

## Running of gauge couplings in string theory

Steven Abel<sup>1,2,\*</sup> Keith R. Dienes<sup>3,4,†</sup> and Luca A. Nutricati<sup>1,‡</sup>

<sup>1</sup>*IPPP and Department of Mathematical Sciences, Durham University,  
Durham, DH1 3LE, United Kingdom*

<sup>2</sup>*CERN, Theoretical Physics Department, CH 1211 Geneva 23 Switzerland*

<sup>3</sup>*Department of Physics, University of Arizona, Tucson, Arizona 85721, USA*

<sup>4</sup>*Department of Physics, University of Maryland, College Park, Maryland 20742, USA*



(Received 26 March 2023; accepted 27 April 2023; published 21 June 2023)

In this paper we conduct a general, model-independent analysis of the running of gauge couplings within closed string theories. Unlike previous discussions in the literature, our calculations fully respect the underlying modular invariance of the string and include the contributions from the infinite towers of string states which are ultimately responsible for many of the properties for which string theory is famous, including an enhanced degree of finiteness and UV/IR mixing. In order to perform our calculations, we adopt a formalism that was recently developed for calculations of the Higgs mass within such theories, and demonstrate that this formalism can also be applied to calculations of gauge couplings. In general, this formalism gives rise to an “on-shell” effective field theory (EFT) description in which the final results are expressed in terms of supertraces over the physical string states, and in which these quantities exhibit an EFT-like “running” as a function of an effective spacetime mass scale. We find, however, that the calculation of the gauge couplings differs in one deep way from that of the Higgs mass: while the latter results depend on purely on-shell supertraces, the former results have a different modular structure which causes them to depend on *off-shell* supertraces as well. In some regions of parameter space, our results demonstrate how certain expected field-theoretic behaviors can emerge from the highly UV/IR-mixed environment. In other situations, by contrast, our results give rise to a number of intrinsically stringy behaviors that transcend what might be expected within an effective field theory approach.

DOI: [10.1103/PhysRevD.107.126019](https://doi.org/10.1103/PhysRevD.107.126019)

### I. INTRODUCTION AND MOTIVATION

String theory is widely regarded as providing the ultimate “UV completion” of theories which successfully describe experimental phenomena at lower energy scales. Such theories include the Standard Model as well as its various extensions. However, it is not always clear how one might draw an explicit map between these full string theories on the one hand and observable low-energy phenomena on the other. Because the fundamental scale of string theory is normally considered to be unreachably remote, and because the particle spectrum of the string is generally quantized in units of this scale, one traditionally attempts to extract low-energy phenomenological

predictions from string theory by focusing on the effects associated with only the lightest of the string modes.

Unfortunately, this approach towards string phenomenology robs us of the full power of string theory to provide new insights into low-energy phenomena. String theory, as a theory of extended objects, does not merely produce light states—it also gives rise to infinite towers of massive states which are also an intrinsic part of the string spectrum. Indeed, the “stringiness” of string theory—i.e., the fundamental features of string theory that transcend our field-theoretic expectations and therefore have the power to suggest new solutions to old puzzles—lies within these states. By disregarding these states and their accumulated contributions to low-energy physics, we are severing the link between the UV-complete theory and its low-energy phenomenology. This reduces us to working within an effective field theory (EFT) whose relevant operators are very hard to explain.

For this reason, it may be argued that a proper approach to understanding many of the low-energy phenomenological implications of string theory is one in which these infinite towers of states are retained and their effects are incorporated in a natural way throughout our calculations.

\*s.a.abel@durham.ac.uk

†dienes@arizona.edu

‡luca.a.nutricati@durham.ac.uk

*Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.*

Indeed, the effects of such states are likely to be the most relevant for fundamental phenomenological questions—such as *hierarchy* problems—which focus on the difficulties of maintaining a peaceful coexistence of both light and heavy scales within a quantum-mechanical universe.

One clue as to the power of these infinite towers of states is that string theories generally have finiteness properties that transcend what can be expected in field theory. One normally attributes these finiteness properties to the extended nature of the string—a feature lacking in theories based on point particles—but this extended nature of the string is precisely what gives rise to these infinite towers of states. For perturbative closed strings (which will be our main focus throughout this paper), worldsheet modular invariance is the exact fundamental symmetry which governs these states and their interactions. Thus, modular invariance holds the key to much of the stringiness of string theory and the finiteness (or softened divergences) associated with its low-energy phenomenological predictions. However, modular invariance also leads to much more, including a unique and surprising form of UV/IR mixing that can severely distort the validity of effective field theories, even at low energies where one might have assumed EFT-based approaches to hold.

For this reason, it is important to develop fully modular-invariant methods of extracting low-energy phenomenological predictions from string theory. By their very nature, these are methods in which the full towers of string states play an important role and cannot be neglected. It is then hoped that the inclusion of these infinite towers of states and the preservation of the underlying modular symmetry can lead to new ways of approaching long-standing phenomenological puzzles. Indeed, as originally advocated in Ref. [1], this might be one route towards developing nontraditional approaches towards addressing hierarchy problems.

In this paper, we shall calculate the running of the one-loop gauge couplings within string theory. This is an old and classic topic within string phenomenology, but we shall employ a formalism for doing this calculation which fully respects modular invariance and which thereby incorporates all of the “magic” to which string theory gives rise. We shall begin in Sec. II by reviewing the framework [2] within which we shall perform this calculation. We shall also summarize the prior results in this field and highlight the ways in which our approach (and our eventual results) will be different. Section III then forms the main body of this paper. Within this section, we shall systematically perform our calculations, ultimately developing a completely general picture of how gauge couplings run within four-dimensional closed string theories. Along the way we shall also discuss several new results which may have wider applicability beyond our specific gauge-coupling calculation. These include new theorems concerning the cancellations of various supertraces of modular-invariant operators. We shall also discuss the effects of *entwinement*,

a phenomenon which emerges within the context of our gauge-coupling calculation and which shifts the meaning of “physicality” when characterizing different states in the string spectrum. We shall then summarize our main results and possible directions for future research in Sec. IV.

## II. PRELIMINARIES: OUR FRAMEWORK AND CONNECTION TO PRIOR LITERATURE

In Ref. [2], a framework was developed for performing calculations of the Higgs mass in a fully modular-invariant way. As discussed there, this framework is completely general and can be applied to any string model (vacuum state). Moreover, although the focus within Ref. [2] centered around calculations of the Higgs mass, this framework can be applied to numerous quantities of phenomenological interest, including the running of the gauge couplings. Explicitly performing such a calculation is thus the primary goal of this paper.

In this section, we shall begin by reviewing the salient features of this framework and the various steps that are involved. With these steps explicitly elucidated, we shall then discuss prior calculations of the running gauge couplings that exist in the literature—including the classic calculation of Kaplunovsky [3]—and discuss precisely which parts of those prior calculations preserve modular invariance and which parts do not. We shall then outline the primary goals of this paper within this language.

### A. Our analysis framework

Within the framework developed in Ref. [2], the calculation of a given low-energy quantity  $\zeta$  proceeds through a number of distinct steps. These steps are illustrated schematically in Fig. 1. For clarity we shall now enumerate these steps individually although many of them are deeply connected to each other and may be performed simultaneously. Explicit examples of each step will be given later in this section.

- (1) As a starting point, one constructs what may be considered to be the “string-theoretic” equivalent of the one-loop field-theoretic contributions to the relevant quantity  $\zeta$  coming from each of the string states. In doing this, one must sum over the contributions from the infinite towers of string states, regardless of their masses. This is a sum of the contributions from the entire tower of states as they propagate around the worldsheet torus, with these contributions weighted appropriately by the naive vertex factors corresponding to their charges and couplings. However, even though we have summed over the entire string spectrum, the resulting expression may not be modular invariant.
- (2) Second, if needed, one then performs a “modular completion” of the above expression for  $\zeta$ . This will generally require the introduction of additional

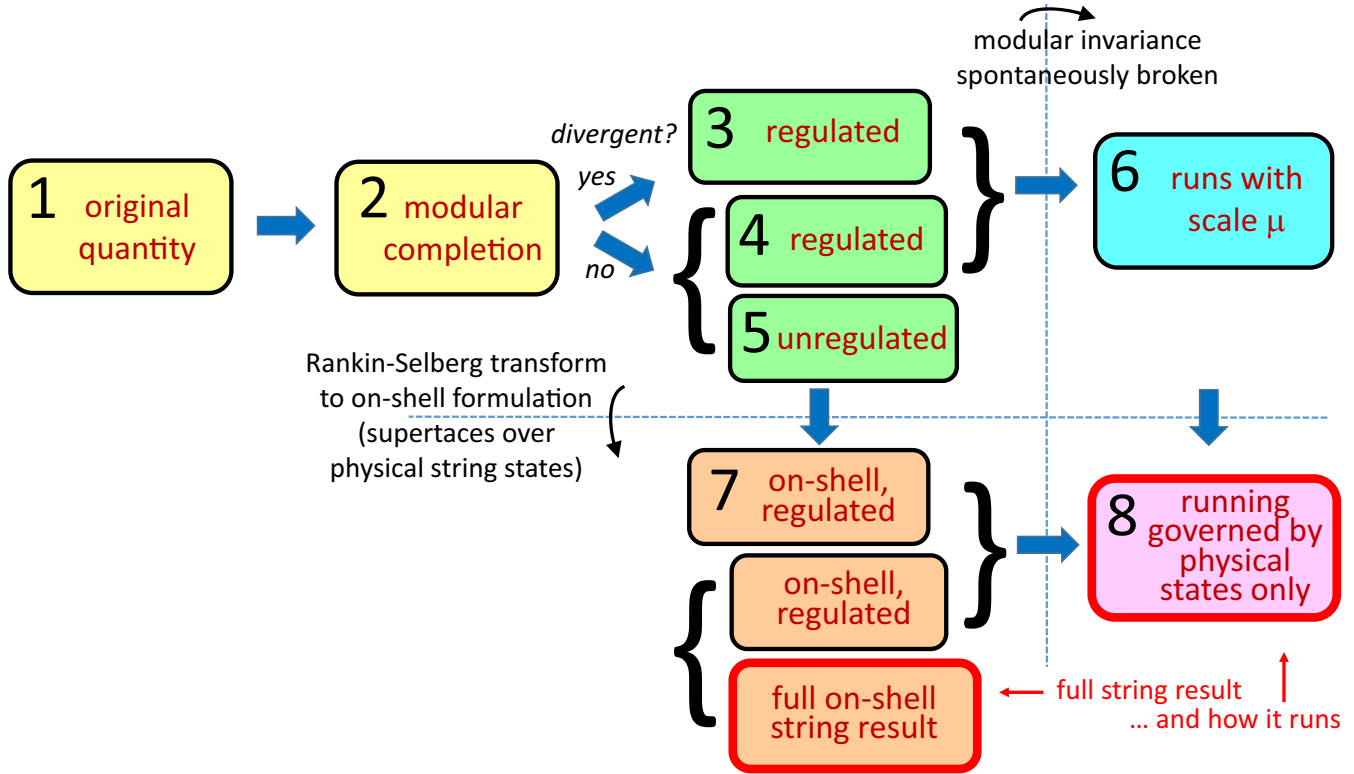


FIG. 1. Distinct steps associated with our analysis of an arbitrary physical quantity, as discussed in the text. The particular sequence of steps to be followed depends on the goal of the analysis. Particularly relevant results are those in the lower portions of this sketch, in which our physical quantity is expressed purely in terms of contributions from on-shell, physical string states and also in a form which runs as a function of a spacetime mass scale  $\mu$ .

terms which may be interpreted as coming from extra intrinsically stringy effects such as gravitational backreactions. In such situations, the tight constraints of modular invariance render these modular completions fairly unique. Thus, after this step, one has obtained a general, fully modular-invariant, string-theoretic expression for the quantity  $\zeta$  under study.

One could, in principle, stop here. However, one natural question that arises is whether this expression for  $\zeta$  is finite, or whether it might diverge in certain string backgrounds. Because this quantity is fully modular invariant, this expression will already exhibit the elimination or softening of the divergences that would have otherwise been expected in field theory. Thus, the divergence structure of  $\zeta$  might be very different from what one would expect in ordinary quantum field theory.

We then have different options, depending on whether  $\zeta$  is finite or divergent.

- (3) If the quantity resulting from Step 2 is potentially divergent, one must regulate this quantity *in a manner which is consistent with modular invariance*. (Indeed, any regulator which breaks modular invariance is likely to introduce precisely the sorts of spurious effects we are hoping to avoid.) This passage from the divergent quantity  $\zeta$  to the regulated

quantity  $\hat{\zeta}$  is indicated as Step 3 within Fig. 1. Thus, after this step, one has a fully string-theoretic (and hence modular-invariant) expression for  $\zeta$  which is also *finite*. We shall let  $\hat{\zeta}$  denote this finite, regulated quantity. In general,  $\hat{\zeta}$  will depend on a regulator parameter  $a$  or collections of parameters  $\{a_i\}$ . The resulting regulated quantity  $\hat{\zeta}(a_i)$  will then be finite for all  $\{a_i\}$  *except* those limiting values of  $a_i$  which correspond to removing the regulator.

- (4) Alternatively, if the quantity resulting from Step 2 is finite, we have two possibilities. One possibility is to nevertheless choose to regulate this quantity in the same way as in Step 3. This then deforms  $\zeta$  into another finite quantity  $\hat{\zeta}(a_i)$  which remains finite even in the limit when the regulator is removed. This is shown as Step 4 within Fig. 1.
- (5) Finally, if the quantity resulting from Step 2 is finite, an alternative possibility is to simply recognize that no regulation is needed. In that case, one can forego the regulator entirely and simply retain the expression obtained in Step 2. We indicate this choice as Step 5 within Fig. 1.

At this stage (green boxes in Fig. 1), we have a quantity  $\hat{\zeta}$  which is fully modular invariant and finite. This quantity will depend on regulator parameters  $a_i$  if we have

employed Steps 3 or 4, but will be independent of  $a_i$  if we have followed Step 5.

There are now several different options for how one might proceed. These different paths ultimately correspond to recasting the finite expression  $\hat{\zeta}$  obtained in Steps 3 through 5 in different forms that are useful for different purposes.

- (6) If we are interested in extracting an EFT-like “running” for  $\hat{\zeta}(a_i)$ , we can start from Step 3 or Step 4 and proceed to identify an appropriate combination  $f(a_i)$  of the  $a_i$ -parameters with a spacetime scale  $\mu$ . As discussed in Ref. [2], such an identification breaks modular invariance by adopting a particular EFT-like direction for spacetime “UV” versus “IR” physics (i.e., a particular UV/IR direction for  $\mu$ ) in terms of the otherwise UV/IR-blind worldsheet combination  $f(a_i)$ . This step nevertheless respects all other aspects of the modular symmetry, and can be viewed as merely breaking the modular symmetry *spontaneously*. One then obtains a running quantity  $\hat{\zeta}(\mu)$ . Indeed, this is the step at which we first introduce the notion of a spacetime energy scale into the theory.

All expressions up to this point receive explicit contributions from the full towers of string states. These include not only physical, “on-shell” level-matched states (whose left- and right-moving mass contributions are equal), but also unphysical “off-shell” states (whose left- and right-moving mass contributions are unequal). Note that the off-shell states can only appear within loops, and thus cannot serve as in-states or out-states in any string amplitude. Indeed, it is the on-shell states which have field-theory analog, while the off-shell states are intrinsically stringy. Thus, if our goal for comparison purposes is to recast our string results into an on-shell form which is as close as possible to what might arise in field theory, we would like to rewrite  $\hat{\zeta}$  in terms of the contributions from only the physical, on-shell states as fully as possible.

To do this, we can utilize certain methods derived from modular-function theory which involve the so-called “Rankin-Selberg” transform [4–6]. The mathematics behind this transform is reviewed in Ref. [2] and ultimately allows us to express a one-loop *string-theoretic* amplitude as the *residue* of a deformed *field-theoretic* amplitude, evaluated at a location in the complex plane associated with the deformation parameter where the field-theoretic amplitude has a pole. This relation between a string amplitude and a (deformed) field-theory amplitude then enables us to obtain an expression for the string amplitude which involves supertraces over the contributions from only the physical string states.

- (7) If we perform a Rankin-Selberg transform starting from the results of Step 3, 4, or 5 (green boxes in Fig. 1), we then obtain corresponding results (orange boxes in Fig. 1) which involve supertraces over only

the physical string states. Such results preserve modular invariance fully and represent an alternative—and often more transparent—formulation for  $\hat{\zeta}$  which enables a direct comparison with what might have been expected in field theory. In particular, if we apply the Rankin-Selberg transform to the results of Steps 3 or 4, we obtain results which also depend on our regulator parameters  $a_i$ . However, if we apply the Rankin-Selberg transform to the results of Step 5, our result depends on the physical supertraces only and does not involve any regulator parameters (orange box with red border in Fig. 1). *This may then be viewed as our final result for the string quantity  $\hat{\zeta}$ —one which is fully modular invariant and involves only the supertraces over physical string states.*

- (8) Alternatively, if we calculate the Rankin-Selberg transform of the results of Step 6 (blue box in Fig. 1)—or equivalently identify  $f(a_i)$  with  $\mu$  within the results of Step 7 (upper two orange boxes in Fig. 1)—we obtain an expression for  $\hat{\zeta}(\mu)$  in which the supertraces over the physical string states govern the running of  $\hat{\zeta}(\mu)$ . This is indicated by the purple box with the red border in Fig. 1. As discussed in Ref. [2], these results preserve modular invariance as fully as possible and yet resemble as closely as possible the running of physical quantities in field theory. *This result thus describes  $\hat{\zeta}$  as a running quantity, where the running is now governed purely by the supertraces of the physical string states.* This formulation for  $\hat{\zeta}(\mu)$  is particularly useful for studying the maximal extent to which an EFT description of  $\hat{\zeta}(\mu)$  at low energies emerges and remains valid within the full modular-invariant string theory.

The most important results of this analysis are those which are indicated in the red-bordered boxes in the lower right portion of Fig. 1. As discussed above, these results respectively express our original quantity in terms of the supertraces over only the physical states in the string spectrum, and also describe how this quantity runs as a function of a spacetime energy scale  $\mu$ . Indeed, the limit in which the regulator is removed will typically correspond to taking the deep-IR limit  $\mu \rightarrow 0$  (or equivalently the deep-UV limit  $\mu \rightarrow \infty$ , given that modular invariance requires an invariance under the scale duality  $\mu \rightarrow M_s^2/\mu$ , as originally pointed out in Ref. [2]). In this limit, the result of Step 8 reduces to the regulator-independent result of Step 7.

Even though we have broken this analysis procedure into distinct steps, we stress that many of these steps are deeply connected and can be performed simultaneously. For example, as discussed in Ref. [2], it is possible to proceed directly from the results of Step 2 to those of Step 7 through a so-called “*regulated* Rankin-Selberg” transform. Likewise, if we are not interested in interpreting our

physical quantities as “running” with respect to a spacetime scale  $\mu$ , we need never be concerned with Step 6 or Step 8.

