_2, \\
e^4 &= \sin\psi\sin\theta_2 d\phi_2 + \cos\psi d\theta_2, \\
e^5 &= d\psi + \cos\theta_1 d\phi_1 + \cos\theta_2 d\phi_2.
\end{aligned}
\tag{7.26}
$$

In terms of these, the $T^{1,1}$ metric takes the form

$$
ds_{T^{1,1}} = \frac{1}{9}(g^5)^2 + \frac{1}{6}\sum_{i=1}^{4}(g^i)^2.
\tag{7.27}
$$

On the conifold, there are harmonic two- and three-forms

$$
\begin{aligned}
\omega_2 &= \frac{1}{2}(g^1 \wedge g^2 + g^3 \wedge g^4), \\
\omega_3 &= g^5 \wedge \omega_2.
\end{aligned}
\tag{7.28}
$$

Their Hodge duals are

$$
\star\omega_2 = \frac{r}{3}dr \wedge \omega_3 \quad \star\omega_3 = -3\frac{dr}{r} \wedge \omega_2.
\tag{7.29}
$$

If we think of $T^{1,1}$ as an S^2 fibration over S^3, then ω_2 has nonzero flux through the S^2 fiber and ω_3 has nonzero flux through the S^3 base of the fibration. To find these fluxes, we take a representative S^2 fiber with $\psi = 0, \theta_1 = \theta_2$ and $\phi_1 = -\phi_2$. The S^3 base can be defined with the equations $\psi_2 = \phi_2 = 0$. Integrating the explicit expressions (7.28) for the harmonic two and three-forms over the cycle representatives gives

$$
\int_{S^2} \omega_2 = 4\pi, \quad \int_{S^3} \omega_3 = 8\pi^2.
\tag{7.30}
$$

The use of these formulas is the following. The cone over $T^{1,1}$ can be slightly resolved or deformed to make a smooth six-manifold X' with a small S^2 or S^3 at its center. (We think of X' as an approximation to part of a compact Calabi-Yau manifold X.) We obtain gauge theory by wrapping D-branes on the vanishing cycle, that is, on the small S^2 or S^3. The gauge coupling is inversely proportional to the volume of the vanishing cycle, and so depends crucially on the details of the resolution or deformation of the cone. The axion that couples to these gauge fields comes from a harmonic two-form or three-form on X, which we approximate by a harmonic two-form on X'. As there is no $\mathbb{L}^2$ harmonic two-form or three-form on X', this form is not supported near the vanishing cycle, and in describing it we can simply approximate X' by the cone. The relevant harmonic two-form and three-form on the cone are simply the pullbacks of the harmonic forms ω_2 and ω_3 on $T^{1,1}$.

Type IIB. Let us first consider the type IIB string theory with gauge symmetry coming from $D5$-branes wrapping the vanishing S^2. If the radius of the S^2 is r_0, the gauge coupling is

$$\alpha_{\text{GUT}} = \frac{g_s \ell_s^2}{4\pi r_0^2}. \tag{7.31}$$

The axion comes from a mode of the RR two-form field C_2 that is the product of a four-dimensional field a with a harmonic two-form field that has nonzero flux over the vanishing cycle S^2. Since this harmonic two-form is not supported near the vanishing cycle, we can approximate it by a harmonic form on the cone, which in fact is a pullback of the harmonic form ω_2 from the five-manifold $T^{1,1}$. Thus, the ansatz for the RR two-form is

$$C_2 = \frac{\omega_2}{4\pi} \frac{a}{2\pi}. \tag{7.32}$$

We used (7.30) to normalize a to have 2π periods. To find the axion decay constant, we substitute the ansatz (7.32) into the formula for the axion decay constant (7.9). With the help of (7.29), we get

$$F_b^2 = \frac{1}{3(32\pi^3)\ell_s^4} \int r\,dr \int_{T^{1,1}} \omega_2 \wedge \omega_3. \tag{7.33}$$

According to (7.30), the integral over $T^{1,1}$ is $\int_{T^{1,1}} \omega_2 \wedge \omega_3 = 32\pi^3$. We estimate the integral over the radial direction of the cone with $\int r\,dr = xR^2/2$. Here, $R = V_X^{1/6}$ is the size of X and x is a number of order one that depends on the details of X. Thus, the axion supported near the vanishing S^2 has

$$F_b = \sqrt{\frac{x}{6}} \frac{R}{\ell_s^2}. \tag{7.34}$$

From (7.2) and (7.34), we express R and $M_s = \ell_s^{-1}$ as

$$R = \ell_s g_s^{1/2} \left(\frac{M_P}{F_b}\right)^{1/2} \left(\frac{x}{24\pi}\right)^{1/4},$$

$$M_s = \frac{F_b}{g_s^{1/2}} \left(\frac{F_b}{M_P}\right)^{1/2} \left(\frac{2^5 3^3 \pi}{x^3}\right)^{1/4}. \tag{7.35}$$