### B. Prior literature: Results to date

To date, this procedure has been applied to two different quantities  $\zeta$  of phenomenological interest: the one-loop cosmological constant  $\Lambda$ , and the one-loop Higgs mass  $m_\phi^2$ . Here the Higgs field  $\phi$  is identified as any scalar field  $\phi$  whose fluctuations can affect the masses of other string states throughout the string spectrum. We shall now present some of the main results of these prior analyses. These results will not only serve to illustrate the different steps of this procedure but will also be relevant later in this paper.

For the one-loop closed-string cosmological constant (vacuum energy)  $\Lambda$ , Step 1 requires that we begin with the standard expression [7] which is nothing other than the one-loop string partition function  $\mathcal{Z}(\tau)$  integrated over the fundamental domain  $\mathcal{F}$  of the modular group. Indeed, if we define the standard four-dimensional one-loop string amplitude for any operator insertion  $A$  as

$$\langle A \rangle \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} \tau_2^{-1} \sum_{m,n} (-1)^F A_{mn} \bar{q}^m q^n, \quad (2.1)$$

then the corresponding one-loop vacuum energy  $\Lambda$  is nothing but

$$\Lambda = -\frac{\mathcal{M}^4}{2} \langle \mathbf{1} \rangle. \quad (2.2)$$

Here  $\tau$  is the one-loop torus modular parameter with real and imaginary parts  $\tau_{1,2}$  respectively;  $q \equiv e^{2\pi i\tau}$ ;  $\mathcal{F}$  is the fundamental domain of the modular group; the sum  $\sum_{m,n}$  is over all discrete string states with right- and left-moving worldsheet energies  $(m, n)$ , normalized so that the corresponding spacetime mass  $M^2$  is given by  $M^2 = \frac{1}{2}(M_R^2 + M_L^2) = \frac{2}{\alpha'}(m + n)$  where the string scale  $M_s$  and reduced string scale  $\mathcal{M}$  are given by  $M_s \equiv 2\pi\mathcal{M} = 1/\sqrt{\alpha'}$ ;  $F$  is the spacetime fermion number of each state contributing in the sum; and  $A_{mn}$  are the eigenvalues of the operator  $A$  when acting on each  $(m, n)$  string state. Note that  $d^2\tau/\tau_2^2$  is the modular-invariant measure for the  $\tau$ -integration, while the extra prefactor  $\tau_2^{-1}$  within the integrand of Eq. (2.1) emerges from the integration over the continuum of modes associated with the uncompactified spacetime coordinates and reflects the fact that the four-dimensional string partition function, prior to insertions, has modular weight  $k = -1$ . The curved shape of the lower portion of the fundamental domain  $\mathcal{F}$  implies that the amplitude in Eq. (2.1) receives contributions from not only the physical (level-matched) “on-shell” string states with  $m = n$  but also the unphysical (intrinsically stringy) “off-shell” states with  $m \neq n$ . Indeed, we see from Eq. (2.2) that for the cosmological constant  $\Lambda$  the only “insertion” into the partition-

function in Eq. (2.1) is given by  $A = \mathbf{1}$ , the identity operator. This makes sense for a vacuum energy, since all states contribute equally and independently of their possible charges or other characteristics.

The result in Eq. (2.2) thus represents Step 1. Given that  $A = \mathbf{1}$ , this expression is fully modular invariant and no modular completion is needed. This result then carries over to Step 2. Proceeding to Step 3, we ask whether this quantity is divergent. In principle, there are indeed certain states within the string spectrum which could cause divergences: these are physical tachyons for which  $m = n < 0$ . Since the presence of such tachyons destabilizes the theory, we shall restrict our attention to string theories in which such states are absent. It then follows that  $\Lambda$  is finite. According to the procedure we have sketched in Fig. 1, we then have two options which amount to whether or not we wish to impose a regulator. For considerations of  $\Lambda$  alone, there is no need to do so, since  $\Lambda$  is already finite. We shall therefore carry this expression for  $\Lambda$  unchanged into Step 5.

Our final step (Step 7 within Fig. 1) is to evaluate the Rankin-Selberg transform of the expression in Eq. (2.2). This is not difficult, and leads immediately to a result first derived in Ref. [8]:

$$\Lambda = \frac{1}{24} \mathcal{M}^2 \text{Str } M^2 \quad (2.3)$$

where our supertrace ‘Str’ notation indicates a statistics-weighted trace over the spectrum of only physical string states [8]:

$$\text{Str } A \equiv \lim_{y \rightarrow 0} \sum_{\text{states } i} (-1)^{F_i} A_i e^{-y\alpha' M_i^2} \quad (2.4)$$

with the index  $i$  labeling the different physical states in the spectrum. This definition of the supertrace will be discussed further in Sec. III. Intimately connected with the result in Eq. (2.3) and emerging from the same analysis is also an additional constraint [8]

$$\text{Str } \mathbf{1} = 0. \quad (2.5)$$

The results in Eqs. (2.3) and (2.5) hold for any tachyon-free closed string theory in four spacetime dimensions and even generalize [8] to other dimensionalities as well, with  $\text{Str } M^{2\beta} = 0$  for all  $0 \leq \beta \leq \frac{1}{2}(D - 4)$  in general and with  $\Lambda_D \sim \mathcal{M}^2 \text{Str } M^{D-2}$  where  $\Lambda_D$  is the corresponding one-loop cosmological constant in  $D$  spacetime dimensions.

These results are truly remarkable. In ordinary four-dimensional quantum field theory, we would expect that  $\Lambda$  would be a divergent quantity for which  $\text{Str } \mathbf{1}$  governs the quartic divergence and  $\text{Str } M^2$  governs the quadratic divergence. However, we now see that in a four-dimensional tachyon-free modular-invariant string theory  $\Lambda$  is actually *finite* and moreover that  $\text{Str } M^2$  gives its *value*. Likewise,  $\text{Str } \mathbf{1}$  actually *vanishes*.

These results are the consequence of a governing “misaligned supersymmetry” [8,9] which has been proven to exist within the spectra of all tachyon-free modular-invariant string theories. Indeed, this symmetry indicates that bosonic and fermionic string states must be distributed across the infinite string spectrum in such a way that the spectrum is either exactly spacetime supersymmetric (a “degenerate” form of misaligned supersymmetry) or configured in a precise mathematical way wherein any surplus of bosonic states at a given mass level triggers the existence of an even greater surplus of fermionic states at an even higher mass level, which in turn triggers the existence of an even larger surplus of bosonic states at an even higher mass level, and so forth. The sizes of these alternating surpluses grow exponentially as a function of mass, thereby explaining how even a nonsupersymmetric string can remain consistent not only with the Hagedorn transition but also with finite supertrace results such as that in Eq. (2.3). Misaligned supersymmetry thus lies at the heart of the remarkable finiteness properties of closed strings [1,8,9] and will ultimately underpin the results of this paper as well.

To date, the only other physical quantity which has been studied within the full framework sketched in Fig. 1 is the Higgs mass  $m_\phi^2$ . This analysis was performed in Ref. [2], and we shall outline the salient results here. In general, as stated above, the Higgs will be viewed as any state whose VEV affects the masses of at least some of the corresponding string states. We shall work within the Higgsed phase of the theory and accordingly assume that the Higgs field has a nonzero VEV and is already sitting at the minimum of its potential. Clearly the Higgs mass then corresponds to the curvature of this potential at that minimum. In order to calculate this curvature, we can regard the masses  $M^2$  of all string states in the Higgsed phase to be functions of  $\phi$ , where  $\phi$  parametrizes the fluctuations of the Higgs field around this minimum. In complete analogy to Eq. (2.2), it then turns out that [2] the one-loop Higgs mass can then be written as

$$m_\phi^2 = -\frac{\mathcal{M}^2}{2} \langle \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \rangle + \dots \quad (2.6)$$

where the insertions  $\mathbb{X}_1$  and  $\mathbb{X}_2$  into the partition function sum are given by

$$\begin{aligned} \mathbb{X}_1 &\equiv -\frac{1}{4\pi} \partial_\phi^2 M^2 \Big|_{\phi=0}, \\ \mathbb{X}_2 &\equiv \frac{1}{16\pi^2 \mathcal{M}^2} (\partial_\phi M^2)^2 \Big|_{\phi=0}. \end{aligned} \quad (2.7)$$

These insertions thus tally the effective Higgs “charges” (or equivalently the contributions to the curvature of the effective Higgs potential) from each state. Indeed, these are the strengths with which each state couples to the Higgs, as measured by the degree to which the mass of the state responds to fluctuations of the Higgs VEV. The result in

Eq. (2.6) can thus serve as the starting point (Step 1) for our analysis.

As shown in Ref. [2], the insertion of these nontrivial  $\mathbb{X}_i$  operators breaks the modular invariance due to a subtle modular anomaly. As a result, a modular completion is needed. It turns out [2] that the appropriate completion in this case can be achieved by introducing an additional constant into the operator insertions, so that Eq. (2.6) now takes the fully modular-invariant (completed) form

$$m_\phi^2 = -\frac{\mathcal{M}^2}{2} \left\langle \frac{\xi}{4\pi^2} + \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \right\rangle \quad (2.8)$$

where  $\xi$  is an  $\mathcal{O}(1)$  parameter which describes the way in which the particular Higgs field under study is embedded within the corresponding string model. Indeed, this extra term can be interpreted as arising from the universal gravitational backreactions associated with the direct Higgs couplings to the individual string states. Equation (2.8) is therefore fully modular invariant and serves as the result of Step 2. Note that use of our result in Eq. (2.2) enables us to rewrite Eq. (2.8) in the form

$$m_\phi^2 = \frac{\xi}{4\pi^2} \frac{\Lambda}{\mathcal{M}^2} - \frac{\mathcal{M}^2}{2} \langle \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \rangle, \quad (2.9)$$

thereby indicating the existence of a surprising string-theoretic connection between the Higgs mass and the cosmological constant [2]. It is intriguing that such relations join together precisely the two quantities whose values lie at the heart of the two most pressing hierarchy problems in modern physics.

In general, the quantity in Eq. (2.9) can diverge at most logarithmically. This is also a striking result, indicating that modular invariance has significantly softened what would otherwise have been a field-theoretic *quadratic* divergence of the Higgs mass. Moreover, we see that this quantity is actually *finite* unless the string model in question happens to contain a net number of massless  $\mathbb{X}_2$ -charged string states. For simplicity, we shall therefore proceed under the assumption that the net number of massless  $\mathbb{X}_2$ -charged string states vanishes in the string model under discussion, and merely note that the analysis presented in Ref. [2] is completely general and considers all possible cases, including those in which the net number of such states is nonzero.

Given these assumptions, we can now continue to express these results in different forms. One possibility is to proceed directly through Step 5 towards Step 7 by taking the Rankin-Selberg transform of our modular-complete result in Eq. (2.9). In this way, one finds that the Higgs mass can generally be expressed in terms of the contributions from only the physical string states [2]:

$$m_\phi^2 = \frac{1}{24} \mathcal{M}^2 \text{Str}[D_\phi^2 M^2(\phi)] \Big|_{\phi=0} \quad (2.10)$$

where we have defined the modular-covariant double- $\phi$  derivative

$$D_\phi^2 \equiv \partial_\phi^2 + \frac{\xi}{4\pi^2 \mathcal{M}^2}. \quad (2.11)$$

The result in Eq. (2.10) is thus the Higgs-mass analog of the  $\Lambda$ -result in Eq. (2.3).

Another possibility is to analyze how our string-theoretic Higgs mass runs as a function of a spacetime mass scale  $\mu$ . For this purpose we start from the result in Eq. (2.9) and proceed towards Step 4 by introducing a suitable regulator. As discussed in Ref. [2], there are many requirements on such regulators, chief among them that they be completely modular invariant. One compelling class of such regulators can be formulated by deforming our one-loop amplitudes

$$\langle A \rangle \rightarrow \langle A \rangle_{\hat{\mathcal{G}}} \quad (2.12)$$

where  $\langle A \rangle_{\hat{\mathcal{G}}}$  is defined exactly as in Eq. (2.1) except that the integrand is now multiplied by an appropriate modular-invariant regulator function  $\hat{\mathcal{G}}(a_i, \tau)$ , with  $a_i$  denoting the internal regulator parameters. We then must demand that  $\hat{\mathcal{G}}(a_i, \tau)$  exhibit certain properties in order to ensure that we have a sensible regulator. In particular, for such a regulator, we demand that there exist a combination or function  $f(a_i)$  of regulator parameters such that taking  $f(a_i) \rightarrow 0$  effectively removes the regulator while taking any nonzero value of  $f(a_i)$  allows the regulator to suppress the unwanted divergences but otherwise leave the theory intact as far as possible. Given that all such divergences must come from those portions of the integration region in which  $\tau \rightarrow \tau_{\text{cusp}}$  (where  $\tau_{\text{cusp}}$  are the so-called ‘‘cusp’’ points  $\tau_{\text{cusp}} = i\infty$  or  $\tau_{\text{cusp}} = p/q$ , where  $p, q \in \mathbb{Z}$ ), we thus have three requirements for suitable modular-invariant regulator functions  $\hat{\mathcal{G}}(a_i, \tau)$ :

- (i) For all  $f(a_i) > 0$ , we require that  $\hat{\mathcal{G}}(a_i, \tau) \rightarrow 0$  sufficiently rapidly as  $\tau \rightarrow \tau_{\text{cusp}}$ . This enables our regulator to suppress divergences and yield a finite one-loop string amplitude.
- (ii) For all  $f(a_i) > 0$ , we also require that  $\hat{\mathcal{G}}(a_i, \tau) \approx 1$  when  $\tau$  is sufficiently far away from the cusp points. This ensures that our regulator, while suppressing divergences near the cusp points, leaves the remainder of the theory intact as much as possible.
- (iii) Finally, as  $f(a_i) \rightarrow 0$ , we require that  $\hat{\mathcal{G}}(a_i, \tau) \rightarrow 1$  for all  $\tau$ . This ensures the existence of a limit in which our regulator is effectively removed and our original theory is obtained.

We shall also need to require for consistency that  $\hat{\mathcal{G}}(\tau)$  satisfy an additional algebraic identity [2] whose significance will be discussed shortly.

In Ref. [2], a suitable modular-invariant regulator function  $\hat{\mathcal{G}}(a_i, \tau)$  meeting all of these criteria was developed. This regulator function will be discussed in detail in

Sec. III. However, using this regulator, we can then take Step 4 by evaluating

$$\hat{m}_\phi^2(\rho, a) = -\frac{\mathcal{M}^2}{2} \left\langle \frac{\xi}{4\pi^2} + \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \right\rangle_{\hat{\mathcal{G}}}. \quad (2.13)$$

We then follow Step 6 by mapping to a spacetime mass scale  $\mu$  via the identification [2]

$$\mu^2 = f(a_i) M_s^2, \quad (2.14)$$

after which we follow Step 8 by evaluating the Rankin-Selberg transform. A detailed discussion of the Rankin-Selberg procedure is provided in Ref. [2]. The end result of this analysis yields our final on-shell result for the running Higgs mass, expressed completely in terms of supertraces over only physical string states. Indeed, this result takes the form

$$\hat{m}_\phi^2(\mu) = \hat{m}_\phi^2(\mu)|_{\mathbb{X}} + \frac{\xi}{4\pi^2 \mathcal{M}^2} \hat{\Lambda}(\mu) \quad (2.15)$$

where the two different terms on the right side represent the contributions ultimately stemming from the different terms in Eq. (2.8).

The algebraic forms of these final results [2] are fairly complicated (involving infinite sums of Bessel functions) and thus not particularly illuminating. However, the total result for the running Higgs mass  $\hat{m}_\phi^2(\mu)$  is plotted in Fig. 3 of Ref. [2]. One important feature of this running is a ‘‘scale-duality’’ invariance [2] under  $\mu \rightarrow M_s^2/\mu$ . As discussed in Ref. [2], the emergence of scale duality is a general phenomenon, an unavoidable consequence of modular invariance and its corresponding UV/IR symmetries.

The existence of scale duality nevertheless places an additional constraint on potential regulator functions  $\hat{\mathcal{G}}(a_i, \tau)$ . Specifically, scale-duality symmetry in conjunction with the identification in Eq. (2.14) together require that our regulator function  $\hat{\mathcal{G}}(a_i, \tau)$  also exhibit an invariance under any transformations on the parameters  $a_i$  for which  $f(a_i) \rightarrow 1/f(a_i)$ . Phrased slightly differently, the transformation  $f(a_i) \rightarrow 1/f(a_i)$  must be a symmetry of the regulator. Otherwise, it would not be possible to identify a spacetime mass scale  $\mu$  consistent with scale duality. Thus, while a regular function without this additional symmetry might have been sufficient if our only goal were to tame divergences, this extra symmetry is required if we wish to further identify some combination  $f(a_i)$  of regulator parameters with a spacetime mass scale  $\mu$  and thereby express our results as quantities that run with  $\mu$ .

Given the explicit expressions for  $\hat{m}_\phi^2(\mu)|_{\mathbb{X}}$  and  $\hat{\Lambda}(\mu)$  in Ref. [2], it is possible to verify that  $\lim_{\mu \rightarrow 0} \hat{\Lambda}(\mu) = \Lambda$ , as

expected when the regulator is removed. Moreover, it turns out that

$$\hat{m}_\phi^2(\mu)|_{\times} = \left. \frac{\partial^2}{\partial \phi^2} \hat{\Lambda}(\mu, \phi) \right|_{\phi=0}. \quad (2.16)$$

Indeed, as discussed in Ref. [2], this result holds independently of the choice of regulator function  $\hat{\mathcal{G}}(a_i, \mu)$ . Given Eq. (2.15), we then have

$$\begin{aligned} \hat{m}_\phi^2(\mu) &= \left( \partial_\phi^2 + \frac{\xi}{4\pi^2 \mathcal{M}^2} \right) \hat{\Lambda}(\mu, \phi) \Big|_{\phi=0} \\ &= D_\phi^2 \hat{\Lambda}(\mu, \phi) \Big|_{\phi=0}, \end{aligned} \quad (2.17)$$

whereupon taking the  $\mu \rightarrow 0$  limit we find

$$\begin{aligned} \lim_{\mu \rightarrow 0} \hat{m}_\phi^2(\mu) &= D_\phi^2 \Lambda(\phi) \Big|_{\phi=0} \\ &= \frac{1}{24} \mathcal{M}^2 \text{Str}[D_\phi^2 M^2(\phi)] \Big|_{\phi=0}, \end{aligned} \quad (2.18)$$

thereby matching the result for  $m_\phi^2$  from Step 7 in Eq. (2.10). This matching is an important cross-check, since taking  $\mu \rightarrow 0$  corresponds to the removal of our regulator. Indeed, pushing this further, we see that  $\Lambda$  and  $m_\phi^2$  are related through the algebraic structure [2]

$$\begin{cases} \Lambda = \Lambda(\phi) \Big|_{\phi=0} \\ m_\phi^2 = D_\phi^2 \Lambda(\phi) \Big|_{\phi=0}, \end{cases} \quad (2.19)$$

with this structure remaining intact even if we extend these quantities to run as functions of  $\mu$ . Finally, the second of these relations suggests that we may view  $\Lambda(\phi)$  as a Higgs Coleman-Weinberg potential for  $\phi$  (at least locally). This is discussed further in Ref. [2].

### C. Goals and results of this paper

As reviewed above, the one-loop cosmological constant  $\Lambda$  and one-loop Higgs mass  $m_\phi^2$  have already been analyzed within the formalism we have presented, with the central results outlined above. In this paper, by contrast, our goal is to analyze a third quantity: the one-loop contributions to the gauge couplings  $\alpha_i \equiv g_i^2/(4\pi)$  associated with the various gauge groups that might be present in a given string model.

For the gauge couplings, it turns out that certain steps in the above procedure have already been performed. In a seminal early paper [3], Kaplunovsky considered the so-called ‘‘threshold corrections’’ that are required to match the full string gauge couplings to an EFT at one loop and constructed an expression for such threshold corrections which we may regard as completing Step 1. He recognized that this quantity generally diverges due to the contributions from certain massless states, and provided a procedure for removing this divergence. Unfortunately, although

sufficient for certain purposes, this procedure was not modular invariant. Indeed, we shall see that even the starting point—the notion of a ‘‘threshold correction’’—is not modular invariant, as it artificially separates the contributions of massless states from those of massive states. This will be discussed further in Sec. III.

Later, in an important series of papers [10–12], Kiritsis and Kounnas revisited this issue and developed a properly modular-invariant regulator for this calculation. In so doing, Kiritsis and Kounnas implicitly completed Steps 2 and 3. Indeed, the regulator which we shall employ in this paper (and which was employed in Ref. [2]) is built upon the regulator they constructed. However, our regulator has been generalized and modified in a certain critical way which allows us to proceed to identify a corresponding spacetime mass scale  $\mu$  for all values of the regulator parameters  $a_i$  and thereby express the gauge couplings as running quantities [2]. Specifically, the regulator function used in Refs. [10–12] satisfied the two bulleted requirements above but did not exhibit the required symmetry under  $f(a_i) \rightarrow 1/f(a_i)$  which is critical for properly identifying a running spacetime mass scale  $\mu$ . This will be corrected in our analysis in Sec. III.

More importantly, however, the primary purpose of this paper is to bring this analysis of the gauge couplings to its natural conclusion. In particular, we shall complete the remaining steps in our procedure outlined above, and seek to obtain an expression for the gauge couplings in terms of the supertraces of the contributions from only the physical string states. Interestingly, we shall find that this cannot be done for all terms in our expressions because of the unique modular structure of the gauge couplings. We shall therefore spend considerable time discussing this issue, and we shall develop a procedure through which these contributions can nevertheless be written as supertraces over certain string states. We will also study the running of the gauge couplings as functions of a spacetime mass scale  $\mu$ . This will enable us to determine the properties—and also the limits of validity—for any associated EFT describing the behavior of the gauge couplings in closed string theories. In particular, we will see how the running EFT emerges from our prescription and evolves as various mass thresholds are crossed.

### III. GAUGE COUPLINGS IN STRING THEORY: GENERAL TREATMENT

We now turn to the principal goal of this paper: to utilize the methods outlined above in order to study the behavior of the one-loop contribution to the gauge coupling  $g_G$  corresponding to any spacetime gauge group  $G$  in closed string theory. We shall normalize these couplings such that the corresponding gauge-kinetic terms are given by

$$\mathcal{L} = -\frac{1}{4g_G^2} F_{\mu\nu}^{(G)} F^{(G)\mu\nu}, \quad (3.1)$$



and we shall isolate the one-loop contributions to  $g_G$  by evaluating these couplings  $g_G$  to one-loop order and then separating out the tree-level contributions. In general, these quantities are related through

$$\frac{16\pi^2}{g_G^2} \Big|_{\substack{\text{total thru} \\ \text{one-loop} \\ \text{order}}} = \frac{16\pi^2}{g_G^2} \Big|_{\text{tree}} + \Delta_G \quad (3.2)$$

where  $\Delta_G$  denotes the one-loop contribution to  $16\pi^2/g_G^2$ . Indeed, in string theory we know that  $g_G|_{\text{tree}} \sim e^{-\langle\phi\rangle}$  where  $\langle\phi\rangle$  denotes the VEV of the dilaton  $\phi$ . Our goal in this paper is thus to study the properties of  $\Delta_G$ .

### A. Operator insertions

In field theory, we know that  $\Delta_G$  receives contributions from all of the states in our theory which transform in nontrivial representations  $R$  of  $G$ . Indeed, for each such state in the theory, the corresponding one-loop contribution to  $\Delta_G$  is given by  $b \cdot \text{tr}_R(Q_G^2)$ , where

- (i)  $Q_G^2$  is the sum of the squares of the charges in the Cartan subalgebra of  $G$ ;
- (ii) the trace tallies the values of  $Q_G^2$  over all the states within the representation  $R$  (following the convention that each  $CPT$ -conjugate particle/antiparticle pair of states is counted only once); and
- (iii) the numerical coefficient  $b$  encapsulates the Lorentz helicity properties of the state, with  $b = \{1/3, 2/3, -11/3\}$  for Lorentz scalars, spinors, and vectors respectively.

Indeed, we note that these  $b$ -coefficients are nothing but  $b = -4(-1)^F(S^2 - 1/12)$  where  $S = \{0, 1/2, 1\}$  is the Lorentz spin of the corresponding state and where  $F$  is the spacetime fermion number.

Given these observations, it is straightforward to generate an analogous expression in string theory. Of course, in string theory, our traces count *all* states in the theory independently and thus tally each member of a  $CPT$ -conjugate particle/antiparticle pair separately. With this effective doubling of the conventions for our traces, our field-theoretic  $b$ -coefficients are effectively rescaled to become  $b = \{1/6, 1/3, -11/6\}$  for Lorentz scalars, spinors, and vectors respectively, or equivalently  $b = -2(-1)^F(S^2 - 1/12)$ . At this stage, then, our QFT-motivated expression for  $\Delta_G$  in string theory can be expected to take the form

$$\Delta_G = -2\langle(S^2 - 1/12)Q_G^2\rangle \quad (3.3)$$

where the brackets signify the full one-loop amplitude of the form given in Eq. (2.1). Indeed, we note from Eq. (2.1) that these brackets already include the factor of  $(-1)^F$  as well as the double sum  $\sum_{m,n}$  which effects the sum over gauge-group representations  $R$  and the traces over  $Q_G^2$  within each  $R$ . Furthermore, without loss of generality, the presence of a

gauge symmetry implies that our string states populate a corresponding lattice of gauge charges. We can then decompose

$$Q_G^2 = \sum_{\ell, \ell'} c_{\ell\ell'}^{(G)} Q_\ell Q_{\ell'} \quad (3.4)$$

where the  $Q_\ell$  component is the charge operator in the  $\ell^{\text{th}}$  lattice direction and where the coefficients  $c_{\ell\ell'}^{(G)}$  describe how the string gauge group  $G$  is *embedded* within the charge lattice.

Equation (3.3) thus represents our Step 1 starting point for our study of the one-loop contributions to the gauge couplings. Indeed, we see that this quantity is written in terms of the product of two insertions,  $Q_G^2$  and  $S^2 - 1/12$ , and thus resembles as closely as possible the field-theory result, only expressed in terms of a full one-loop string amplitude. Note that if our theory is spacetime-supersymmetric, then we are free to drop the factor of  $-1/12$ , since the contributions from this term will be proportional to  $\text{Tr}(-1)^F$  for each representation of the gauge group and thus vanish. We shall nevertheless keep this factor for generality.

According to the procedure outlined in Sec. II, our next step is to perform a *modular completion* of this expression. Clearly, there are two separate insertions in play:  $Q_G^2$  and  $S^2 - 1/12$ . We shall discuss each of these in turn, since neither insertion preserves the modular invariance of the full string amplitude.

Let us first discuss the modular completion of  $Q_G^2$ . In general, it was shown in Ref. [2] that the product of any two charge bilinears can be modular completed by substituting

$$Q_\ell Q_{\ell'} \rightarrow Q_\ell Q_{\ell'} - \frac{1}{4\pi\tau_2} \delta_{\ell, \ell'}. \quad (3.5)$$

Given the embedding in Eq. (3.4), we thus find the modular completion of  $Q_G^2$  is given by

$$Q_G^2 \rightarrow Q_G^2 - \frac{\xi}{4\pi\tau_2} \quad (3.6)$$

where  $\xi \equiv \sum_{\ell} c_{\ell\ell}^{(G)}$ . Indeed, with this result, we see that  $\xi$  is ultimately related to the *affine level*  $k_G$  at which the gauge group  $G$  is realized.

We now turn to the modular completion of the helicity factor  $S^2 - 1/12$  in Eq. (3.3). In general, a given string theory gives rise to infinite towers of states with higher and higher spins. However, in the heterotic string, these states can ultimately be organized in terms of the CFT sector from which they arise, where the CFT in question is that associated with the transverse right-moving Lorentz group  $SO(D-2)$ . In the heterotic string, there are only three such sectors: the identity (or scalar) sector, the spinor sector, and the vector sector. The ground states of these sectors have

spins  $S = \{0, 1/2, 1\}$  respectively. Loosely speaking, every other string state can be viewed as a member of one of these sectors in the sense that it can be realized through tensor products of this vacuum state (or one of its CFT descendants) with additional vector representations arising from excitations of the left-moving coordinate bosons. In this way, states with arbitrarily high spins can be generated.

Disregarding the contributions from the purely internal degrees of freedom and the two transverse spacetime-coordinate bosons, the contribution to the total partition function from the states in each of these three sectors takes the form  $\bar{\Theta}/\bar{\eta}$ , where  $\eta$  is the Dedekind eta-function and where  $\Theta$  is given by

$$\begin{aligned} \text{scalar: } \Theta &= \frac{1}{2}(\vartheta_3 + \vartheta_4), \\ \text{spinor: } \Theta &= \frac{1}{2}\vartheta_2, \\ \text{vector: } \Theta &= \frac{1}{2}(\vartheta_3 - \vartheta_4). \end{aligned} \quad (3.7)$$

Here  $\vartheta_i$  are the three Jacobi theta-functions. Indeed, in each of these cases we find that

$$\Theta \sim e^{\pi i \tau S^2} (1 + \dots), \quad (3.8)$$

thereby already suggesting a relationship between  $S$  and a modular derivative.

Given this, we now seek to understand how to incorporate the helicity factor  $S^2 - 1/12$  in a fully modular-invariant way into the sum over string states. A direct string calculation [3] tells us that the proper procedure to generate the helicity part is to modify the total partition function of the string theory in question, replacing

$$\frac{\bar{\Theta}}{\bar{\eta}} \rightarrow \frac{\partial}{\partial \bar{\tau}} \left( \frac{\bar{\Theta}}{\bar{\eta}} \right). \quad (3.9)$$

This is the result of a full string calculation, and thus this replacement does not disturb the modular invariance of the total partition function. In particular, the  $\tau$ -derivative  $d/d\bar{\tau}$  is modular-covariant when acting on a modular-covariant function of modular weight  $k = 0$  such as  $\bar{\Theta}/\bar{\eta}$ . Thus, no further modular completion is required after this replacement is implemented. Or, to phrase this another way, the simple insertion  $S^2 - 1/12$  has been “modular completed” by instead implementing the replacement in Eq. (3.9).

The issue that remains for us, however, is to express the replacement in Eq. (3.9) as an *insertion* into the numerator of the partition-function trace. We wish to do this in order to eventually express our results in terms of (weighted) traces over our original string spectrum. To accomplish this, we observe that

$$\begin{aligned} \frac{\partial}{\partial \bar{\tau}} \left( \frac{\bar{\Theta}}{\bar{\eta}} \right) &= \frac{1}{\bar{\eta}} \frac{\partial \bar{\Theta}}{\partial \bar{\tau}} + \bar{\Theta} \frac{\partial}{\partial \bar{\tau}} \frac{1}{\bar{\eta}} \\ &= \frac{1}{\bar{\eta}} \left[ \frac{\partial \bar{\Theta}}{\partial \bar{\tau}} - \bar{\Theta} \frac{\partial}{\partial \bar{\tau}} \log \bar{\eta} \right] \\ &= \frac{1}{\bar{\eta}} \left[ \frac{\partial}{\partial \bar{\tau}} + \frac{\pi i}{12} \bar{E}_2(\bar{\tau}) \right] \bar{\Theta} \end{aligned} \quad (3.10)$$

where in passing to the final line we have utilized the identity

$$E_2(\tau) = \frac{1}{2\pi i} \frac{\partial}{\partial \tau} \log \eta^{24}(\tau) \quad (3.11)$$

where  $E_2(\tau)$  is the normalized weight-two holomorphic Eisenstein function

$$\begin{aligned} E_2(\tau) &\equiv 1 - 24 \sum_{n=1}^{\infty} \sigma(n) e^{2\pi i n \tau} \\ &= 1 - 24q - 72q^2 - 96q^3 - 168q^4 - \dots \end{aligned} \quad (3.12)$$

with  $\sigma(n) \equiv \sum_{d|n} d$ . We shall find it convenient to simplify this notation slightly by writing  $E_2(\tau) = \sum_{n=0}^{\infty} \chi_n q^n$  where

$$\chi_n = \begin{cases} 1 & n = 0 \\ -24\sigma(n) & n > 0. \end{cases} \quad (3.13)$$

We thus see that the replacement in Eq. (3.9) is tantamount to the insertion of the modular-covariant derivative  $\bar{D}_{\bar{\tau}}$  into that portion of the total partition-function trace corresponding to the spacetime Lorentz group, where

$$D_{\bar{\tau}} \equiv \frac{\partial}{\partial \bar{\tau}} - \frac{i\pi}{12} E_2(\bar{\tau}). \quad (3.14)$$

In this sense  $D_{\bar{\tau}}$  is the operator that represents  $S^2 - 1/12$  in string theory.

As is evident from this discussion, the operator  $d/d\bar{\tau}$  acting purely on  $\bar{\Theta}$  represents the spin  $S^2$ . Indeed, we can identify the spin  $S$  as the “helicity charge”  $\bar{Q}_H$  of the state relative to the spacetime Lorentz symmetry, where the subscript  $H$  can be identified as that right-moving lattice direction  $\ell$  whose trace yields  $\bar{\Theta}$ . We can therefore identify  $\bar{Q}_H^2 = \frac{i}{\pi} \partial/\partial \bar{\tau}$ , allowing us to express our modular completion in the form

$$\bar{Q}_H^2 - \frac{1}{12} \rightarrow \bar{Q}_H^2 - \frac{1}{12} \bar{E}_2(\bar{\tau}). \quad (3.15)$$

At first glance, it might have seemed from Eq. (3.5) that the modular completion of  $\bar{Q}_H^2$  would simply be  $\bar{Q}_H^2 \rightarrow \bar{Q}_H^2 - 1/(4\pi\tau_2)$ , just as occurred for the gauge charges. However, the critical difference here is that we are not seeking the modular completion of  $Q_H^2$ ; we are seeking the

modular completion of  $Q_H^2 - 1/12$ . It is the presence of the extra term  $-1/12$  which induces the subtlety and ultimately requires the Eisenstein function in Eq. (3.14). Although it might have seemed that the extra shift  $-1/12$  is only a pure number and thus should be completely harmless, this neglects the fact that we must preserve modular invariance. While the insertion of  $S^2$  raises the modular weight of the corresponding portion of the partition function by two, the insertion of a pure number such as  $-1/12$  does not affect the modular weight at all. *We thus cannot subtract  $1/12$  directly from  $S^2$  or  $Q_H^2$  in a modular-invariant theory; rather, the  $-1/12$  must first be “modular completed” into a modular function (or in this case, a quasimodular function) of weight two.* As it turns out, a theorem in modular-function theory asserts that there is only one (quasi)modular function of weight  $k = 2$ : this is the Eisenstein function  $E_2(\tau)$ . It is thus natural and expected that the modular completion in Eq. (3.15) would involve the Eisenstein function. Indeed, in this sense we may regard  $E_2/12$  as the properly normalized modular completion of  $1/12$ , with  $E_2/12 = 1/12 + \mathcal{O}(q)$ .

As noted above, the Eisenstein series  $E_2$  (unlike the Eisenstein series  $E_{2k}$  for  $k > 1$ ) is not a strict modular function. Instead,  $E_2$  is only *quasimodular*, transforming under modular transformations as

$$E_2\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^2 E_2(\tau) - \frac{6}{\pi} ic(c\tau + d). \quad (3.16)$$

It is the latter “anomaly” term in this result which spoils the true modular covariance for  $E_2$ . However, this is precisely what is needed because the derivative  $\partial/\partial\tau$  in Eq. (3.14) also fails to be modular invariant in exactly the opposite way. Thus, it is precisely the combination in Eq. (3.14) that yields a fully modular-invariant result.

Given these results, we see that our modular-completed expression for  $\Delta_G$  now takes the form

$$\Delta_G = -2 \left\langle \tau_2^2 \left( \bar{Q}_H^2 - \frac{1}{12} \bar{E}_2 \right) \left( Q_G^2 - \frac{\xi}{4\pi\tau_2} \right) \right\rangle. \quad (3.17)$$

Note that the extra factor of  $\tau_2^2$  that has been inserted into Eq. (3.17) is another element of our modular completion. This reflects the fact that the insertions of the helicity and gauge factors—although preserving modular invariance—also together raise the modular weight of the resulting integrand in Eq. (2.1) by two units (from  $k = -1$  to  $k = +1$ ) for any four-dimensional string theory. Modular invariance then dictates that such an increase in the modular weight of the integrand be accompanied by a corresponding increase in the number of leading  $\tau_2$  prefactors. The result in Eq. (3.17) then completes Step 2 of the procedure outlined in Sec. II.

At this stage, it may be worthwhile to compare with the classic results of Kaplunovsky in Ref. [3]. First, we

emphasize that in this paper we are simply calculating the one-loop contributions to the gauge coupling. In particular, despite the algebraic resemblance of Eq. (3.2) to a renormalization-group equation (RGE) for a running gauge coupling, at this stage we have not introduced any notion of running or scale. Second, this conceptual difference notwithstanding, there is a further critical difference in that the contributions from the massless states were explicitly removed within the calculation of the  $\Delta_G$ -term in Ref. [3]. This was done because a separate field-theoretic logarithmic running (assumed to be contributed from the massless states) was explicitly introduced into the renormalization-group version of Eq. (3.2) in Ref. [3]. This rendered  $\Delta_G$  in Ref. [3] a mere tally of the contributions from only the massive modes. Thus, in this sense, the version of  $\Delta_G$  in Ref. [3] became a mere threshold correction, one which is devoid of its own running.

*By contrast, in this paper  $\Delta_G$  will always represent the full one-loop contribution to the gauge coupling, with the contributions from both massless and massive states included together in a unified way.* Indeed, it is only in such a manner that we can ever hope to preserve modular invariance throughout our calculations. Moreover, once we proceed to introduce a scale dependence into our eventual results and consider how these quantities run, we shall even find that the contributions from the massless string states are not strictly logarithmic, but instead take a more complex form which is dictated by modular invariance and which only reduces to a logarithmic running in a certain EFT-like limit.

Certain aspects of the modular completions we have discussed here also appear in Ref. [3] and in the subsequent work reported in Refs. [10–12]. In particular, our modular completion of  $Q_G^2$  is already implicit in Refs. [10–12] and further discussed/reviewed in Refs. [13,14]. Likewise, the effective “modular completion” of the helicity factor whereby the factor of  $S^2 - 1/12$  is dropped in favor of the replacement in Eq. (3.9) already appears in Ref. [3]. However, our subsequent reformulation of this replacement as a partition-function insertion involving the Eisenstein function  $E_2$ —as given in Eqs. (3.14) and (3.15)—is, as far as we are aware, new and does not appear in the prior literature.