For F_b in the range preferred by phenomenological considerations $10^9 \text{GeV} \lesssim F_a \lesssim 10^{12} \text{GeV}$, with $x = 1$, we have

$$1.5 \times 10^5 \text{GeV} \lesssim M_s g_s^{1/2} \lesssim 5 \times 10^9 \text{GeV},$$

$$1.6 \times 10^4 \ell_s \gtrsim \frac{R}{g_s^{1/2}} \gtrsim 5 \times 10^2 \ell_s. \tag{7.36}$$

To assess whether the explicit breaking of the Peccei-Quinn symmetry is sufficiently small, we estimate the actions of instantons that break it. Generic instantons are Euclidean $D1$-branes wrapping cycles of size $\sim R$. They have very large action $I \sim 2\pi R^2/g_s \ell_s^2$, since $R \gg \ell_s$. It follows that these instantons break Peccei-Quinn symmetry negligibly. The main violation of the PQ-symmetry comes from $D1$-brane instantons that wrap the vanishing S^2. They have action $I = 2\pi V_Q/g_s \ell_s^2 = 2\pi/\alpha_{\text{GUT}} \sim 157$. Hence, with some help from low-energy supersymmetry, the PQ-symmetry might be able to explain the strong CP-problem.

Type IIA. In IIA string theory, we get gauge symmetry by wrapping $D6$-branes around the small S^3 of the deformed conifold. If the S^3 has radius r_0, the gauge coupling is

$$\alpha_{\text{GUT}} = \frac{g_s \ell_s^3}{2\pi^2 r_0^3}. \tag{7.37}$$

The axions is a four-dimensional scalar b coming from a zero mode of the RR three-form field C_3:

$$C_3 = \frac{\omega_3}{8\pi^2} \frac{b}{2\pi}. \tag{7.38}$$

ω_3 is a harmonic three-form on X with a nonzero flux through the vanishing S^3. We approximate it by a harmonic form on the cone, which is a pullback of the harmonic form ω_3 (7.28) on $T^{1,1}$. With the help of (7.30), we normalized the C-field so that the axion b has period 2π. We find F_b from the general formula for the decay constant of an RR-axion (7.9)

$$F_b^2 = \frac{1}{2\pi \ell_s^2} \left(\frac{1}{8\pi^2} \right)^2 \int_X \omega_3 \wedge \star \omega_3 = \frac{3x}{4\pi^2 \ell_s^2} \ln \left(\frac{R}{r_0} \right), \tag{7.39}$$

where x is a dimensionless number of order one. In the evaluation of the integral, we approximated the space X with just the conical region. The integral over the radial direction of the cone diverges both for large and small radius. We cut off the large distance divergence of the integral at the radius R of X and the short distance divergence at the radius r_0 of the vanishing S^3. Since the harmonic three-form is not supported near the vanishing cycle, the main contribution to the integral comes from the logarithm $\ln(R/r_0)$. We are justified to neglect the corrections from the region near the tip of the deformed conifold and from the region that compactifies the conifold, as long as $R \gg r_0$. If we assume that the gauge coupling at the string scale is $\alpha_{\text{GUT}} \sim 1/25$, it follows from (7.37) that $r_0 \sim \ell_s$. But we already know from our estimate (7.14) that $R \gg \ell_s$, whence it follows that $R \gg r_0$ and our approximations are self-consistent.

To find the range of the string compactification parameters that lead to phenomenologically acceptable axion, we express M_s and R from (7.39) and (7.2), as

$$M_s = \frac{2\pi F_b}{\sqrt{3x \ln (R/r_0)}},$$
$$R = \ell_s \left(\frac{M_P}{F_a} \right)^{1/3} g_s^{1/3} \left(\frac{3x \ln (R/r_0)}{2^4 \pi^3} \right)^{1/6}. \tag{7.40}$$

For $10^9 \text{GeV} \lesssim F_b \lesssim 10^{12} \text{GeV}$ and $x = 1$, we have

$$1.4 \times 10^9 \text{GeV} \lesssim M_s \lesssim 1.8 \times 10^{12} \text{GeV},$$
$$800 \ell_s \gtrsim \frac{R}{g_s^{1/3}} \gtrsim 73 \ell_s. \tag{7.41}$$

8. Type I string theory

Just like heterotic string theory, compactifications of type I string has a model dependent-axion and a number of model-dependent axions. These come from zero modes of the RR

two-form field C. The computations of axion properties in type I string are analogous to the ones in heterotic string. Here, we will illustrate them in the case of model-independent axion. For some aspects of axion physics in Type I orbifold compactifications, see [89].