Given our expression in Eq. (3.17), we can now continue along the path outlined in Sec. II. In particular, following Eq. (2.6), we see that our total operator insertion for the gauge couplings is given by

$$\mathcal{X} \equiv -2\tau_2^2 \left( \bar{Q}_H^2 - \frac{1}{12} \bar{E}_2 \right) \left( Q_G^2 - \frac{\xi}{4\pi\tau_2} \right). \quad (3.18)$$

Expanding  $\mathcal{X}$  in leading powers of  $\tau_2$  then yields

$$\mathcal{X} = \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \quad (3.19)$$

where we now identify

$$\begin{aligned}\mathbb{X}_1 &\equiv \frac{\xi}{2\pi} \left( \bar{Q}_H^2 - \frac{\bar{E}_2}{12} \right), \\ \mathbb{X}_2 &\equiv -2 \left( \bar{Q}_H^2 - \frac{\bar{E}_2}{12} \right) \mathcal{Q}_G^2.\end{aligned}\quad (3.20)$$

This division of the total insertion into two separate terms  $\mathbb{X}_1$  and  $\mathbb{X}_2$  is based on their leading powers of  $\tau_2$  and will be important when we discuss how our expressions diverge and what kinds of running these quantities ultimately experience. However, we stress that neither  $\langle \tau_2 \mathbb{X}_1 \rangle$  nor  $\langle \tau_2^2 \mathbb{X}_2 \rangle$  is modular invariant by itself. Indeed, these two terms serve as modular completions of each other, and only their sum in Eq. (3.19) is modular invariant. Phrased slightly differently, the splitting of the total insertion into an  $\mathbb{X}_1$  piece and an  $\mathbb{X}_2$  piece based on their leading powers of  $\tau_2$  is not unique. This nonuniqueness arises because modular transformations (especially the Poisson resummations that often underlie these transformations) can change the apparent powers of leading  $\tau_2$  factors that appear. Thus such resummations have the power to induce a reclassification of various terms as belonging to either  $\mathbb{X}_1$  and  $\mathbb{X}_2$ . However, once a given separation into  $\mathbb{X}_1$  and  $\mathbb{X}_2$  is given, it will be consistent to perform all calculations within the framework of that separation without further Poisson resummations. This issue will be discussed further below.

Our result in Eq. (3.19) tells us that our calculation of the gauge couplings shares the same basic algebraic structure as our calculation of the Higgs mass in Ref. [2]. However, one important difference is the fact that  $\mathbb{X}_{1,2}$  now depend on the worldsheet modular parameters  $\tau_{1,2}$  through the Eisenstein function  $E_2(\tau)$ . In other words, we now have more than simple charge insertions—we also have the insertion of an entire modular function. We will shortly see that this difference will have important ramifications.

## B. Divergences and regulator function

Our next step is to study the potential divergence structure of  $\Delta_G$ . Indeed, just as in the Higgs case, it is possible for  $\Delta_G$  to diverge. For example, any level-matched massless state which carries a nonzero  $\mathbb{X}_2$  charge will induce a divergence in  $\Delta_G$  unless this state is balanced against another similar state of opposite statistics. Indeed, in a rough sense to be clarified shortly, the divergence in  $\Delta_G$  will be proportional to  $\text{Str}_{M=0} \mathbb{X}_2$ . Likewise, we see that this divergence is at most *logarithmic*.

The fact that  $\Delta_G$  formally diverges means that we must introduce a regulator. It is here that we pass to Step 3 within Fig. 1. It might seem natural to proceed by simply subtracting the contributions from the massless states (or more precisely the  $\mathbb{X}_2$ -charged massless states) from  $\Delta_G$ . This is reminiscent of what was done in Ref. [3], but introducing this sort of artificial distinction between

massless and massive states necessarily breaks the modular invariance of  $\Delta_G$ .

Instead, following what was done in Ref. [2] for the Higgs mass, we shall regulate our theory by deforming the one-loop amplitude as described in Eq. (2.12), introducing a new regulator function  $\hat{\mathcal{G}}(\tau, \bar{\tau})$  into the integrand:

$$\begin{aligned}\Delta_G &\rightarrow \hat{\Delta}_G \equiv \langle \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \rangle_{\hat{\mathcal{G}}} \\ &= \langle (\tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2) \hat{\mathcal{G}} \rangle.\end{aligned}\quad (3.21)$$

The issue at hand is thus to choose a suitable  $\hat{\mathcal{G}}$  regulator function.

Below Eq. (2.12) we have listed a number of properties that such a regulator function should exhibit. A function satisfying all of these properties was given in Ref. [2], adapting prior results in Ref. [10], and we shall use this function here as well. This function  $\hat{\mathcal{G}}_\rho(a, \tau)$  has two free regulator parameters  $a_i \equiv \{\rho, a\}$  with  $\rho \in \mathbb{R}^+$  and  $\rho \neq 1$ , and is given by

$$\hat{\mathcal{G}}_\rho(a, \tau) \equiv \frac{1}{1 + \rho a^2} \frac{\rho}{\rho - 1} a^2 \frac{\partial}{\partial a} [Z_{\text{circ}}(\rho a, \tau) - Z_{\text{circ}}(a, \tau)] \quad (3.22)$$

where

$$Z_{\text{circ}}(a, \tau) \equiv \sqrt{\tau_2} \sum_{k, \ell \in \mathbb{Z}} \bar{q}^{(ka - \ell/a)^2/4} q^{(ka + \ell/a)^2/4}. \quad (3.23)$$

Note that  $Z_{\text{circ}}(a, \tau)$  represents the sum over the Kaluza-Klein (KK) and winding modes that would be associated with a bosonic worldsheet field compactified on a circle of radius  $(M_s a)^{-1}$ , with  $k$  and  $\ell$  respectively indexing the KK and winding modes, while the leading factor of  $\sqrt{\tau_2}$  is inserted into Eq. (3.23) in order to ensure that  $Z_{\text{circ}}$  is modular invariant. We stress that for our purposes  $Z_{\text{circ}}$  is merely an ingredient in the definition of our regulator and does not correspond to any actual physical compactification of our theory. It turns out that  $\hat{\mathcal{G}}_\rho(a, \tau) \rightarrow 1$  as  $a \rightarrow 0$ , indicating that taking  $a \rightarrow 0$  removes the regulator. Indeed, for this function we have

$$f(a_i) = \rho a^2. \quad (3.24)$$

It then turns out that all of the numbered requirements below Eq. (2.12) are satisfied, in addition to the requirement that  $\hat{\mathcal{G}}_\rho(a, \tau)$  exhibit an invariance under  $f(a_i) \rightarrow 1/f(a_i)$ , or equivalently under  $a \rightarrow (\rho a)^{-1}$ . Indeed, we will eventually identify our spacetime running scale  $\mu$  according to Eq. (2.14) with  $f(a_i)$  given in Eq. (3.24).

### C. The Rankin-Selberg transform: From amplitudes to supertraces

Given this choice of regulator function, our result for  $\hat{\Delta}_G$  in Eq. (3.21) can then be viewed as representing Steps 3 and/or 4 in Fig. 1. From this point, there are several options. One possibility is to proceed directly to Step 6 by identifying a running mass scale  $\mu$  according to Eq. (2.14), where  $f(a_i)$  is given in Eq. (3.24). However, in order to extract a description of the running of the gauge couplings which is as close as possible to what we might expect from ordinary quantum field theory, we are more interested in performing the Rankin-Selberg transform of our result in Eq. (3.21) in order to pass to Steps 7 and 8.

Operationally, this transform can be performed in a number of different ways. In this section, we shall describe three different approaches to evaluating this transform in order to understand their relative advantages and disadvantages. As we shall see, the first two approaches lead to results which hold only under certain simplifying assumptions. Indeed, we describe these approaches because they connect to our previous calculations in Ref. [2]. Unfortunately, these approaches lack the complete generality that we will ultimately require for some of the later calculations in this paper. For this reason, after describing these two approaches, we shall then proceed to outline our third and preferred approach. As we shall see, this approach will be completely general and lead to results which have not been previously described in the literature.

To begin, let us assume that the partition function of our string theory in  $D$  uncompactified spacetime dimensions takes the form

$$Z(\tau) \equiv \tau_2^k \sum_{m,n} a_{mn} \bar{q}^m q^n \quad (3.25)$$

where  $a_{mn}$  is the net (bosonic minus fermionic) number of string states with right- and left-moving worldsheet energies  $(m, n)$  in the string spectrum and where  $k \equiv 1 - D/2$ . These worldsheet energies are related to the total spacetime mass  $M$  of the corresponding string state via  $M^2 = \frac{1}{2}(M_L^2 + M_R^2)$  where  $m = \alpha' M_R^2/4$  and  $n = \alpha' M_L^2/4$ . Let us further assume that we wish to consider the corresponding one-loop amplitude  $\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \rangle$  where the operator  $\sum_{\ell} \tau_2^{\ell} A^{(\ell)}$  being inserted gives the value  $\sum_{\ell} \tau_2^{\ell} A_{mn}^{(\ell)}$  when acting on a state with worldsheet energies  $(m, n)$ . Our one-loop amplitude is then given by

$$\left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} \sum_{\ell} \tau_2^{k+\ell} \sum_{m,n} a_{mn} A_{mn}^{(\ell)} \bar{q}^m q^n. \quad (3.26)$$

For the sake of this discussion we shall assume that this amplitude is already finite and therefore does not require any regulation. We shall return to this issue later when we discuss what happens when we also insert our regulator.

As long as the amplitude in Eq. (3.26) is finite and modular invariant, the Rankin-Selberg transform [4–6] then tells us that this amplitude may equivalently be expressed as

$$\left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle = \frac{\pi}{3} \text{Res}_{s=1} \int_0^{\infty} d\tau_2 \tau_2^{s-2} g(\tau_2) \quad (3.27)$$

where

$$\begin{aligned} g(\tau_2) &\equiv \sum_{\ell} \tau_2^{k+\ell} \int_{-1/2}^{1/2} d\tau_1 \sum_{m,n} a_{mn} A_{mn}^{(\ell)} \bar{q}^m q^n \\ &= \sum_{\ell} \tau_2^{k+\ell} \sum_n a_{nn} A_{nn}^{(\ell)} e^{-\pi\alpha' M_n^2 \tau_2} \end{aligned} \quad (3.28)$$

where  $\alpha' M_n^2 = 4n$ . Inserting Eq. (3.28) into Eq. (3.27) and exchanging the order of the  $n$ -sum and the  $s$ -integral/residue (an operation whose validity will be discussed below), we obtain

$$\begin{aligned} \left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle &= \frac{\pi}{3} \sum_{\ell} \sum_n a_{nn} A_{nn}^{(\ell)} \\ &\quad \times \text{Res}_{s=1} \int_0^{\infty} d\tau_2 \tau_2^{k+\ell+s-2} e^{-\pi\alpha' M_n^2 \tau_2} \\ &= \frac{\pi}{3} \sum_{\ell} \sum_n a_{nn} A_{nn}^{(\ell)} \\ &\quad \times \text{Res}_{s=1} [(\pi\alpha' M_n^2)^{1-k-\ell-s} \\ &\quad \quad \times \Gamma(k + \ell + s - 1)]. \end{aligned} \quad (3.29)$$

Taking  $k = -1$  (as appropriate for  $D = 4$ ) and evaluating the residue of the Euler  $\Gamma$ -function we thus obtain

$$\begin{aligned} \left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle &= \frac{\pi}{3} \sum_{\ell \leq 1} \frac{(-1)^{\ell+1}}{(1-\ell)!} \\ &\quad \times \sum_n a_{nn} A_{nn}^{(\ell)} (\pi\alpha' M_n^2)^{1-\ell}, \end{aligned} \quad (3.30)$$

or equivalently

$$\left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle = \frac{\pi}{3} \sum_{\ell \leq 1} \frac{(-1)^{\ell+1}}{(1-\ell)!} \text{Str}[A^{(\ell)} (\pi\alpha' M^2)^{1-\ell}] \quad (3.31)$$

where for this purpose we temporarily identify the supertrace simply as  $\text{Str} A \equiv \sum_n a_{nn} A_{nn}$ . We thus see that the Rankin-Selberg transform procedure has allowed us to express our original amplitude  $\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \rangle$  in Eq. (3.26) as a sum of supertraces over only physical string states (i.e., states with  $m = n$ ).

As an example of Eq. (3.31), we can consider the amplitude with no insertions at all. Assuming a tachyon-free theory (so that the amplitude is finite, as required), we obtain

$$\langle \mathbf{1} \rangle = -\frac{1}{12} \text{Str } M^2 / \mathcal{M}^2 \quad (3.32)$$

where  $\mathcal{M} \equiv M_s / (2\pi) = (2\pi\sqrt{\alpha'})^{-1}$ . We thus find [8]

$$\Lambda \equiv -\frac{\mathcal{M}^4}{2} \langle \mathbf{1} \rangle = \frac{1}{24} \mathcal{M}^2 \text{Str } M^2, \quad (3.33)$$

as already noted in Eq. (2.3).

Interestingly, we see that our final expression in Eq. (3.31) receives no apparent contributions from off-shell states (i.e., states with  $m \neq n$ ). Likewise, our final result receives no contributions from  $A^{(\ell)}$  insertions with  $\ell \geq 2$ . Because of these features, it might initially seem that the value of the amplitude  $\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \rangle$  is independent of these states and insertions. However, this is not correct: our result in Eq. (3.31) holds only under the assumption that this amplitude is modular-invariant, and modular invariance certainly requires these off-shell states and the occasional inclusion of  $A^{(\ell)}$  insertions with  $\ell \geq 2$ , as we have seen above. Rather, what we are learning from Eq. (3.31) that modular invariance is so powerful a symmetry that the contributions from the off-shell states and  $A^{(\ell)}$  insertions with  $\ell \geq 2$  are already implicitly determined by—and thus can be written in terms of—the contributions from the on-shell and  $\ell \leq 1$  insertions.

Finally, before proceeding further, we note that our derivation of the result in Eq. (3.31) rests on a number of algebraic manipulations which have important qualifications and implications. One of these assumptions is relevant for Eq. (3.29), in which it was assumed that all  $M_n$  are positive in passing to the second line. However, if massless states are present, one can always imagine deforming our theory to give these states very small nonzero masses. One can then perform the integration in Eq. (3.29) rigorously and remove these masses at the end. As discussed in Sec. IV of Ref. [2], this procedure is valid as long as such massless states do not cause the amplitude under study to diverge. Indeed, we see from the final result in Eq. (3.31) that it is difficult for massless states to make contributions in any case unless  $\ell = 1$ .

However, a more important assumption was made in passing from Eq. (3.28) to Eq. (3.29), where we exchanged the order of the  $n$ -summation and  $s$ -integration/residue extraction. This exchange is valid only if the  $n$ -summation over the spectrum does not itself introduce any new divergences. In general, this will indeed be the case. However, there can be limits of our theories in which the spectrum becomes so dense as to be effectively continuous. In such cases, this procedure breaks down,

additional divergences can emerge from the sums over states, and there can be nonzero contributions from the operator insertions  $A^{(\ell)}$  with  $\ell > 1$ . However, it is easy to understand what is happening in such situations: the theory is becoming effectively higher-dimensional. As a result, in such cases we could equivalently shift to a higher-dimensional description from the start. The value of  $k$  would then drop below  $-1$ , with a corresponding change in the values of  $\ell$  for which Eq. (3.31) would receive explicit contributions. Indeed, in a limit in which our four-dimensional theory becomes effectively six-dimensional, we have  $k = 1 - D/2 \rightarrow -2$ . Following the same derivation as above, we then find that the range of the  $\ell$ -summation in Eq. (3.31) becomes  $\ell \leq 2$ , and we now have explicit contributions from  $A^{(2)}$ . This phenomenon can also be understood in terms of the discussion in the paragraph below Eq. (3.20). As we decompactify our theory and the string spectrum becomes dense, it becomes appropriate to perform a Poisson resummation over the corresponding Kaluza-Klein states (momentum modes). This Poisson resummation introduces extra powers of  $\tau_2$  and thus effectively reshuffles certain terms between  $\mathbb{X}_1$  and  $\mathbb{X}_2$ .

Fortunately, it is possible to reformulate the Rankin-Selberg transform procedure in such a way as to avoid exchanging the order of the  $n$ -sum and the  $s$ -integral/residue after reaching Eq. (3.28). This then produces a more general result which holds even when the string spectra become dense. Indeed, rather than integrate over  $\tau_2$  and then take the residue at  $s = 1$  as in Eq. (3.29), we can instead proceed by recognizing from Eq. (3.27) that the original string amplitude  $\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \rangle$  is nothing but the Mellin transform of  $g(\tau_2) / \tau_2$ . We can therefore write  $g(\tau_2)$  directly as the inverse Mellin transform of this amplitude, thereby ultimately leading to the alternative result [15,16]

$$\left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle = \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} g(\tau_2), \quad (3.34)$$

where  $g(\tau_2)$  continues to be given by Eq. (3.28). This result is equivalent to Eq. (3.27), but has the primary advantage that we can now evaluate  $\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \rangle$  simply by taking the  $\tau_2 \rightarrow 0$  limit of  $g(\tau_2)$  rather than by evaluating the residue of the  $\tau_2$  integral of  $g(\tau_2)$ , as in Eq. (3.29). Indeed, inserting Eq. (3.28) into Eq. (3.34), we find

$$\left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle = \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \sum_{\ell} \tau_2^{k+\ell} \sum_n a_{nn} A_{nn}^{(\ell)} e^{-\pi\alpha' M_n^2 \tau_2}. \quad (3.35)$$

The issue then boils down to how we evaluate the right side of Eq. (3.35). Since we are taking the  $\tau_2 \rightarrow 0$  limit, one natural possibility would be to Taylor-expand

the exponential. Taking  $k = -1$  (as appropriate for a four-dimensional theory) and recognizing that terms with  $r > 1 - \ell$  vanish in the  $\tau_2 \rightarrow 0$  limit, we would then obtain

$$\begin{aligned} & \left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle \\ &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \sum_{\ell} \sum_{r=0}^{1-\ell} \tau_2^{\ell+r-1} \frac{(-1)^r}{r!} \sum_n a_{nn} A_{nn}^{(\ell)} (\pi \alpha' M_n^2)^r \\ &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \sum_{\ell \leq 1} \sum_{r=0}^{1-\ell} \tau_2^{\ell+r-1} \frac{(-1)^r}{r!} \text{Str}[A^{(\ell)} (\pi \alpha' M^2)^r] \quad (3.36) \end{aligned}$$

where the supertrace continues to be defined as below Eq. (3.31) and where in passing to the final line we have recognized that the conditions on the  $r$ -sum have imposed an upper limit on the  $\ell$ -sum. However, for theories in which the original string amplitude  $\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \rangle$  is finite, we know that the right side of Eq. (3.36) cannot diverge as  $\tau_2 \rightarrow 0$ . We thus obtain a set of auxiliary conditions which must hold in all such theories, namely

$$\text{Str}[A^{(\ell)} (\pi \alpha' M^2)^r] = 0 \quad (3.37)$$

for all  $0 \leq r \leq -\ell$  with  $\ell \leq 1$ . As long as these auxiliary conditions are satisfied, we then have

$$\left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle = \frac{\pi}{3} \sum_{\ell \leq 1} \frac{(-1)^{1-\ell}}{(1-\ell)!} \text{Str}[A^{(\ell)} (\pi \alpha' M^2)^{1-\ell}], \quad (3.38)$$

thereby exactly matching the result we obtained in Eq. (3.31). We thus see that the Mellin-transformed result has not only reproduced the result in Eq. (3.31), but has also furnished the explicit extra conditions in Eq. (3.37) which must be satisfied in order for this result to be valid. Indeed, these auxiliary conditions may be viewed as the conditions under which the exchange of the order of  $n$ -summation and  $s$ -integration/residue extraction are valid. For example, in the case of the vacuum amplitude  $\langle \mathbf{1} \rangle$ , these auxiliary conditions reduce to the condition  $\text{Str} \mathbf{1} = 0$ . This condition, which we have already mentioned in Eq. (2.5), is quite remarkable, indicating that even when spacetime supersymmetry is broken in a given string model, the residual misaligned supersymmetry continues to ensure that this supertrace vanishes—even though such theories do not permit any possible pairing of bosonic and fermionic states.

The results in Eqs. (3.37) and (3.38) hold for the vast majority of string theories as long as we are avoiding the edges of moduli space corresponding to decompactification limits. Indeed, we continue to issue this proviso because we have still made a further critical algebraic assumption in this analysis. This occurred when we evaluated the right

side of Eq. (3.35) by Taylor-expanding the exponential and then passing the  $r$ -summation past the  $n$ -summation. While this may be a valid step in many string theories, in this paper we shall need to consider cases in which the insertions  $A^{(\ell)}$  have eigenvalues  $(m, n)$ -eigenvalues which are growing functions of  $m$  and  $n$ , thereby rendering our sums over states badly divergent. In such cases, the exponential suppression is critical for a finite result and it is therefore not possible to Taylor-expand this exponential and consider the different terms in the Taylor expansion separately.

As we have explained above, the fundamental difficulty is that the different  $\ell$ -terms which contribute to  $g(\tau_2)$  have different *apparent* powers of  $\tau_2$ , but in reality these powers of  $\tau_2$  can be exchanged between these different  $\ell$ -terms as the result of algebraic manipulations (such as Poisson resummations) that become appropriate in certain limiting regions of moduli space. For this reason, we should really consider  $g(\tau_2)$  as a single object with its own overall  $\tau_2$ -dependence without attempting to draw a correspondence between this overall  $\tau_2$ -dependence and the different  $\ell$ -terms within  $g(\tau_2)$ .

We can accomplish this by following the approach originally taken in Ref. [8]. This approach has also been generalized in Ref. [17], and in the remainder of this subsection we shall quickly derive the important results from Refs. [8,17] that we shall require in the rest of this paper. In particular, returning to Eq. (3.35), we can begin by isolating our sum over states with all modular-invariant operator insertions included, i.e.,

$$S(\tau_2) \equiv \sum_{\ell} \tau_2^{\ell} \sum_n a_{nn} A_{nn}^{(\ell)} e^{-\pi \alpha' M_n^2 \tau_2}. \quad (3.39)$$

However, let us now assume that  $S$  has an overall  $\tau_2$ -dependence as  $\tau_2 \rightarrow 0$  given by

$$S(\tau_2) \sim \sum_j C_j \tau_2^j \quad (3.40)$$

where the coefficients  $C_j$  are completely arbitrary. Note that our assumption that this sum is finite as  $\tau_2 \rightarrow 0$  allows us to assume that  $j \geq 0$ . Indeed, this sum is finite because we have assumed that our theory is free of physical tachyons; because the degeneracies grow no more rapidly than  $|a_{nn}| \sim e^{\sqrt{n}}$  according to the Hagedorn phenomenon; and because the charge-operator eigenvalues  $A_{nn}^{(\ell)}$  typically grow no faster than a polynomial in  $n$ . By contrast, the exponential suppression goes as  $e^{-n}$  since  $\alpha' M_n^2 \sim n$ . In this context, we remark that we are *not* imposing the condition that  $C_j \neq 0$  only for integer values of  $j$ . There indeed exist examples for which noninteger values of  $j$  contribute within the sum in Eq. (3.40). However, we shall be concerned with the lowest-lying values of  $j$ , and for these we can take  $j \in \mathbb{Z}$ .

Once we assume a  $\tau_2$ -dependence of the form in Eq. (3.40), we can take the  $\tau_2$ -derivative of both sides of Eq. (3.40) to obtain

$$\begin{aligned} & \frac{d}{d\tau_2} \left[ \sum_{\ell} \tau_2^{\ell} \sum_n a_{nn} A_{nn}^{(\ell)} e^{-\pi\alpha' M_n^2 \tau_2} \right] \\ &= \sum_n a_{nn} \frac{d}{d\tau_2} \left[ \sum_{\ell} \tau_2^{\ell} A_{nn}^{(\ell)} \right] e^{-\pi\alpha' M_n^2 \tau_2} \\ & \quad + \sum_{\ell} \tau_2^{\ell} \sum_n a_{nn} A_{nn}^{(\ell)} (\pi\alpha' M_n^2) e^{-\pi\alpha' M_n^2 \tau_2} \\ & \stackrel{\text{set}}{=} \frac{d}{d\tau_2} \left( \sum_j C_j \tau_2^j \right) = \sum_j (j+1) C_{j+1} \tau_2^j. \end{aligned} \quad (3.41)$$

Indeed, these relations hold for  $\tau_2 \ll 1$ . Taking the  $\tau_2 \rightarrow 0$  limits of Eqs. (3.40) and (3.41) then yields explicit expressions for the leading coefficients [8,17]

$$\begin{aligned} C_0 &= \lim_{\tau_2 \rightarrow 0} \left[ \sum_{\ell} \tau_2^{\ell} \sum_n a_{nn} A_{nn}^{(\ell)} e^{-\pi\alpha' M_n^2 \tau_2} \right] \\ C_1 &= \lim_{\tau_2 \rightarrow 0} \left\{ \sum_n a_{nn} \frac{d}{d\tau_2} \left[ \sum_{\ell} \tau_2^{\ell} A_{nn}^{(\ell)} \right] e^{-\pi\alpha' M_n^2 \tau_2} \right\} \\ & \quad - \lim_{\tau_2 \rightarrow 0} \left[ \sum_{\ell} \tau_2^{\ell} \sum_n a_{nn} A_{nn}^{(\ell)} (\pi\alpha' M_n^2) e^{-\pi\alpha' M_n^2 \tau_2} \right]. \end{aligned} \quad (3.42)$$

Likewise, the coefficients  $C_j$  with  $j \geq 2$  can be calculated in a similar fashion by taking additional  $\tau_2$ -derivatives (and will ultimately be useful only for theories in spacetime dimensions  $D > 4$ ).

Substituting Eq. (3.40) into Eq. (3.35) and taking  $k = -1$  as appropriate for four-dimensional string theories then yields

$$\begin{aligned} \left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \tau_2^k \sum_{j=0}^{\infty} C_j \tau_2^j \\ &= \frac{\pi}{3} C_1 \\ &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \left\{ \sum_n a_{nn} \frac{d}{d\tau_2} \left[ \sum_{\ell} \tau_2^{\ell} A_{nn}^{(\ell)} \right] e^{-\pi\alpha' M_n^2 \tau_2} \right\} \\ & \quad - \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \left[ \sum_{\ell} \tau_2^{\ell} \sum_n a_{nn} A_{nn}^{(\ell)} (\pi\alpha' M_n^2) e^{-\pi\alpha' M_n^2 \tau_2} \right]. \end{aligned} \quad (3.43)$$

Critically, we now see that all of our expressions are properly convergent as a result of the exponential damping factors. Indeed, motivated by comparison with our earlier results, we define the *regulated supertrace* [8]

$$\text{Str } A \equiv \lim_{\tau_2 \rightarrow 0} \sum_n a_{nn} A_{nn} e^{-\pi\alpha' M_n^2 \tau_2}. \quad (3.44)$$

Indeed, this is nothing but the supertrace we introduced in Eq. (2.4), with  $y$  now identified as  $\pi\tau_2$ . Moreover, we now

see that Eq. (3.44) serves as the correct formal definition of the supertrace previously defined below Eq. (3.31).

Written in terms of these supertraces we thus have our general result that expresses our full one-loop string amplitude in terms of supertraces over physical string states:

$$\begin{aligned} \left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle &= \frac{\pi}{3} \text{Str} \left[ \frac{d}{d\tau_2} \left( \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right) \right] \\ & \quad - \frac{\pi}{3} \text{Str} \left[ \left( \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right) (\pi\alpha' M^2) \right]. \end{aligned} \quad (3.45)$$

Equivalently, for any modular-invariant operator insertion  $\mathcal{X}$  in four dimensions, we have [17]

$$\begin{aligned} \langle \mathcal{X} \rangle &= \frac{\pi}{3} \text{Str} \left( \frac{d}{d\tau_2} \mathcal{X} \right) - \frac{\pi}{3} \text{Str} [\mathcal{X} (\pi\alpha' M^2)] \\ &= \frac{\pi}{3} \text{Str} (D_{\tau_2} \mathcal{X}) \end{aligned} \quad (3.46)$$

where

$$\begin{aligned} D_{\tau_2} &\equiv \frac{d}{d\tau_2} - \pi\alpha' M^2 \\ &= \frac{d}{d\tau_2} - \frac{1}{4\pi\mathcal{M}^2} M^2. \end{aligned} \quad (3.47)$$

For modular-invariant insertions  $\mathcal{X}$  that are  $\tau_2$ -independent (so that we can write  $\mathcal{X} = \mathbb{X}$  where  $\mathbb{X}$  is itself modular invariant), this result simplifies to

$$\langle \mathbb{X} \rangle = -\frac{1}{12\mathcal{M}^2} \text{Str} (\mathbb{X} M^2). \quad (3.48)$$

Likewise, for operator insertions of the form  $\mathcal{X} = \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2$ , we have

$$\begin{aligned} \langle \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \rangle &= \frac{\pi}{3} \text{Str} (\mathbb{X}_1 + 2\tau_2 \mathbb{X}_2) \\ & \quad - \frac{1}{12\mathcal{M}^2} \text{Str} [(\tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2) M^2]. \end{aligned} \quad (3.49)$$

Moreover, under our assumption that the string amplitude is finite, we learn from the first line of Eq. (3.43) that we must have  $C_0 = 0$  as an auxiliary condition. Given the coefficients in Eq. (3.42), this then yields [17]

$$\text{Str } \mathcal{X} = 0. \quad (3.50)$$

This auxiliary condition thus accompanies our result in Eq. (3.46) or its simplifications in Eqs. (3.48) or (3.49).

Before proceeding further, it is worth emphasizing that the quantity in Eq. (3.44) is properly viewed as a regulated supertrace only if the inserted operator  $A$  is itself



$\tau_2$ -independent. In such cases, the  $\tau_2$ -dependent exponential factor in Eq. (3.44) functions as a true regulator, with  $\tau_2$  functioning as a dummy regulator variable whose regulating effects are ultimately removed by taking the  $\tau_2 \rightarrow 0$  limit. By contrast, in cases in which our inserted operator  $A$  has its own  $\tau_2$ -dependence, we are no longer free to view  $\tau_2$  as an independent regulator variable; rather, the quantity whose limit must be taken in Eq. (3.44) becomes inextricably identified with the  $\tau_2$  that appears within  $A$ , so that the  $\tau_2 \rightarrow 0$  limit not only removes the damping exponential but also deforms the operator  $A$ . However, in either case, the operational prescription is clear: we insert our full operator  $A$  within Eq. (3.44), along with any  $\tau_2$ -dependent factors which may appear, and then evaluate the sum and limit accordingly.

The results in Eqs. (3.48) and (3.50) were originally derived in Ref. [8] for the case  $\mathcal{X} = \mathbb{X} = \mathbf{1}$ , but we now see [17] that they hold for all modular-invariant  $\tau_2$ -independent insertions  $\mathbb{X}$  that lead to finite string amplitudes. Moreover, we now also see that the result in Eq. (3.48) is actually only a special case of the more general result in Eq. (3.46). Indeed, as we have stressed above, our formulation of the Rankin-Selberg transform in Eq. (3.46) is completely general and holds for *any* operator insertion  $\mathcal{X} \equiv \sum_{\ell} \tau_2^{\ell} A^{(\ell)}$  regardless of the values of  $\ell$  involved, so long as the regulated supertrace in Eq. (3.44) is used and the inserted operator  $\mathcal{X}$  does not disturb the modular invariance of the amplitude integrand or the finiteness of the resulting amplitude.

Thus far, we have shown that our desired amplitude  $\langle \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \rangle$  is given in Eq. (3.49). However, there is an alternate way in which we might have evaluated this amplitude [17]. Given the sum  $S$  in Eq. (3.39), we immediately recognize that

$$g(\tau_2) = \tau_2^{-1} S(\tau_2). \quad (3.51)$$

We therefore directly have

$$\begin{aligned} \left\langle \sum_{\ell} \tau_2^{\ell} A^{(\ell)} \right\rangle &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \tau_2^{-1} S(\tau_2) \\ &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \left[ \sum_{\ell} \tau_2^{\ell-1} \sum_n a_{nn} A_{nn}^{(\ell)} e^{-\pi \alpha' M_n^2 \tau_2} \right] \\ &= \frac{\pi}{3} \text{Str} \left( \sum_{\ell} \tau_2^{\ell-1} A^{(\ell)} \right). \end{aligned} \quad (3.52)$$

Indeed, for any modular-invariant operator insertion  $\mathcal{X}$ , this becomes [17]

$$\langle \mathcal{X} \rangle = \frac{\pi}{3} \text{Str}(\tau_2^{-1} \mathcal{X}). \quad (3.53)$$

Comparing this with our result in Eq. (3.46), we thus obtain the remarkable identity [17]

$$\text{Str}(\tau_2^{-1} \mathcal{X}) = \text{Str}(D_{\tau_2} \mathcal{X}) \quad (3.54)$$

where the derivative  $D_{\tau_2}$  is given in Eq. (3.47). This identity applies to any modular-invariant operator insertion  $\mathcal{X}$  in four dimensions.

Note that this identity does *not* imply that  $\tau_2^{-1} \mathcal{X}$  and  $D_{\tau_2} \mathcal{X}$  are equal. Rather, it implies that both of these quantities have the same *supertrace* when this supertrace is evaluated over the entire string spectrum. For  $\tau_2$ -independent insertions  $\mathcal{X} = \mathbb{X}$  this identity takes the simple form [17]

$$\text{Str}(\tau_2^{-1} \mathbb{X}) = -\frac{1}{4\pi \mathcal{M}^2} \text{Str}(\mathbb{X} M^2). \quad (3.55)$$

Moreover, for the special case in which  $\mathcal{X} = \mathbb{X} = \mathbf{1}$ , we find

$$\text{Str}(\tau_2^{-1}) = -\frac{1}{4\pi \mathcal{M}^2} \text{Str} M^2. \quad (3.56)$$

In conjunction with Eq. (2.3), this then provides an alternate supertrace expression for the cosmological constant  $\Lambda$  in Eq. (2.2), specifically

$$\Lambda = -\frac{\pi}{6} \mathcal{M}^4 \text{Str}(\tau_2^{-1}). \quad (3.57)$$

Finally, for  $\mathcal{X} = \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2$ , the identity in Eq. (3.54) takes the form

$$\text{Str}(\tau_2 \mathbb{X}_2) = \frac{1}{4\pi \mathcal{M}^2} \text{Str}[(\tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2) M^2]. \quad (3.58)$$

This then allows us to express the amplitude in Eq. (3.49) in the simpler form

$$\langle \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2 \rangle = \frac{\pi}{3} \text{Str}(\mathbb{X}_1 + \tau_2 \mathbb{X}_2), \quad (3.59)$$

in agreement with Eq. (3.53).

The identity in Eq. (3.54) is rather astonishing, leading to results such as those in Eqs. (3.55), (3.56), and (3.58) in which a change in the power of  $\tau_2$  within a supertrace can be traded for the insertion of an additional factor of the squared *mass*. Moreover, within this context, the condition in Eq. (3.50) helps us to interpret results such as those in Eq. (3.55). Recall that the definition of the supertrace in Eq. (3.44) involves taking the  $\tau_2 \rightarrow 0$  limit. Thus, when we insert a factor of  $\tau_2^{-1}$  into a supertrace—such as on the left side of Eq. (3.55)—or when we equivalently insert a factor of  $M^2$  into the supertrace—such as on the right side of Eq. (3.55)—it would *a priori* seem that we are pushing this supertrace towards a divergence, especially since our supertrace involves summing over an infinite tower of states with ever-increasing masses. However, these relations come with the additional constraint in Eq. (3.50)

which tells us that the supertrace *without* these factors of  $\tau_2^{-1}$  or  $M^2$  actually vanishes. These extra factors thus “lift” the value of this supertrace away from zero and thereby allow it to have the nonzero result which matches the value of the corresponding amplitude.

Because of its central importance, it is critical that we understand the implications of Eq. (3.50). As we have shown, this relation holds *regardless* of whether  $\mathcal{X}$  contains leading  $\tau_2$  factors. If  $\mathcal{X}$  does not contain any explicit leading  $\tau_2$  factors of its own (e.g.,  $\mathcal{X} = \mathbb{X} = \mathbf{1}$ ), then this constraint is at its most powerful, requiring that

$$\sum_n a_{nn} \mathbb{X}_{nn} e^{-\pi\alpha' M_n^2 \tau_2} \rightarrow 0 \quad \text{as } \tau_2 \rightarrow 0. \quad (3.60)$$

More precisely, in four uncompactified dimensions, we have from Eq. (3.43) that

$$\left| \sum_n a_{nn} \mathbb{X}_{nn} e^{-\pi\alpha' M_n^2 \tau_2} \right| \lesssim \tau_2 \quad \text{as } \tau_2 \rightarrow 0. \quad (3.61)$$

However, even if  $\mathcal{X}$  has leading  $\tau_2$  factors, the constraint in Eq. (3.50) has teeth. As an example, let us suppose that  $\mathcal{X}$  has a leading power  $\tau_2^\ell$  with some value  $\ell > 0$ , so that we can write  $\mathcal{X} = \tau_2^\ell \mathbb{X}$ . In such cases the constraints in Eqs. (3.50) and (3.61) tell us that

$$\left| \sum_n a_{nn} \mathbb{X}_{nn} e^{-\pi\alpha' M_n^2 \tau_2} \right| \lesssim \tau_2^{1-\ell} \quad \text{as } \tau_2 \rightarrow 0. \quad (3.62)$$

This constraint continues to provide a significant bound on the spectral sum on the left side of this equation: as  $\tau_2 \rightarrow 0$ , we find that this spectral sum can grow no more rapidly than a *power* of  $1/\tau_2$ , with the power given by  $\ell - 1$ . This constraint is obviously strongest for small values of  $\ell$ , but nevertheless rules out all exponential growth as  $\tau_2 \rightarrow 0$ . This is a significant exclusion, since exponential growth would be the naïve expectation in string theory given that the numbers of bosonic and fermionic states in string theory generically grow exponentially with mass.