The type I supergravity action is (12.1.34) of [54]

$$S = \frac{2\pi}{\ell_s^8 g_s^2} \int d^{10}x (-g)^{1/2} R - \frac{4\pi}{\ell_s^4} \int \frac{1}{2} F_3 \wedge \star F_3 - \frac{\sqrt{2}}{4(2\pi)g_s\ell_s^6} \int \mathrm{tr}\, F_2 \wedge \star F_2. \qquad (8.1)$$

where F_2 is the SO(32) gauge field strength and tr is the trace in the **32** vector representation of SO(32). We fixed the normalization of the string coupling g_s using the convention

$$\kappa^2 = g_s^2 \kappa_{10}^2 = g_s^2 \ell_s^8/4\pi, \qquad (8.2)$$

and we substituted the gauge coupling from eq. (13.3.31) of [54]:

$$g_{YM}^2 = 2(2\pi)^{3/2}\ell_s^2\kappa = \sqrt{2}(2\pi)g_s\ell_s^6. \qquad (8.3)$$

We normalized the RR-two form, so that its field-strength, (12.1.35) of [54]

$$F_3 = dC - \frac{1}{16\pi^2}\omega_3(A) \qquad (8.4)$$

has integer periods. Thus, our C-field is related to C_{conv} of [54] via $C_{\mathrm{conv}} = \sqrt{2}\ell_s^2 C$. From (8.1) we read off the four-dimensional gauge and gravitational couplings

$$\alpha_{\mathrm{GUT}} = \frac{g_s\ell_s^6}{\sqrt{2}V_X} \qquad M_P^2 = \frac{4\pi V_X}{g_s^2\ell_s^8}. \qquad (8.5)$$

Here, we assume usual embedding SU(5) $\subset$ SO(10) $\subset$ SO(32) of SU(5) in SO(32).

The model-independent axion, comes from a mode of the C field constant on X with all indices along the Minkowski space. As explained in section 3 in our study of model-independent axion in heterotic string theory, the axion decay constant is the inverse of the coefficient of the C-field kinetic energy

$$F_a = \frac{\ell_s^2}{\sqrt{4\pi V_X}} = \frac{M_P}{\sqrt{2}} \frac{\alpha_{\mathrm{GUT}}}{2\pi}. \qquad (8.6)$$

This gives $F_a = 1.1 \times 10^{16}$GeV, which is same as the axion decay constant of the heterotic string model-independent axion. Indeed, under the heterotic-type I duality, the RR C-field model independent axion of type I string goes into the NS-NS B-field model independent axion of heterotic string theory.

9. Axion coupling to photons

This concluding section is devoted to a topic in axion physics that has nothing to do with string theory. We will reconsider a matter discussed in section 2, namely the axion-photon coupling. This coupling has a contribution from short distance axion physics and another contribution from low energy QCD strong coupling effects. In section 2, we evaluated

the first type of contribution in a unified SU(5) gauge theory. To get the physical axion-photon coupling, one must also evaluate QCD strong coupling effects that mix the axion with the π^0. These effects have been previously determined in [90 – 92] using current algebra techniques. Here we will provide a new derivation of this coupling based on the relevant low energy effective Lagrangian.

In section 2, we described the kinetic term and the mass term of the low energy effective Lagrangian describing pion physics

$$S_0 = -\frac{F_\pi^2}{16} \int \text{tr}\left(U^{-1}\partial_\mu U\right)\left(U^{-1}\partial^\mu U\right) + \frac{v}{2}\text{tr}\left(MU + \overline{M}U^\dagger\right). \tag{9.1}$$

Here U takes values in the group manifold SU(3) and $F_\pi = 184\text{MeV}$ is the pion decay constant. U is conventionally parametrized as

$$\text{U}(x) = \exp\left(\frac{2i}{F_\pi}\sum_{a=1}^{8}\lambda^a\pi^a(x)\right), \tag{9.2}$$

where λ^a are the generators of the SU(3) algebra normalized as $\text{Tr}\,\lambda^a\lambda^b = 2\delta^{ab}$. If we integrate out the s quark, which is much heavier than the u, d quarks, then U is an element of SU(2) and we take $\lambda^a = \sigma^a$, $a = 1, 2, 3$ to be the Pauli matrices.

The anomalous couplings of pions are summarized in the Wess-Zumino-Witten term [93, 94]. To write it down, we let D be a five-dimensional space bounding our four-dimensional spacetime. We extend U to a map from D to SU(3). The WZW term is [94]

$$\Gamma = -\frac{iN_c}{2\pi^2 \times 5!}\int_D d\Sigma^{ijklm}(U^{-1}\partial_i U)(U^{-1}\partial_j U)(U^{-1}\partial_k U)(U^{-1}\partial_l U)(U^{-1}\partial_m U), \tag{9.3}$$

where $d\Sigma^{ijklm}$ is the five-dimensional volume element on D. The integrand is N_c times the generator of $H^5(\text{SU}(3), 2\pi\mathbb{Z})$, where $N_c = 3$ is the number of colors. The factor of 2 in the denominator comes from the Bott periodicity theorem [95].

9.1 Gauging electromagnetism

The WZW action (9.3) captures the anomalous couplings between the pions and axions. The couplings of the pion and axion to the photon can be described by a gauged version of the WZW term. Let us discuss briefly how one gauges the action $S = S_0 + \Gamma$. For further details, see [94]. The action S is invariant under the global U(1)$_{\text{EM}}$ symmetry

$$\delta U = i\epsilon[\mathcal{Q}, U], \tag{9.4}$$

where $\mathcal{Q}$ is the electric charge matrix of the u, d, and s quarks

$$\mathcal{Q} = \begin{pmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix}. \tag{9.5}$$

We now promote U(1)$_{EM}$ to a local gauge symmetry. This is accomplished as follows. The kinetic term S_0 becomes gauge-invariant if we replace the ordinary derivatives with covariant ones

$$D_\mu = \partial_\mu + ieA_\mu[\mathcal{Q}, \ ]. \tag{9.6}$$

The U(1)$_{EM}$ gauge field is canonically normalized so it transforms as $A_\mu \to A_\mu - \partial_\mu \lambda$ under a gauge transformation λ.

The WZW term cannot be written in four dimensions as an integral of a manifestly SU(3) × SU(3) invariant expression, so the standard gauging procedure is not applicable to it. Instead, the gauge invariant generalization of Γ can be obtained using the Noether procedure. A short calculation shows that the desired modification of Γ that makes it gauge invariant is [94]

$$\widetilde{\Gamma} = \Gamma - e \int d^4 x A_\mu J^\mu + \frac{ie^2}{24\pi^2} \int d^4 x \, \epsilon^{\mu\nu\alpha\beta} \left(\partial_\mu A_\nu \right) A_\alpha$$
$$\times \operatorname{tr} \left[\mathcal{Q}^2 \left(\partial_\beta U \right) U^{-1} + \left(\partial_\beta U \right) \mathcal{Q}^2 U^{-1} + \mathcal{Q} \left(\partial_\beta U \right) \mathcal{Q} U^{-1} \right], \tag{9.7}$$

where

$$J^\mu = \frac{1}{48\pi^2} \epsilon^{\alpha\beta\gamma\delta} \operatorname{tr} \Big[\mathcal{Q} \left(\partial_\nu U U^{-1} \right) \left(\partial_\alpha U U^{-1} \right) \left(\partial_\beta U U^{-1} \right)$$
$$+ \mathcal{Q} \left(U^{-1} \partial_\nu U \right) \left(U^{-1} \partial_\alpha U \right) \left(U^{-1} \partial_\beta U \right) \Big]. \tag{9.8}$$

The reason that we started with three flavors was that this gave a convenient way to determine the anomalous interactions, which are summarized in (9.7). Now that we have determined those interactions, we can for our purposes here omit the strange quarks and reduce to the case of two flavors.

9.2 Axion-photon coupling

In eq. (2.23), we determined the part of the axion-photon coupling coming from explicit coupling of the axion to the SU(5) gauge fields:

$$\frac{4r}{3} \frac{\alpha}{8\pi F_a} a \, \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} = \frac{4r}{3\pi} \frac{\alpha}{F_a} a \, \vec{E} \cdot \vec{B}. \tag{9.9}$$

Here a is the canonically normalized axion which has periods $2\pi F_a$. Now, we will compute the additional contribution to the axion-photon coupling due to the mixing of the axion with the neutral pion. The axion enters the effective action as a phase of the determinant of the quark mass matrix. In the two-flavor approximation, we can take this mass matrix to be

$$M = \begin{pmatrix} \exp(-ic_u a/F_a) m_u & 0 \\ 0 & \exp(-ic_d a/F_a) m_d \end{pmatrix}, \tag{9.10}$$

where $c_u + c_d = 1$ so that $\det(M) \propto \exp(-ia)$. The (π^0, a) mass matrix comes from expanding the mass term (9.1) to quadratic order

$$\mathcal{L}_m = \frac{v}{2} \operatorname{tr} \left(M U + \overline{M} U^\dagger \right)$$
$$= -\frac{2v}{F_\pi^2} (m_u + m_d) \left[\pi^0 - \frac{a F_\pi}{2 F_a} \frac{c_u m_u - c_d m_d}{m_u + m_d} \right]^2 - \frac{v}{2 F_a^2} \frac{m_u m_d}{m_u + m_d} a^2. \tag{9.11}$$