To illustrate this phenomenon numerically, let us examine a spectrum with alternating, exponentially growing boson/fermion surpluses, as predicted by misaligned supersymmetry, where the growth rates scale as  $\sim e^{\sqrt{n}}$  (as required by conformal invariance and the Hagedorn transition). For numerical simplicity, we shall model  $a_{nn}$  (the net number of bosons minus the number of fermions at a given level  $n$ ) as having the functional form

$$a_{nn} = (-1)^n e^{\sqrt{n}} \quad (3.63)$$

where  $n$  schematically represents the string level and thus can be associated with the eigenvalue of the

corresponding mass  $\alpha' M^2$  or charge  $Q^2$ . Here the factor  $(-1)^n$  indicates that the even levels are presumed to have surpluses of bosons relative to fermions while the odd levels have surpluses of fermions relative to bosons. Of course, such a functional form cannot correspond to any actual modular-invariant string theory—for example, the  $a_{nn}$  values in this little exercise are not even integers—and far more sophisticated functional forms of this general type emerge in actual string models [1,8,9]. However, this simple functional form does capture the essential consequence of misaligned supersymmetry, namely that we have alternating bosonic and fermionic surpluses for which no boson/fermion pairings are possible anywhere in the infinite string spectrum, with the degeneracies  $a_{nn}$  lying along equal but opposite bosonic and fermionic “envelope functions”  $\sim e^{\sqrt{n}}$  [1,8,9]. We shall also imagine that our insertion  $\mathbb{X}$  has eigenvalues  $\mathbb{X}_{nn} \sim n^\beta$  for some exponent  $\beta$ , and consider the spectral sums

$$f_\beta(\tau_2) \equiv \sum_{n=0}^{\infty} a_{nn} n^\beta e^{-n\tau_2}. \quad (3.64)$$

For example, the different values of  $\beta$  might correspond to different powers of mass/charge insertions, with an insertion of  $n^\beta$  corresponding to an insertion of  $2\beta$  powers of mass  $M$  or charge  $Q$ . Indeed, in such cases the fully modular-invariant insertion would also include an overall factor  $\tau_2^\beta$ , and thus the different values of  $\beta$  correspond to different values of  $\ell$  in Eq. (3.62).

In general, we know that for all  $\beta > 0$  we must have  $f_\beta(\tau_2) \rightarrow 0$  as  $\tau_2 \rightarrow \infty$ , since in this limit only the contributions from the massless ( $n = 0$ ) states survive, and these vanish for all  $\beta > 0$ . However, as we dial  $\tau_2$  to smaller values, there will be less and less suppression of the contributions from the heavier states. Our functions  $f_\beta(\tau_2)$  therefore become increasingly sensitive to the exponentially growing oscillations that exist throughout the massive levels with  $n > 0$ . Thus, as  $\tau_2 \rightarrow 0$ , we expect that  $f_\beta(\tau_2)$  will diverge exponentially while simultaneously experiencing rapid oscillations which prevent the extraction of any smooth  $\tau_2 \rightarrow 0$  limit.

In Fig. 2 we plot  $f_\beta(\tau_2)$  as a function of  $\tau_2$  for  $0 \leq \beta \leq 6$ . (Such plots are similar in spirit to those in Ref. [18].) As expected, we see that  $f_\beta(\tau_2) \rightarrow 0$  as  $\tau_2 \rightarrow \infty$  for all  $\beta > 0$ , as discussed above. However, as  $\tau_2$  becomes smaller, we find that for each  $\beta$  our function  $|f_\beta(\tau_2)|$  *does not diverge exponentially as  $\tau_2 \rightarrow 0$ , but instead remains within the bounds indicated in Eq. (3.62). Indeed, this happens even without the ability to realize any boson/fermion pairings within the associated spectrum. Moreover, for the simple functional form in Eq. (3.63), we even find that  $f_\beta(\tau_2)$  approaches a finite value as  $\tau_2 \rightarrow 0$ . Thus, we see that the spectrum in Eq. (3.63) already does a good job of satisfying our supertrace constraints, and even has a  $\tau_2 \rightarrow 0$  limit which comes close to vanishing in the  $\beta = 0$  case.*

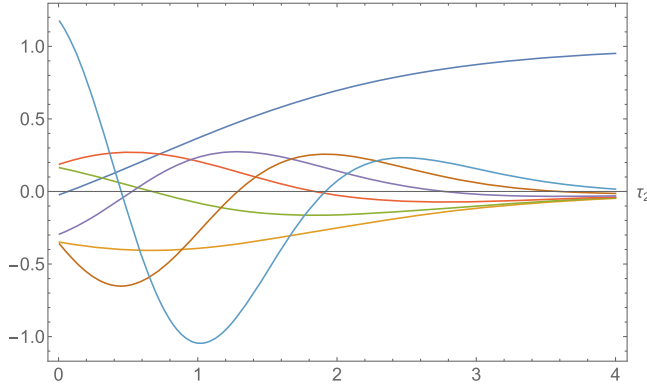


FIG. 2. The infinite sums  $f_\beta(\tau_2)$  in Eq. (3.64), plotted as functions of  $\tau_2$  for  $0 \leq \beta \leq 6$ . The  $\beta = 0$  curve is the upper blue curve which asymptotes to 1 as  $\tau_2 \rightarrow \infty$ , while the  $\beta = 1, 2, \dots, 6$  curves all asymptote to 0 as  $\tau_2 \rightarrow \infty$  and can be identified according to the increasing values that they have at any fixed large value of  $\tau_2$  (with  $n = 1$  orange,  $n = 2$  green,  $n = 3$  red, etc.). As  $\tau_2$  drops to zero from a large value, we see that our functions  $f_\beta(\tau_2)$  do not diverge (as would have been expected as we slowly remove the  $\tau_2$  cutoff), but instead remain within the bounds indicated in Eq. (3.62), leading to finite values for  $f_\beta(\tau_2)$  as  $\tau_2 \rightarrow 0$ .

Once again, we stress that the simple spectrum in Eq. (3.63) is only a caricature of an actual fully modular-invariant string spectrum. This exercise nevertheless illustrates how even the constraint in Eq. (3.62) has teeth.

As a final remark, we note that not every oscillating functional form for  $a_{nn}$  will exhibit this behavior. Indeed, the functional form in Eq. (3.63) is particularly “stringy”: rather than relying on boson/fermion pairings at any mass level, the controlled behavior as  $\tau_2 \rightarrow 0$  occurs as the result of tight constraints that involve the numbers of states across the entire (infinite) string spectrum. From this perspective the critical aspect of the spectrum in Eq. (3.63) is that the bosonic and fermionic states share the same exponentially growing degeneracy profile function  $e^{\sqrt{n}}$  while nevertheless sampling this function at “misaligned” values (in this case, with even  $n$  for bosons and odd  $n$  for bosons). This is the underpinning of misaligned supersymmetry, as discussed in Ref. [9].

To see that this is the critical feature, let us imagine a more “field-theoretic” spectrum in which the bosonic states continue to have degeneracies  $a_{nn} \sim e^{\sqrt{n}}$  but in which their fermionic would-be superpartners have these same degeneracies but with masses lifted by some small supersymmetry-breaking scale  $\delta n$ :

$$\begin{aligned} f_\beta^{(B)}(\tau_2) &\equiv \sum_n a_{nn} e^{-n\tau_2} \\ f_\beta^{(F)}(\tau_2) &\equiv \sum_n a_{nn} e^{-(n+\delta n)\tau_2}. \end{aligned} \quad (3.65)$$

Here the two distributions in Eq. (3.65) indicate bosonic and fermionic states respectively. Of course, this is a very

natural spectrum from a field-theoretic perspective, exhibiting a clear boson/fermion pairing. However, such cases lack modular invariance, and indeed we find that the corresponding  $f_\beta(\tau_2) \equiv f_\beta^{(B)}(\tau_2) - f_\beta^{(F)}(\tau_2)$  functions not only grow exponentially as  $\tau_2 \rightarrow 0$ , thereby violating the constraints in Eq. (3.62), but also exhibit increasingly violent oscillations that preclude numerical extraction of any smooth limiting values as  $\tau_2 \rightarrow 0$ . Similar problematic results arise for other paired bosonic/fermionic splitting patterns as well.

Thus, to summarize the results of this section, we find that any four-dimensional string amplitude  $\langle \mathcal{X} \rangle$  with a modular-invariant insertion  $\mathcal{X}$  can be written as [17]

$$\langle \mathcal{X} \rangle = \frac{\pi}{3} \text{Str}(D_{\tau_2} \mathcal{X}) = \frac{\pi}{3} \text{Str}(\tau_2^{-1} \mathcal{X}) \quad (3.66)$$

where  $D_{\tau_2}$  is defined in Eq. (3.47). In conjunction with Eq. (3.66), we also have the constraint in Eq. (3.50) which tightly constrains the spectrum of states and renders the quantities in Eq. (3.66) finite. Given the discussion above, we see that supertrace expressions such as these are meaningful and convergent precisely because we are working within a string-theoretic context wherein the corresponding spectra are governed by modular invariance and misaligned supersymmetry, even though bosonic/fermionic pairings are no longer possible. Indeed, it is in this way that string theory maintains its finiteness—even without space-time supersymmetry, and even in the face of not only exponentially growing towers of states but also exponentially growing *net* (bosonic minus fermionic) numbers of states [1,8,9].

#### D. Entwined amplitudes and entwined supertraces

In the previous subsection we derived the general results in Eqs. (3.46) and (3.50) for the Rankin-Selberg transform. At first glance, our next task would then appear to be to apply these results for the specific operator insertions that arise in our calculation of the gauge couplings. In particular, as evident from Eq. (3.20), there are four specific operator insertions that will be required in our gauge-coupling calculation:

$$\bar{Q}_H^2, \quad \bar{Q}_H^2 Q_G^2, \quad \bar{E}_2, \quad \text{and} \quad \bar{E}_2 Q_G^2. \quad (3.67)$$

The first and third of these come from the insertion of  $\mathbb{X}_1$ , while the second and fourth come from the insertion of  $\mathbb{X}_2$ .

Unfortunately, we cannot yet apply the Rankin-Selberg formalism to all of these operator insertions. The insertions  $\bar{Q}_H^2$  and  $\bar{Q}_H^2 Q_G^2$  behave as assumed above for  $A^{(\ell)}$ , depositing their corresponding  $(m, n)$  eigenvalues  $A_{mn}^{(\ell)}$  within the partition-function trace as in Eq. (3.26). The same is also true of the factor  $Q_G^2$  within the fourth operator in Eq. (3.67). However the “operator”  $\bar{E}_2$  is actually a

function of the complex parameter  $\tau$  and thus has its own double power-series expansion in  $(q, \bar{q})$ . Moreover, because we are to insert the same  $\bar{E}_2$  function for each state within the string spectrum, insertion of this  $\bar{E}_2$  function into any partition-function trace yields nothing but a *product* of the original partition-function trace and the  $\bar{E}_2$  function. As a result, if  $Z_X$  denotes the partition-function trace with an operator  $X$  inserted (i.e., if  $Z_X \sim \sum_{m,n} a_{mn} X_{mn} \bar{q}^m q^n$ ), then  $Z_{\bar{E}_2} = Z \cdot \bar{E}_2$  and  $Z_{X\bar{E}_2} = Z_X \cdot \bar{E}_2$ . The integrand of the (unregulated) amplitude for our gauge-coupling calculation thus generally takes the form

$$Z_{\mathcal{A}+\mathcal{B}\bar{E}_2} = Z_{\mathcal{A}} + Z_{\mathcal{B}} \cdot \bar{E}_2 \quad (3.68)$$

where  $\mathcal{A} = \sum_{\ell} \tau_2^{\ell} \mathbb{A}^{(\ell)}$  and  $\mathcal{B} = \sum_{\ell} \tau_2^{\ell} \mathbb{B}^{(\ell)}$ . Here  $\mathbb{A}^{(\ell)}$  and  $\mathbb{B}^{(\ell)}$  are the analogs of  $\mathbb{X}_1$  and  $\mathbb{X}_2$ . Indeed, from Eq. (3.20) we have

$$\begin{aligned} \mathbb{A}^{(1)} &= \frac{\xi}{2\pi} \bar{Q}_H^2, & \mathbb{A}^{(2)} &= -2\bar{Q}_H^2 Q_G^2, \\ \mathbb{B}^{(1)} &= -\frac{\xi}{24\pi}, & \mathbb{B}^{(2)} &= \frac{1}{6} Q_G^2. \end{aligned} \quad (3.69)$$

Note that  $\mathbb{B}^{(1)}$  is just a constant. Given the form in Eq. (3.68), we thus expect the results of our calculations to involve not only the supertraces emerging from  $Z_{\mathcal{A}}$ , as discussed in the previous subsection, but also the supertraces emerging from the *product*  $Z_{\mathcal{B}}\bar{E}_2$  in Eq. (3.68). Unfortunately, as we shall now see, this product structure renders the extraction of the corresponding supertraces more complicated than before.

To see the implications of this product structure, let us begin by considering a completely general product of the form  $Z_1 \cdot Z_2$ , where the factors  $Z_i$  are arbitrary modular-invariant functions which each have their own  $(q, \bar{q})$  power-series expansions of the forms

$$\begin{aligned} Z_1 &\sim \sum_{m,n} b_{mn} \bar{q}^m q^n, \\ Z_2 &\sim \sum_{r,s} c_{rs} \bar{q}^r q^s. \end{aligned} \quad (3.70)$$

In order to apply the Rankin-Selberg procedure as in the previous subsection, we must first recast the product of these two power-series expansions into the form of a single power-series expansion, as in Eq. (3.26). We then wish to consider the corresponding  $g$ -function, as in Eq. (3.28).

Given Eq. (3.70), we immediately see that

$$Z_1 \cdot Z_2 \sim \sum_{m,n} \sum_{r,s} b_{m,n} c_{r,s} \bar{q}^{m+r} q^{n+s}. \quad (3.71)$$

It then follows that

$$g_{Z_1 \cdot Z_2} \sim \sum_p d_{pp} (\bar{q}q)^p \quad (3.72)$$

where the ‘‘on-shell’’ degeneracies  $d_{pp}$  associated with the product are given by the discrete convolution

$$d_{pp} \equiv \sum_{m,n} b_{mn} c_{p-m,p-n}. \quad (3.73)$$

This truncation to a single summation in Eq. (3.72) occurs because the  $\tau_1$ -integration within the definition of  $g_{Z_1 \cdot Z_2}(\tau_2)$  selects only those terms in Eq. (3.71) for which  $m+r = n+s = p$ . Thus the new ‘‘degeneracies’’  $d_{pp}$  within  $g_{Z_1 \cdot Z_2}$  depend on the degeneracies of both the level-matched *and* nonlevel-matched states within the original  $Z_1$  and  $Z_2$  factors. In other words, the two factors  $Z_1$  and  $Z_2$  are now *entwined* within the product  $Z_1 \cdot Z_2$ . Nevertheless, as is evident upon comparing Eq. (3.28) and (3.72), the quantities  $\{d_{pp}\}$  now play the same role for a product of two modular functions as previously played by the degeneracies  $\{a_{nn} A_{nn}\}$  in the case of a single modular function.

If  $Z_2$  is purely antiholomorphic, we can set  $s = 0$  above. We then find that  $p = n$ , whereupon

$$d_{pp} \equiv \sum_{r=0}^{r_{\max}} b_{p-r,p} c_{r0}. \quad (3.74)$$

In general, right-moving worldsheet energies in string theory are bounded from below by the right-moving worldsheet vacuum energy  $\Delta$ . For example, we have  $\Delta = -1$  for the bosonic string and  $\Delta = -1/2$  for the superstring and heterotic string. In all cases, we must therefore have  $p-r \geq \Delta$ . For any value of  $p$ , as in Eq. (3.74), this therefore imposes an upper bound

$$r_{\max} \equiv \lfloor p - \Delta \rfloor \quad (3.75)$$

where  $\lfloor x \rfloor$  denotes the greatest integer  $\leq x$ .

For the special case with  $Z_2 = \bar{E}_2$ , we find  $c_{r0} = \chi_r$  in Eq. (3.74), where  $\chi_r$  is defined in Eq. (3.13). This then yields

$$\begin{aligned} d_{pp} &= b_{pp} - 24 \sum_{r=1}^{r_{\max}} \sigma(r) b_{p-r,p} \\ &= \sum_{r=0}^{r_{\max}} \chi_r b_{p-r,p}. \end{aligned} \quad (3.76)$$

We thus see that we obtain the expected result  $d_{pp} = b_{pp}$  (for  $r = 0$ ) along with a ‘‘correction’’ term (for  $r \geq 1$ ) which reflects the entwining of the theories and which is induced by the modular completion.

There is interesting physics in this entwinement. For the lowest-lying levels with  $p = 0$ ,  $p = 1$ , and so forth, we have

$$\begin{aligned} d_{00} &= b_{00} \\ d_{11} &= b_{11} - 24\sigma(1)b_{01} \\ d_{22} &= b_{22} - 24\sigma(1)b_{12} - 24\sigma(2)b_{02} \\ d_{33} &= b_{33} - 24\sigma(1)b_{23} - 24\sigma(2)b_{13} - 24\sigma(3)b_{03}. \end{aligned} \quad (3.77)$$

Of course, if  $Z_1$  represents the string partition function itself without any insertions, we then have  $b_{mn} = a_{mn}$  where  $a_{mn}$  tallies the number of bosonic minus fermionic string states with worldsheet energies  $(m, n)$ , as in Eq. (3.25). Otherwise, if  $Z_1$  represents the string partition function with an insertion whose  $(m, n)$  eigenvalues are given by  $A_{mn}$ , then  $b_{mn} = a_{mn}A_{mn}$ . In either case, we see from Eq. (3.77) that off-shell (nonlevel-matched) *purely stringy* states are entering into the entwinement. Moreover, we should also remember that these are not the only states in the theory. For example, for  $p = 1/2$ ,  $p = 3/2$ , *etc.*, we also have

$$\begin{aligned} d_{\frac{1}{2}\frac{1}{2}} &= b_{\frac{1}{2}\frac{1}{2}} - 24\sigma(1)b_{-\frac{1}{2}\frac{1}{2}} \\ d_{\frac{3}{2}\frac{3}{2}} &= b_{\frac{3}{2}\frac{3}{2}} - 24\sigma(1)b_{\frac{1}{2}\frac{3}{2}} - 24\sigma(2)b_{-\frac{1}{2}\frac{3}{2}}. \end{aligned} \quad (3.78)$$

Thus even off-shell states with right-moving *tachyonic* mass contributions (i.e., with  $\alpha' M_R^2 < 0$ ) are now entering into the entwinement, even though our theory is presumed to lack physical on-shell tachyons. (Indeed, if this had been the holomorphic  $E_2$ -function rather than the antiholomorphic  $\bar{E}_2$ -function, even the proto-graviton [19] would have entered into the entwinement.) Of course, this analysis pertains to heterotic strings. For Type II theories, by contrast, the proto-graviton states will enter the entwinement even for  $\bar{E}_2$ , given that the Type II string has a holomorphic/antiholomorphic (or worldsheet left-moving/right-moving) symmetry.