From (9.11) we read off the masses of the pion and the axion

$$m_\pi^2 = \frac{4v}{F_\pi^2} (m_u + m_d), \qquad m_a^2 = \frac{m_\pi^2 F_\pi^2}{4 F_a^2} \frac{m_u m_d}{(m_u + m_d)^2}. \tag{9.12}$$

The anomalous QCD contribution to the coupling of axion and pion to two photons comes from the last term in the gauged WZW term (9.7), since this is the only piece quadratic in the $U(1)_{EM}$ gauge field. To find these couplings, we perform the $SU(2)_L$ axial rotation

$$U \to \begin{pmatrix} \exp(+ic_u\frac{a}{F_a}) & 0 \\ 0 & \exp(+ic_d\frac{a}{F_a}) \end{pmatrix} U \tag{9.13}$$

and expand (9.7) to first order in π^0 and a. The couplings to two photons are

$$-\frac{\alpha}{8\pi}\left(\frac{a}{3F_a}(4c_u + c_d) + 2\frac{\pi^0}{F_\pi}\right)\epsilon^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta}. \tag{9.14}$$

To determine the axion-photon coupling from (9.14) we take a shortcut and set $c_u m_u - c_d m_d = 0$ so that the axion does not mix with the neutral pion in the axion-pion mass matrix (9.11). This condition together with the constraint $c_u + c_d = 1$ determines $c_u = m_d/(m_u + m_d), c_d = m_u/(m_u + m_d)$. Substituting this into the axion-photon vertex (9.14) gives the axion-photon coupling

$$-\frac{r}{3}\frac{\alpha}{8\pi F_a}\frac{m_u + 4m_d}{m_u + m_d}a\,\epsilon^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta}. \tag{9.15}$$

To get the complete coupling of the axion to two photons we add to this the $SU(5)$ contribution (9.9)

$$\frac{r}{3}\frac{\alpha}{8\pi F_a}\left(4 - \frac{m_u + 4m_d}{m_u + m_d}\right)a\,\epsilon^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta} = -\frac{r\alpha}{\pi F_a}\frac{m_u}{m_u + m_d}a\,\vec{E}\cdot\vec{B}. \tag{9.16}$$

The coupling (9.16) depends on the mass ratio of of the light quarks and it vanishes for $m_u = 0$. In the nature $m_d/m_u \simeq 1.8/1$, hence the axion to two photons coupling is suppressed by a factor of ~ 4 compared to the $SU(5)$ result (9.9) alone without the QCD correction.

Acknowledgments

References

[1] A.A. Belavin, A.M. Polyakov, A.S. Shvarts and Y.S. Tyupkin, *Pseudoparticle solutions of the Yang-Mills equations*, *Phys. Lett.* **B 59** (1975) 85.

[2] G. 't Hooft, *Symmetry breaking through Bell-Jackiw anomalies*, *Phys. Rev. Lett.* **37** (1976) 8.

[3] R. Jackiw and C. Rebbi, *Vacuum periodicity in a Yang-Mills quantum theory*, *Phys. Rev. Lett.* **37** (1976) 172.

[4] C.G. Callan Jr., R.F. Dashen and D.J. Gross, *The structure of the gauge theory vacuum*, *Phys. Lett.* **B 63** (1976) 334.

[5] R.D. Peccei and H.R. Quinn, *CP conservation in the presence of instantons*, *Phys. Rev. Lett.* **38** (1977) 1440.

[6] P.G. Harris et al., *New experimental limit on the electric dipole moment of the neutron*, *Phys. Rev. Lett.* **82** (1999) 904.

[7] M.V. Romalis, W.C. Griffith and E.N. Fortson, *A new limit on the permanent electric dipole moment of ^{199}Hg*, *Phys. Rev. Lett.* **86** (2001) 2505 [hep-ex/0012001].

[8] M. Pospelov and A. Ritz, *Theta induced electric dipole moment of the neutron via QCD sum rules*, *Phys. Rev. Lett.* **83** (1999) 2526 [hep-ph/9904483].

[9] S. Weinberg, *The problem of mass*, *Trans. N.Y. Acad. Sci.* **38** (1977) 185

[10] D.B. Kaplan and A.V. Manohar, *Current mass ratios of the light quarks*, *Phys. Rev. Lett.* **56** (1986) 2004.

[11] T. Banks, Y. Nir and N. Seiberg, *Missing (up) mass, accidental anomalous symmetries and the strong CP problem*, hep-ph/9403203.

[12] MILC collaboration, C. Aubin et al., *Light pseudoscalar decay constants, quark masses and low energy constants from three-flavor lattice QCD*, *Phys. Rev.* **D 70** (2004) 114501 [hep-lat/0407028].

[13] S. Weinberg, *A new light boson?*, *Phys. Rev. Lett.* **40** (1978) 223.

[14] F. Wilczek, *Problem of strong P and T invariance in the presence of instantons*, *Phys. Rev. Lett.* **40** (1978) 279.

[15] J.E. Kim, *Weak interaction singlet and strong CP invariance*, *Phys. Rev. Lett.* **43** (1979) 103.

[16] M. Dine, W. Fischler and M. Srednicki, *A simple solution to the strong CP problem with a harmless axion*, *Phys. Lett.* **B 104** (1981) 199.

[17] P. Sikivie, *Experimental tests of the 'invisible' axion*, *Phys. Rev. Lett.* **51** (1983) 1415.

[18] R. Bradley et al., *Microwave cavity searches for dark-matter axions*, *Rev. Mod. Phys.* **75** (2003) 777.

[19] CAST collaboration, K. Zioutas et al., *First results from the Cern axion solar telescope (cast)*, *Phys. Rev. Lett.* **94** (2005) 121301 [hep-ex/0411033].

[20] R.N. Mohapatra and G. Senjanovic, *Natural suppression of strong P and T noninvariance*, *Phys. Lett.* **B 79** (1978) 283.

[21] A.E. Nelson, *Naturally weak CP-violation*, *Phys. Lett.* **B 136** (1984) 387.

[22] S.M. Barr, *Solving the strong CP problem without the Peccei-Quinn symmetry*, *Phys. Rev. Lett.* **53** (1984) 329.

[23] G. Hiller and M. Schmaltz, *Solving the strong CP problem with supersymmetry*, *Phys. Lett.* **B 514** (2001) 263 [hep-ph/0105254].

[24] K.S. Babu, B. Dutta and R.N. Mohapatra, *Solving the strong CP and the SUSY phase problems with parity symmetry*, *Phys. Rev.* **D 65** (2002) 016005 [hep-ph/0107100].

[25] J.E. Kim, *Light pseudoscalars, particle physics, and cosmology*, *Phys. Rept.* **150** (1987) 1.

[26] M. Turner, *Windows on the axion*, *Phys. Rept.* **197** (1990) 67.

[27] J. Preskill, M.B. Wise and F. Wilczek, *Cosmology of the invisible axion*, *Phys. Lett.* **B 120** (1983) 127.

[28] L.F. Abbott and P. Sikivie, *A cosmological bound on the invisible axion*, *Phys. Lett.* **B 120** (1983) 133.

[29] M. Dine and W. Fischler, *The not-so-harmless axion*, *Phys. Lett.* **B 120** (1983) 137.

[30] M.B. Green and J.H. Schwarz, *Anomaly cancellation in supersymmetric $D = 10$ gauge theory and superstring theory*, *Phys. Lett.* **B 149** (1984) 117.

[31] E. Witten, *Some properties of $O(32)$ superstrings*, *Phys. Lett.* **B 149** (1984) 351.

[32] M.B. Green, J.H. Schwarz and E. Witten, *Superstring theory*, vol. 2, Cambridge University Press, 1987.

[33] S.M. Barr, *Harmless axions in superstring theories*, *Phys. Lett.* **B 158** (1985) 397.

[34] A.D. Linde, *Inflation and axion cosmology*, *Phys. Lett.* **B 201** (1988) 437.

[35] K. Choi and J.E. Kim, *Harmful axions in superstring models*, *Phys. Lett.* **B 154** (1985) 393; *Compactification and axions in $E_8 \times E_8'$ superstring models*, *Phys. Lett.* **B 165** (1985) 689.

[36] P. Fox, A. Pierce and S.D. Thomas, *Probing a QCD string axion with precision cosmological measurements*, `hep-th/0409059`.

[37] T. Banks and M. Dine, *Couplings and scales in strongly coupled heterotic string theory*, *Nucl. Phys.* **B 479** (1996) 173 [`hep-th/9605136`].

[38] T. Banks, M. Dine, P.J. Fox and E. Gorbatov, *On the possibility of large axion decay constants*, *JCAP* **06** (2003) 001 [`hep-th/0303252`].

[39] J.P. Conlon, *The QCD axion and moduli stabilisation*, *JHEP* **05** (2006) 078 [`hep-th/0602233`].

[40] M. Dine, N. Seiberg, X.G. Wen and E. Witten, *Nonperturbative effects on the string world sheet*, *Nucl. Phys.* **B 278** (1986) 769.

[41] K. Becker, M. Becker and A. Strominger, *Five-branes, membranes and nonperturbative string theory*, *Nucl. Phys.* **B 456** (1995) 130 [`hep-th/9507158`].