We also note that in all cases we must always have  $p \geq 0$ . This restriction arises for two reasons: because our original theory is presumed tachyon-free (with  $b_{pp} = 0$  for all  $p < 0$ ), and because  $d_{pp}$  with  $p < 0$  cannot emerge via entwinement (i.e., from some state  $b_{p-r,p}$  with  $r \geq 1$ ) because right-moving worldsheet energies in string theory are never smaller than  $-1$ , even for the bosonic string. The values of  $p$  within the  $p$ -sum are otherwise unconstrained, and depend on the spectrum of string modes, Kaluza-Klein modes, and winding modes of the particular string under study.

Given this understanding of the nature of the entwinement induced by the presence of the  $\bar{E}_2$  factor, we can now proceed to derive the general Rankin-Selberg transformation of the amplitude whose integrand is given in Eq. (3.68). Following from Eq. (3.34), we have

$$\langle \mathcal{A} + \mathcal{B}\bar{E}_2 \rangle = \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} g_{\mathcal{A} + \mathcal{B}\bar{E}_2}(\tau_2) \quad (3.79)$$

where

$$g_{\mathcal{A} + \mathcal{B}\bar{E}_2}(\tau_2) = g_{\mathcal{A}}(\tau_2) + g_{\mathcal{B}\bar{E}_2}(\tau_2). \quad (3.80)$$

As before, the first contribution is given by

$$g_{\mathcal{A}}(\tau_2) = \sum_{\ell} \tau_2^{\ell-1} \sum_p a_{pp} \mathbb{A}_{pp}^{(\ell)} e^{-4\pi p \tau_2} \quad (3.81)$$

where in the exponential we identify  $p$  with  $\alpha M^2/4$ , as usual, with  $M$  denoting the total mass of the given string state within the  $p$ -sum. By contrast, for the second contribution in Eq. (3.80), we see from Eqs. (3.72) and (3.74) that

$$g_{\mathcal{B}\bar{E}_2}(\tau_2) = \sum_{\ell} \tau_2^{\ell-1} \left[ \sum_p \sum_{r=0}^{r_{\max}} \chi_r a_{p-r,p} \mathbb{B}_{p-r,p}^{(\ell)} \right] e^{-4\pi p \tau_2} \quad (3.82)$$

where  $\chi_r$  is given in Eq. (3.13) and where the quantity in square brackets is nothing but  $d_{pp}$  in Eq. (3.72). Indeed, the double  $(p, r)$ -sum in Eq. (3.82) is essentially a sum over all of the string states with right- and left-moving worldsheet energies  $(m, n)$  with  $m \leq n$ , even those with  $m \neq n$ . Of course, all string states have  $m - n \in \mathbb{Z}$  as a result of the invariance of the string partition function under  $\tau \rightarrow \tau + 1$ .

The expression in Eq. (3.82) is properly convergent: the  $r$ -summation has only a finite range  $0 \leq r \leq r_{\max}$  for any fixed value of  $p$ , while the subsequent  $p$ -sum—although infinite—is kept convergent by the  $p$ -suppressed exponential. However, it is possible to rearrange the terms of this double  $(p, r)$  sum in a manner which continues to include all  $(m, n)$  string states with  $m \leq n$  but which renders this convergence more explicit and is closer to the fundamental string symmetries.

To see this, let us imagine the different  $(m, n) \equiv (p - r, p)$  string states as populating a matrix, with  $m \equiv p - r$  and  $n \equiv p$  indicating the corresponding row and column, respectively. Such an arrangement is illustrated in Fig. 3. Indeed, we see that for any fixed  $p$  there is a maximum corresponding value of  $r$ , as already noted in Eq. (3.75). The  $(p, r)$  double sum as written in Eq. (3.82) then corresponds to tallying our states vertical column by vertical column within the upper triangular  $m \leq n$  portion of the matrix (i.e., the portion on or above the diagonal). However, as is evident from Fig. 3, there is another way in which we might sum these states. First, we can sum the states which lie along the diagonal (which we will now call the *principal diagonal*)—these are the physical string states, all of which have  $r = 0$ . Next, we can sum the states along the “first shifted diagonal” which is one column/row displaced (or shifted) from the principal

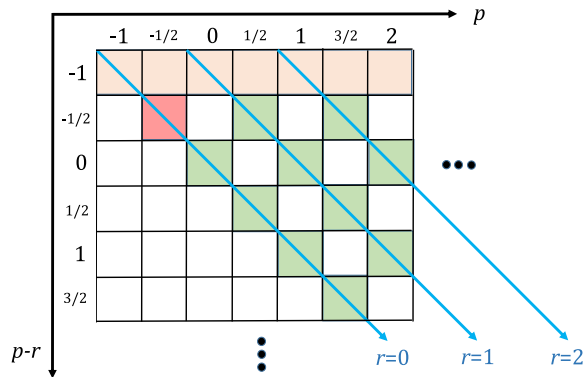


FIG. 3. String states arranged as a matrix according to their right-moving (vertical) and left-moving (horizontal) worldsheet energies  $(m, n) \equiv (p - r, p)$ , respectively. The states with  $r = 0$  lie along the principal diagonal, while the states with  $r = 1, 2, \dots$  lie along successive shifted diagonals. The requirement  $r \geq 0$  selects only those string states along or above the principal diagonal, and the  $\tau \rightarrow \tau + 1$  invariance of the partition function ensures that only those “squares” shaded in green with  $m - n \in \mathbb{Z}$  can be populated. The pink square is necessarily empty in any tachyon-free theory, and the row shaded in orange is excluded for heterotic strings because such strings have right-moving worldsheet energies  $\geq -1/2$ . In drawing this figure we have assumed that states populate only integer or half-integer mass levels, but in general the spectrum of states can be far denser and may even approach a continuum in  $(m, n)$  [or equivalently in  $p$ ] for exceedingly large or small compactification radii.

diagonal. These entries tally the contributions from the unphysical string states whose left- and right-moving worldsheet energies differ by one unit, i.e., states with  $r = 1$ . Next, we can sum over the “second shifted diagonal” which is two units removed from the principal diagonal, and so forth. In this way, we can equivalently reach all of our  $(m, n)$  states with  $m \leq n$  and  $m - n \in \mathbb{Z}$ . However, organizing our states according to the diagonals on which they lie is tantamount to organizing our states according to their  $L_0 - \bar{L}_0$  eigenvalues, where  $L_0$  is the zero-mode Virasoro operator. Indeed, this method of summing along diagonals is even suggested by the  $p$  and  $r$  variables themselves, since  $r$  essentially specifies the diagonal on which a given state lies while  $p$  then indicates the location along this diagonal.

Performing our summation along successive diagonals rather than column-by-column is tantamount to replacing

$$\sum_p \sum_{r=0}^{r_{\max}} \longrightarrow \sum_{r=0}^{\infty} \sum_p^{\infty} \quad (3.83)$$

within Eq. (3.82), where as always the  $p$ -sum is an infinite one which includes all of the values relevant for the particular string model in question. These values include not only integers [for states such as those in Eq. (3.77)] but also half-integers [for states such as those in Eq. (3.78)],

etc. Of course, while the  $p$ -sum on the left side of Eq. (3.83) nominally begins at  $p = 0$ , the  $p$ -sum on the right side of Eq. (3.83) begins at higher values of  $p$  so that  $p - r \geq \Delta$ . Implementing Eq. (3.83) within Eq. (3.82) is thus equivalent to reshuffling the contributions from the different terms within Eqs. (3.77) and (3.78) so that we sum along the vertical columns (rather than horizontal rows) within these equations. We therefore obtain

$$g_{B\bar{E}_2}(\tau_2) = \sum_{\ell} \tau_2^{\ell-1} \sum_{r=0}^{\infty} \sum_p \chi_r a_{p-r,p} \mathbb{B}_{p-r,p}^{(\ell)} e^{-4\pi p \tau_2}. \quad (3.84)$$

Our next step is to understand the exponential suppression factor  $e^{-4\pi p \tau_2}$  that appears in Eq. (3.84). In Eq. (3.81), a similar exponential factor appeared, and we identified  $p$  in the exponential with the total squared mass  $\alpha' M^2/4$  of the corresponding string state, where  $M^2 \equiv (M_L^2 + M_R^2)/2$ . However, this identification is no longer appropriate for Eq. (3.84). Because of the entwinement that appears in Eq. (3.74), we see that our exponential factor  $(\bar{q}q)^p$  actually represents the product

$$(\bar{q}q)^p = (\bar{q}^{p-r} q^p) \cdot \bar{q}^r \quad (3.85)$$

where only the parenthesized first factor encapsulates the mass contributions from the states in the original theory. By contrast, the second factor emerges purely from the entwinement function. Indeed, for  $r \neq 0$  this is precisely why our results are sensitive to the off-shell string states. However, this means that our identification of  $p$  and  $r$  with the masses of our underlying string states must now take the form

$$\alpha' M_R^2 = 4(p - r), \quad \alpha' M_L^2 = 4p, \quad (3.86)$$

implying that  $r = \alpha'(M_L^2 - M_R^2)/4 \equiv \alpha'(\Delta M^2)/4$ . This is consistent with our identification of  $r$  as indicating how far from the principal diagonal a given state lies. However, we now see directly from Eq. (3.86) that  $p$  represents only the *left-moving* contribution  $\alpha' M_L^2/4$  to the total mass of the state. Thus our exponential suppression factor is given by  $e^{-\pi \alpha' M_L^2 \tau_2}$ .

This result may seem strange, especially given that the exponential damping factor within the  $g$ -function has the level-matched form  $(\bar{q}q)^p$ , as it must. However, in the present circumstance we identify

$$\begin{aligned} (\bar{q}q)^p &= (\bar{q}^{p-r} q^p) \cdot \bar{q}^r \\ &= (\bar{q}^{\alpha' M_R^2/4} q^{\alpha' M_L^2/4}) \bar{q}^r \\ &= e^{-2\pi \tau_2 (\frac{1}{2} \alpha' M^2 + r)} e^{2\pi i \tau_1 [\frac{1}{4} \alpha' (M_L^2 - M_R^2) - r]}. \end{aligned} \quad (3.87)$$

Moreover, we know that  $r = \alpha'(M_L^2 - M_R^2)/4$ . Inserting this result for  $r$  then eliminates the second exponential in

Eq. (3.87)—as it must, given that  $g(\tau_2)$  cannot have any residual  $\tau_1$ -dependence—and the first exponential becomes

$$\begin{aligned} e^{-2\pi\tau_2(\alpha M^2+r)} &= e^{-2\pi\tau_2[\frac{1}{4}\alpha(M_L^2+M_R^2)+\frac{1}{4}\alpha(M_L^2-M_R^2)]} \\ &= e^{-\pi\alpha M_L^2\tau_2}, \end{aligned} \quad (3.88)$$

thereby reproducing the same exponential suppression as given above.

Thus, putting the pieces together, we find that the  $E$ -entwined portion of our  $g$ -function takes the form

$$g_{B\bar{E}_2}(\tau_2) = \sum_{\ell} \tau_2^{\ell-1} \sum_{r=0}^{\infty} \sum_p \chi_r a_{p-r,p} \mathbb{B}_{p-r,p}^{(\ell)} e^{-\pi\alpha M_L^2\tau_2}, \quad (3.89)$$

whereupon we have

$$\begin{aligned} \langle \mathcal{A} + B\bar{E}_2 \rangle &= \frac{\pi}{3} \lim_{\tau_2 \rightarrow 0} \tau_2^{-1} \left[ \sum_{\ell} \tau_2^{\ell} \sum_p a_{pp} \mathbb{A}_{pp}^{(\ell)} e^{-\pi\alpha M^2\tau_2} \right. \\ &\quad \left. + \sum_{\ell} \tau_2^{\ell} \sum_{r=0}^{\infty} \sum_p \chi_r a_{p-r,p} \mathbb{B}_{p-r,p}^{(\ell)} e^{-\pi\alpha M_L^2\tau_2} \right]. \end{aligned} \quad (3.90)$$

The rest of our analysis proceeds precisely as for the case without entwinement. Following Eq. (3.40), we can assume that the total quantity within square brackets in Eq. (3.90) has an overall  $\tau_2$ -dependence of the form  $\sum_{j=0}^{\infty} C_j \tau_2^j$ . We then have

$$\begin{aligned} C_0 &= \lim_{\tau_2 \rightarrow 0} [\dots], \\ C_1 &= \lim_{\tau_2 \rightarrow 0} \frac{d}{d\tau_2} [\dots] \end{aligned} \quad (3.91)$$

where [...] represents the quantity in square brackets in Eq. (3.90). The presumed finiteness of  $\langle \mathcal{A} + B\bar{E}_2 \rangle$  then leads us to conclude that

$$\begin{aligned} C_0 &= 0, \\ \langle \mathcal{A} + B\bar{E}_2 \rangle &= \frac{\pi}{3} C_1. \end{aligned} \quad (3.92)$$

Just as in the case without entwinement, these results can be given a direct interpretation in terms of supertraces over our string states. We have already remarked that our unentwined supertrace in Eq. (3.44) is nothing but an operator eigenvalue-weighted sum over the states that lie along the principal diagonal. Given this, let us also define an analogous *shifted supertrace* as the sum over the states that lie along the  $r$ th *shifted* diagonal:

$$\text{Str}^{(r)} X \equiv \lim_{\tau_2 \rightarrow 0} \sum_p a_{p-r,p} X_{p-r,p} e^{-\pi\alpha M_L^2\tau_2} \quad (3.93)$$

where  $\alpha M_L^2 = 4p$  and where the  $p$ -sum, as always, is over all of the states in the spectrum of the string model under consideration. Note that for  $r = 0$ , level-matching implies that  $M_L = M_R = M$ . We thus find that the  $r = 0$  shifted supertrace is nothing but our ordinary supertrace:

$$\text{Str}^{(r=0)} X = \text{Str} X. \quad (3.94)$$

The shifted supertraces with  $r > 0$  may thus be considered to be the generalizations of the ordinary supertrace to off-shell states—i.e., states that lie along nonprincipal diagonals. As noted above, the  $p$ -sums along nonprincipal diagonals typically begin with nonzero values of  $p$ , so that  $p - r$  continues to exceed the right-moving vacuum energy  $\Delta$  of the string model under consideration (with  $\Delta = -1/2$  for the heterotic string). However, this restriction merely characterizes the existing states in the theory. No states are excluded by these observations, and indeed these sums continue to tally all of the (off-shell) states that exist in the theory.

We can also define an  $E$ -entwined supertrace  $\text{Str}_E$  as the  $\chi_r$ -weighted sum of all of these shifted supertraces:

$$\text{Str}_E X \equiv \sum_{r=0}^{\infty} \chi_r \text{Str}^{(r)} X. \quad (3.95)$$

As such, this entwined supertrace  $\text{Str}_E X$  not only tallies both the physical and the unphysical string states but also organizes the latter naturally according to how nonlevel matched they are. In this way the  $E$ -entwined supertraces elegantly capture the string-theoretic nature of our full string spectrum, where the  $\bar{E}_2$  function determines the  $\chi_r$  coefficients and thereby determines the precise nature of the entwinement (motivating us to refer to this as an  $E_2$ -entwinement, or  $E$ -entwinement for short).

The definition in Eq. (3.95) has an important simplification if our supertrace is restricted to only those states with  $M_L = 0$ . Such states have  $p = 0$ , but for heterotic strings this in turn implies that we can only have  $r = 0$ , since any greater value of  $r$  would result in a right-moving worldsheet energy  $\bar{L}_0$  less than  $-1/2$ . We therefore have

$$\text{Str}_{M_L=0} X = \chi_0 \text{Str}_{M=0}^{(0)} X = \text{Str}_{M=0} X \quad (3.96)$$

where the first equality is a consequence of the restriction to  $r = 0$  while the second equality is a consequence of Eq. (3.94) in conjunction with the fact that  $\chi_0 = 1$ . Indeed, the quantity in Eq. (3.96) is nothing but  $a_{00} X_{00}$ .

Our results in Eqs. (3.91) and (3.92) can be easily expressed in terms of these shifted and entwined supertraces. In particular, we see that

$$\begin{aligned}
C_0 &= \text{Str } \mathcal{A} + \text{Str}_E \mathcal{B}, & \text{Str}^{(r)} \mathbf{1} &= 0 \quad \text{for all } r \\
C_1 &= \text{Str} \left( \frac{d\mathcal{A}}{d\tau_2} \right) - \text{Str}[\mathcal{A}(\pi\alpha' M_L^2)] & \Rightarrow \text{Str}_E \mathbf{1} &= 0. \\
&+ \text{Str}_E \left( \frac{d\mathcal{B}}{d\tau_2} \right) - \text{Str}_E[\mathcal{B}(\pi\alpha' M_L^2)] & & (3.101) \\
&= \text{Str}[D_{\tau_2} \mathcal{A}] + \text{Str}_E[D_{\tau_2}^{(L)} \mathcal{B}] & & (3.97)
\end{aligned}$$

where, in analogy to Eq. (3.47), we have

$$D_{\tau_2}^{(L)} \equiv \frac{d}{d\tau_2} - \pi\alpha' M_L^2. \quad (3.98)$$

Our Rankin-Selberg relations for the entwined case then become

$$\begin{aligned}
\text{Str } \mathcal{A} &= -\text{Str}_E \mathcal{B} \\
\langle \mathcal{A} + \mathbb{B}\bar{E}_2 \rangle &= \frac{\pi}{3} \{ \text{Str}[D_{\tau_2} \mathcal{A}] + \text{Str}_E[D_{\tau_2}^{(L)} \mathcal{B}] \}. \quad (3.99)
\end{aligned}$$

In cases for which the operator insertions  $\mathcal{A}$  and  $\mathcal{B}$  are  $\tau_2$ -independent, this last result reduces to

$$\langle \mathbb{A} + \mathbb{B}\bar{E}_2 \rangle = -\frac{1}{12M^2} [\text{Str}(\mathbb{A}M^2) + \text{Str}_E(\mathbb{B}M_L^2)]. \quad (3.100)$$

We also note that the results in Eq. (3.99) reduce to those in Eqs. (3.50) and (3.46) for the case in which the entwinement is removed, i.e., the case in which  $\bar{E}_2 \rightarrow 0$ , or  $\chi_r \rightarrow 0$  for all  $r \geq 0$ .

We have already seen in Eq. (2.5) that  $\text{Str } \mathbf{1} = 0$  in any closed-string theory which is free of physical (on-shell) tachyons. Indeed, this is a completely general result which was first obtained in Ref. [8] and which holds even if the string model in question lacks spacetime supersymmetry. Indeed, this result is one of the fundamental predictions of *misaligned supersymmetry* [8,9], a hidden symmetry which must always exist in any string spectrum and which plays a critical role in ensuring the finiteness of closed-string amplitudes. Given this result, it is natural to wonder whether the corresponding *entwined* supertrace  $\text{Str}_E \mathbf{1}$  vanishes as well. From Eq. (3.95) we see that this would require either that the shifted supertraces  $\text{Str}^{(r)} \mathbf{1}$  each vanish individually or take values which cancel in the sum over  $r$  in Eq. (3.95). However, just as  $\text{Str } \mathbf{1}$  vanishes when evaluated along the principal diagonal for any self-consistent tachyon-free closed string theory, it can be shown [20] that  $\text{Str}^{(r)} \mathbf{1} = 0$  as well—i.e., that this supertrace also necessarily vanishes along each of the shifted diagonals. This result is ultimately the result of an *off-shell* misaligned supersymmetry that exists within the *off-shell* structure of any tachyon-free string theory [9,20]. We thus find the general result

Given these results, it is now straightforward to evaluate the supertraces of the overall operator insertions  $\mathbb{X}_1$  and  $\mathbb{X}_2$  in Eq. (3.20). We thus obtain the relations

$$\begin{aligned}
\text{Str } \mathbb{X}_1 &= \frac{\xi}{2\pi} \text{Str } \bar{Q}_H^2, \\
\text{Str } \mathbb{X}_2 &= -2 \text{Str } \bar{Q}_H^2 Q_G^2 + \frac{1}{6} \text{Str}_E Q_G^2 \quad (3.102)
\end{aligned}$$

where  $\text{Str}_E$  is defined as in Eq. (3.95) in terms of principal and shifted supertraces and where we have eliminated a term proportional to  $\text{Str}_E \mathbf{1}$  that would otherwise have appeared in the first line of Eq. (3.102). Indeed, with these results we have succeeded in writing the supertraces of our operator insertions  $\mathbb{X}_1$  and  $\mathbb{X}_2$  in terms of the supertraces of our physical charges  $\bar{Q}_H^2$  and  $Q_G^2$  across the states in the string spectrum.