[42] G.R. Dvali, *Removing the cosmological bound on the axion scale*, `hep-ph/9505253`.

[43] M. Kawasaki, T. Moroi and T. Yanagida, *Can decaying particles raise the upperbound on the Peccei-Quinn scale?*, *Phys. Lett.* **B 383** (1996) 313 [`hep-ph/9510461`].

[44] T. Banks and M. Dine, *The cosmology of string theoretic axions*, *Nucl. Phys.* **B 505** (1997) 445 [`hep-th/9608197`].

[45] T. Banks, M. Dine and M. Graesser, *Supersymmetry, axions and cosmology*, *Phys. Rev.* **D 68** (2003) 075011 [`hep-ph/0210256`].

[46] P. Steinhardt and N. Turok, *A cyclic model of the universe*, *Science* **296** (2002) 1436; *The cyclic model simplified*, *New Astron. Rev.* **49** (2005) 43 [`astro-ph/0404480`].

[47] F. Wilczek, *A model of anthropic reasoning, addressing the dark to ordinary matter coincidence*, `hep-ph/0408167`.

[48] M. Romalis, private communication.

[49] S. Thomas, private communication.

[50] S. Weinberg, *The quantum theory of fields*, Vol. II, Cambridge University Press, 1996.

[51] C. Beasley and E. Witten, *New instanton effects in supersymmetric QCD*, *JHEP* **01** (2005) 056 [hep-th/0409149].

[52] J. Polchinski, *Monopoles, duality and string theory*, *Int. J. Theor. Phys.* **A19S1** (2004) 145 [hep-th/0304042].

[53] T. Banks, L.J. Dixon, D. Friedan and E.J. Martinec, *Phenomenology and conformal field theory or can string theory predict the weak mixing angle?*, *Nucl. Phys.* **B 299** (1988) 613.

[54] J. Polchinski, *String theory*, Cambridge University Press, Cambridge 1998.

[55] L.E. Ibáñez, *Gauge coupling unification: strings versus SUSY GUTs*, *Phys. Lett.* **B 318** (1993) 73 [hep-ph/9308365].

[56] M. Roček and E.P. Verlinde, *Duality, quotients and currents*, *Nucl. Phys.* **B 373** (1992) 630 [hep-th/9110053].

[57] E. Witten, *Strong coupling expansion of Calabi-Yau compactification*, *Nucl. Phys.* **B 471** (1996) 135 [hep-th/9602070].

[58] P. Hořava and E. Witten, *Eleven-dimensional supergravity on a manifold with boundary*, *Nucl. Phys.* **B 475** (1996) 94 [hep-th/9603142].

[59] G. Curio and A. Krause, *Four-flux and warped heterotic M-theory compactifications*, *Nucl. Phys.* **B 602** (2001) 172 [hep-th/0012152].

[60] G. Curio and A. Krause, *Enlarging the parameter space of heterotic M-theory flux compactifications to phenomenological viability*, *Nucl. Phys.* **B 693** (2004) 195 [hep-th/0308202].

[61] C. Vafa, *Evidence for F-theory*, *Nucl. Phys.* **B 469** (1996) 403 [hep-th/9602022].

[62] D.R. Morrison and C. Vafa, *Compactifications of F-theory on Calabi-Yau threefolds, I*, *Nucl. Phys.* **B 473** (1996) 74 [hep-th/9602114].

[63] D.R. Morrison and C. Vafa, *Compactifications of F-theory on Calabi-Yau threefolds, II*, *Nucl. Phys.* **B 476** (1996) 437 [hep-th/9603161].

[64] N. Arkani-Hamed, M. Porrati and L. Randall, *Holography and phenomenology*, *JHEP* **08** (2001) 017 [hep-th/0012148].

[65] J.J. Atick, L.J. Dixon and A. Sen, *String calculation of Fayet-Iliopoulos D terms in arbitrary supersymmetric compactifications*, *Nucl. Phys.* **B 292** (1987) 109.

[66] M. Dine, N. Seiberg and E. Witten, *Fayet-Iliopoulos terms in string theory*, *Nucl. Phys.* **B 289** (1987) 589.

[67] M. Dine, I. Ichinose and N. Seiberg, *F terms and D terms in string theory*, *Nucl. Phys.* **B 293** (1987) 253.

[68] J. March-Russell, *The Fayet-Iliopoulos term in type-I string theory and M-theory*, *Phys. Lett.* **B 437** (1998) 318 [hep-ph/9806426].

[69] N. Arkani-Hamed, M. Dine and S.P. Martin, *Dynamical supersymmetry breaking in models with a Green-Schwarz mechanism*, *Phys. Lett.* **B 431** (1998) 329 [hep-ph/9803432].

[70] S.M. Barr, *Harmless axions in superstring theories*, *Phys. Lett.* **B 158** (1985) 397.

[71] E. Poppitz, *On the one loop Fayet-Iliopoulos term in chiral four dimensional type-I orbifolds*, *Nucl. Phys.* **B 542** (1999) 31 [hep-th/9810010].

[72] Z. Lalak, S. Lavignac and H.P. Nilles, *String dualities in the presence of anomalous* U(1) *symmetries*, *Nucl. Phys.* **B 559** (1999) 48 [hep-th/9903160].

[73] L.E. Ibáñez, R. Rabadán and A.M. Uranga, *Anomalous* U(1)*'s in type-I and type-IIB D = 4, N = 1 string vacua*, *Nucl. Phys.* **B 542** (1999) 112 [hep-th/9808139].

[74] A. Lawrence and J. McGreevy, *D-terms and d-strings in open string models*, *JHEP* **10** (2004) 056 [hep-th/0409284].

[75] S. Kachru and J. McGreevy, *Supersymmetric three-cycles and (super)symmetry breaking*, *Phys. Rev.* **D 61** (2000) 026001 [hep-th/9908135].

[76] M.R. Douglas, *D-branes, categories and N = 1 supersymmetry*, *J. Math. Phys.* **42** (2001) 2818 [hep-th/0011017].

[77] T. Friedmann and E. Witten, *Unification scale, proton decay and manifolds of G_2 holonomy*, *Adv. Theor. Math. Phys.* **7** (2003) 577 [hep-th/0211269].

[78] R. Bryant and S. Salamon, *On the construction of some complete metrics with exceptional holonomy*, *Duke Math. J.* **58** (1989) 829.

[79] G.W. Gibbons, D.N. Page and C.N. Pope, *Einstein metrics on S^3, R^3 and R^4 bundles*, *Commun. Math. Phys.* **127** (1990) 529.

[80] M. Atiyah, J.M. Maldacena and C. Vafa, *An M-theory flop as a large-N duality*, *J. Math. Phys.* **42** (2001) 3209 [hep-th/0011256].

[81] M. Atiyah and E. Witten, *M-theory dynamics on a manifold of G_2 holonomy*, *Adv. Theor. Math. Phys.* **6** (2003) 1 [hep-th/0107177].

[82] T. Friedmann, *On the quantum moduli space of M-theory compactifications*, *Nucl. Phys.* **B 635** (2002) 384 [hep-th/0203256].

[83] S.B. Giddings, S. Kachru and J. Polchinski, *Hierarchies from fluxes in string compactifications*, *Phys. Rev.* **D 66** (2002) 106006 [hep-th/0105097].

[84] B. Acharya and E. Witten, *Chiral fermions from manifolds of G_2 holonomy*, hep-th/0109152.

[85] M. Cvetič, G.W. Gibbons, H. Lu and C.N. Pope, *Supersymmetric non-singular fractional D2-branes and NS-NS 2-branes*, *Nucl. Phys.* **B 606** (2001) 18 [hep-th/0101096].

[86] V. Balasubramanian, P. Berglund, J.P. Conlon and F. Quevedo, *Systematics of moduli stabilisation in Calabi-Yau flux compactifications*, *JHEP* **03** (2005) 007 [hep-th/0502058].

[87] P. Candelas and X.C. de la Ossa, *Comments on conifolds*, *Nucl. Phys.* **B 342** (1990) 246.

[88] C.P. Herzog, I.R. Klebanov and P. Ouyang, *Remarks on the warped deformed conifold*, hep-th/0108101.

[89] C. Angelantonj, I. Antoniadis, E. Dudas and A. Sagnotti, *Type-I strings on magnetised orbifolds and brane transmutation*, *Phys. Lett.* **B 489** (2000) 223 [hep-th/0007090].

[90] D.B. Kaplan, *Opening the axion window*, *Nucl. Phys.* **B 260** (1985) 215.

[91] M. Srednicki, *Axion couplings to matter, 1. CP conserving parts*, *Nucl. Phys.* **B 260** (1985) 689.

[92] H. Georgi, D.B. Kaplan and L. Randall, *Manifesting the invisible axion at low-energies*, *Phys. Lett.* **B 169** (1986) 73.

[93] J. Wess and B. Zumino, *Consequences of anomalous ward identities*, *Phys. Lett.* **B 37** (1971) 95.

[94] E. Witten, *Global aspects of current algebra*, *Nucl. Phys.* **B 223** (1983) 422.

[95] R. Bott and R. Seeley, *Some remarks on the paper of Callias*, *Commun. Math. Phys.* **62** (1978) 235.