We see, then, that the entwinement has had a profound effect on our theory, leading to supertraces over more than merely the physical string states. At first glance, this might seem to violate our claim—as expressed in Sec. II—that the Rankin-Selberg procedure leads to supertraces over only the physical string states. However, the Rankin-Selberg procedure always involves performing the  $\tau_1$ -integral when defining  $g(\tau_2)$ , and thus it always projects onto those overall  $\bar{q}^m q^n$  contributions for which  $m = n$ . By contrast, what has occurred is that the presence of the  $\bar{E}_2$  insertion—which has its own intrinsic  $\bar{\tau}$ -dependence—has effectively *deformed* how the different  $(m, n)$  states ultimately contribute within the relevant modular integrands, ultimately allowing off-shell states within the original theory to become “physical” (i.e., level-matched) within the modular-completed calculation of the gauge couplings. This then allows such states to contribute in the large- $\tau_2$  limit (i.e., in the deep IR), just as we expect for physical states. Thus, in this sense, it is the *modular completion*—along with the appearance of the  $\bar{E}_2$  factor—which has deformed the notion of “physicality” insofar as the Rankin-Selberg procedure is concerned, allowing string states which were originally nonlevel matched to behave as physical states in our gauge-coupling calculation. Entwinement thereby widens the class of states which can ultimately contribute to the supertraces when calculating string amplitudes.

### E. Generic picture of running gauge couplings in string theory

Having assembled all of the relevant conceptual ingredients, we are now ready to tackle our main task: to utilize the Rankin-Selberg procedure in order to evaluate the string amplitude that yields the regulated one-loop contribution to the gauge coupling, and to recast this amplitude in terms of



spectral supertraces. As we have seen in Eqs. (3.21) and (3.68), this amplitude is given by

$$\begin{aligned}\hat{\Delta}_G &= \langle (\tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2) \hat{\mathcal{G}} \rangle \\ &= \langle (\mathcal{A} + \mathcal{B} \bar{E}_2) \hat{\mathcal{G}} \rangle \\ &= \langle [(\tau_2 \mathbb{A}^{(1)} + \tau_2^2 \mathbb{A}^{(2)}) \\ &\quad + (\tau_2 \mathbb{B}^{(1)} + \tau_2^2 \mathbb{B}^{(2)}) \bar{E}_2] \hat{\mathcal{G}} \rangle\end{aligned}\quad (3.103)$$

where the  $\mathbb{X}_\ell$  are given in Eq. (3.20), where  $\mathcal{A}, \mathcal{B}, \mathbb{A}^{(\ell)}$ , and  $\mathbb{B}^{(\ell)}$  are given in Eq. (3.69), and where the regulator function  $\hat{\mathcal{G}}_\rho(a, \tau)$  is given in Eq. (3.22).

The first thing we notice from the last term on the final line of Eq. (3.103) is that we are now dealing not with a mere entwinement between two modular functions  $Z_B(\tau, \bar{\tau})$  and  $\bar{E}_2(\bar{\tau})$ , but rather a *triple entwinement* between  $Z_B(\tau, \bar{\tau})$ ,  $\bar{E}_2(\bar{\tau})$ , and our regulator function  $\hat{\mathcal{G}}_\rho(a, \tau, \bar{\tau})$ . Thus, in principle, we should first develop a formalism for handling such a triple entwinement. Moreover, as we have discussed in Sec. III B, we would like to evaluate this amplitude as a function of  $a$ , since  $\rho a^2$  will eventually be identified with our running scale  $\mu$  in units of  $M_s$ . However, the value of  $a$  affects the nature of the triple entwinement in a highly nontrivial way.

For these reasons, we shall adopt a different approach. In particular, we shall follow the methodology first established in Appendix A of Ref. [2] for calculations of the Higgs mass, only suitably adapted for our gauge-coupling calculation. To do this, we observe upon comparing Eqs. (2.6) and (3.103) that the one-loop Higgs mass has the same algebraic structure as the one-loop contribution to the gauge coupling. Indeed, the only difference between the two expressions is a change in the particular operator insertions  $\mathbb{X}_\ell$ . However, the calculation in Appendix A of Ref. [2] does not rely on the precise operator insertions as long as they have the general modular structure  $\tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2$ . For this reason, we can borrow the results from Appendix A of Ref. [2] and then simply update these results using the new operator insertions appropriate for our gauge-coupling calculation.

This procedure is greatly facilitated by first observing that the form of the regulator function  $\hat{\mathcal{G}}_\rho(a, \tau)$  in Eq. (3.22) allows us to reduce the calculation of  $\hat{\Delta}_G$  to a calculation of the “reduced” amplitude

$$P(a) = \langle (\tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2) Z_{\text{circ}}(a, \tau) \rangle \quad (3.104)$$

where  $Z_{\text{circ}}(a, \tau)$  is the circle-compactification function in Eq. (3.23). Indeed, once we have evaluated  $P(a)$ , it follows from Eq. (3.22) that we can then easily evaluate  $\hat{\Delta}_G$  through the relation

$$\hat{\Delta}_G(\rho, a) = \frac{a^2}{1 + \rho a^2} \frac{\rho}{\rho - 1} \frac{\partial}{\partial a} [P(\rho a) - P(a)]. \quad (3.105)$$

Our task is therefore to evaluate  $P(a)$ . However, this is precisely what is done in Ref. [2] for cases in which the operators  $\mathbb{X}_\ell$  do not contain any entwinements. Indeed, in such cases it is shown that  $P(a)$  is given by

$$\begin{aligned}P(a) &= \text{Str}_{M=0} \mathbb{X}_1 [f_1(a) + f_2(a)] \\ &\quad + \text{Str}_{M=0} \mathbb{X}_2 [f_3(a)] \\ &\quad + \text{Str}_{M>0} \mathbb{X}_1 [f_2(a) + f_4(M, a)] \\ &\quad + \text{Str}_{M>0} \mathbb{X}_2 [f_5(M, a)]\end{aligned}\quad (3.106)$$

where

$$\begin{aligned}f_1(a) &= \frac{\pi a}{3}, \\ f_2(a) &= \frac{\pi}{3a}, \\ f_3(a) &= -\frac{2}{a} \log a \\ f_4(M, a) &= \frac{2}{\pi} \sum_{r=1}^{\infty} \left( \frac{M}{r\mathcal{M}} \right) K_1 \left( \frac{rM}{a\mathcal{M}} \right), \\ f_5(M, a) &= \frac{4}{a} \sum_{r=1}^{\infty} K_0 \left( \frac{rM}{a\mathcal{M}} \right).\end{aligned}\quad (3.107)$$

Here  $\mathcal{M} = M_s/(2\pi)$  is the reduced string scale and  $K_\nu(z)$  denotes the modified Bessel function of the second kind. Likewise, the notations  $\text{Str}_{M=0} \mathbb{X} f(M)$  and  $\text{Str}_{M>0} \mathbb{X} f(M)$  respectively indicate contributions from purely massless and massive string states. The result in Eq. (3.106) is exact for all  $a$  and holds for general unentwined modular-invariant operator insertions of the form  $\mathcal{X} = \tau_2 \mathbb{X}_1 + \tau_2^2 \mathbb{X}_2$ . Thus we see from Eq. (3.106) that our reduced amplitude in Eq. (3.104) for *unentwined* operators  $\mathbb{X}_1$  and  $\mathbb{X}_2$  can ultimately be expressed in terms of combinations of supertraces of the form  $\text{Str}[\mathbb{X} f(M)]$  for various combinations of  $\mathbb{X}_1$  and  $\mathbb{X}_2$  and for various functions  $f(M)$ . We further note that the functions  $f_1$  through  $f_3$  are wholly independent of any aspect of the spectrum of the string theory under study, and thus these functions can be taken outside their respective supertraces. By contrast, the functions  $f_4$  and  $f_5$  depend not only on the regulator parameter  $a$  but also on the mass  $M$  of the contributing state. Such functions are thus intrinsically part of the supertrace and cannot be factored out.

Although the result in Eq. (3.106) was derived for operators  $\mathbb{X}_\ell$  that do not involve any entwinement, it is not difficult to generalize this result to the case in which a given operator  $\mathbb{X}$  is entwined, i.e., takes the form  $\mathbb{X} = \mathbb{A} + \mathbb{B} \cdot \bar{E}_2$

with  $\mathbb{B} \neq 0$ . Indeed, tracing through the derivations that originally led to Eq. (3.106), we find that in such cases we can simply replace

$$\begin{aligned} \text{Str}[\mathbb{X}f(M)] \\ \rightarrow \text{Str}[Af(M)] + \text{Str}_E[\mathbb{B}f(M_L)] \end{aligned} \quad (3.108)$$

within Eq. (3.106). This illustrates the power of the entwined-supertrace formalism we have developed. Moreover, when restricting to massive states [i.e., states whose contributions to  $g(\tau_2)$  have an exponential  $\tau_2$ -dependent suppression], we have

$$\begin{aligned} \text{Str}_{M>0}[\mathbb{X}f(M)] \\ \rightarrow \text{Str}_{M>0}[Af(M)] + \text{Str}_{M_L>0}[\mathbb{B}f(M_L)]. \end{aligned} \quad (3.109)$$

The critical point here is that the restriction to massive states for  $\mathbb{X}$  becomes a restriction to states with positive *left-moving* mass  $M_L$  for the  $E$ -entwined supertrace. This makes sense since the exponential suppression within  $g(\tau_2)$  for the entwined supertrace depends on  $M_L^2$  rather than  $M^2$ . By contrast, when restricting to *massless* states [i.e., states that contribute to  $g(\tau_2)$  without exponential suppression], the  $f(M)$  function becomes a constant which can be pulled outside the trace. We can then push this one step further and write

$$\begin{aligned} \text{Str}_{M=0} \mathbb{X} &\rightarrow \text{Str}_{M=0} \mathbb{A} + \text{Str}_{M_L=0} \mathbb{B} \\ &= \text{Str}_{M=0} \mathbb{A} + \text{Str}_{M=0} \mathbb{B}. \end{aligned} \quad (3.110)$$

The top line is of course analogous to what occurs in Eq. (3.109), but the additional step—the passage to the second line—follows from Eq. (3.96).

As an example of Eq. (3.110), let us consider the cases when  $\mathbb{X} = \mathbb{X}_1$  and  $\mathbb{X} = \mathbb{X}_2$ . In these cases we have

$$\begin{aligned} \text{Str}_{M=0} \mathbb{X}_1 &= \frac{\xi}{2\pi} \text{Str}_{M=0} \bar{Q}_H^2 - \frac{\xi}{24\pi} \text{Str}_{M_L=0} \mathbf{1} \\ &= \frac{\xi}{2\pi} \text{Str}_{M=0} \left( \bar{Q}_H^2 - \frac{1}{12} \right) \end{aligned} \quad (3.111)$$

and

$$\begin{aligned} \text{Str}_{M=0} \mathbb{X}_2 &= -2 \text{Str}_{M=0} \bar{Q}_H^2 Q_G^2 + \frac{1}{6} \text{Str}_{M_L=0} Q_G^2 \\ &= -2 \text{Str}_{M=0} \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2. \end{aligned} \quad (3.112)$$

Interestingly, we see that in both cases the restriction to massless states has completely eliminated the modular completion that we originally performed in Eq. (3.15).

We further note the identity

$$\begin{aligned} \text{Str}_{M_L>0} \mathbf{1} &= \text{Str}_E \mathbf{1} - \text{Str}_{M_L=0} \mathbf{1} \\ &= -\text{Str}_{M=0} \mathbf{1} = \text{Str}_{M>0} \mathbf{1}. \end{aligned} \quad (3.113)$$

Here the first equality follows from the observation originally made in the paragraph below that containing Eq. (3.78), namely that  $p \geq 0$  (implying that  $M_L \geq 0$ ). Likewise, the second equality follows from Eq. (3.101) and the final equality follows from Eq. (2.5).

Let us now start with Eq. (3.106) and insert our expressions for  $\mathbb{X}_1$  and  $\mathbb{X}_2$ , where these expressions are given in Eq. (3.20). Bearing in mind the substitutions and simplifications in Eqs. (3.108) through (3.113), we obtain

$$\begin{aligned} P(a) &= \text{Str}_{M=0} \left\{ \frac{\xi}{2\pi} \left( \bar{Q}_H^2 - \frac{1}{12} \right) [f_1(a) + f_2(a)] \right\} \\ &\quad - \text{Str}_{M=0} \left[ 2 \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2 f_3(a) \right] \\ &\quad + \text{Str}_{M>0} \left[ \frac{\xi}{2\pi} \left( \bar{Q}_H^2 - \frac{1}{12} \right) f_2(a) \right] \\ &\quad + \text{Str}_{M>0} \left[ \frac{\xi}{2\pi} \bar{Q}_H^2 f_4(M, a) \right] \\ &\quad - \text{Str}_{M_L>0} \left[ \frac{\xi}{24\pi} f_4(M_L, a) \right] \\ &\quad - \text{Str}_{M>0} [2 \bar{Q}_H^2 Q_G^2 f_5(M, a)] \\ &\quad + \text{Str}_{M_L>0} \left[ \frac{1}{6} Q_G^2 f_5(M_L, a) \right]. \end{aligned} \quad (3.114)$$

This expression represents the line-by-line result of substituting the values of the  $\mathbb{X}_i$  from Eq. (3.20) into Eq. (3.106) and implementing the identities in Eqs. (3.108) through (3.113).

Additional manipulations can further simplify this expression and render it more compact while also simultaneously elucidating its algebraic structure. This rewriting will also be useful for understanding certain properties of the resulting running of the gauge couplings. In particular, we note that the final  $f_2(a)$  term on the first line of Eq. (3.114) can be joined with the third line of this equation in order to remove the  $M > 0$  restriction on the latter. The removal of this restriction then further allows us to eliminate the  $-1/12$  term as a result of the identity  $\text{Str} \mathbf{1} = 0$ . Likewise, the  $f_1(a)$  term on the first line can be combined with the expressions on the fourth and fifth lines in order to remove their  $M > 0$  and  $M_L > 0$  restrictions as well; indeed, these latter observations follow as a result of the fact that  $K_1(z) \sim z^{-1}$  as  $z \rightarrow 0$ , whereupon we see that

$$\lim_{M \rightarrow 0} \sum_{r=1}^{\infty} \left( \frac{M}{r\mathcal{M}} \right) K_1 \left( \frac{rM}{a\mathcal{M}} \right) = \sum_{r=1}^{\infty} \frac{a}{r^2} = \frac{\pi^2 a}{6}, \quad (3.115)$$

or equivalently

$$\lim_{M \rightarrow 0} f_4(M, a) = f_1(a). \quad (3.116)$$

Interpreting the quantity in Eq. (3.116) as  $f_4(0, a)$ , we thus find that Eq. (3.114) simplifies to

$$\begin{aligned} P(a) = & -\text{Str}_{M=0} \left[ 2 \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2 f_3(a) \right] \\ & + \text{Str} \left[ \frac{\xi}{2\pi} \bar{Q}_H^2 f_2(a) \right] \\ & + \text{Str} \left[ \frac{\xi}{2\pi} \bar{Q}_H^2 f_4(M, a) \right] \\ & - \text{Str}_E \left[ \frac{\xi}{24\pi} f_4(M_L, a) \right] \\ & - \text{Str}_{M>0} [2\bar{Q}_H^2 Q_G^2 f_5(M, a)] \\ & + \text{Str}_E \left[ \frac{1}{6} Q_G^2 f_5(M_L, a) \right]. \end{aligned} \quad (3.117)$$

At first glance it might seem that the first line of Eq. (3.117) could be rewritten in an analogous manner as the  $M \rightarrow 0$  limit of the terms in the final two lines. This would require that  $f_3(a)$  somehow emerge as the  $M \rightarrow 0$  limit of  $f_5(M, a)$ . However, this is ultimately not the case: we instead find that

$$f_5(M, a) \sim f_3(a) + \frac{M_s}{M} + \frac{2}{a} \log \left( e^\gamma \frac{M}{2M_s} \right) \quad \text{as } M \rightarrow 0. \quad (3.118)$$

While the extra  $a$ -independent constant  $M_s/M$  would ultimately prove irrelevant under the operation in Eq. (3.105), the logarithmic divergence in Eq. (3.118) spoils the uniform convergence of the Bessel-function sum. This issue has important implications and is discussed in detail in Sec. V of Ref. [2].

There is also a third way of rewriting these expressions which can be useful for understanding the ramifications of the entwinement in these theories. From Eqs. (3.95), (3.94), and (3.13) we can write

$$\text{Str}_E X = \text{Str} X + \sum_{r=1}^{\infty} \chi_r \text{Str}^{(r)} X. \quad (3.119)$$

However, as we discussed below Eq. (3.94), we must have  $p - r \geq \Delta$  where  $\Delta$  is the right-moving vacuum energy within the type of string theory under study. We thus find that  $p \geq \Delta + r$ , which implies that we cannot have  $r \geq 1$

unless  $p \geq \Delta + 1$ . This last constraint is evident in Fig. 3 for the heterotic case in which  $\Delta = -1/2$ . From Eq. (3.86) this last constraint corresponds to  $\alpha' M_L^2 \geq 4(\Delta + 1)$ . We can thus write

$$\text{Str}_E X = \text{Str}_{M < M_E} X + \text{Str}_{M_L \geq M_E} X \quad (3.120)$$

where  $M_E$  denotes the *entwinement scale*

$$M_E \equiv 2\sqrt{\Delta + 1} M_s \quad (3.121)$$

at which the entwinement first appears. For heterotic strings we have  $M_E = \sqrt{2} M_s$ . Note that Eq. (3.120) is a general result, valid for all strings. Despite the  $M < M_E$  upper limit on the first of the sums in Eq. (3.120), this sum can nevertheless involve a large number of states; this is especially true for cases in which compactification radii are far from the string scale.

Using Eq. (3.120), we can finally rewrite Eq. (3.117) in the form

$$\begin{aligned} P(a) = & -\text{Str}_{M=0} \left[ 2 \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2 f_3(a) \right] \\ & + \text{Str} \left[ \frac{\xi}{2\pi} \bar{Q}_H^2 f_2(a) \right] \\ & + \text{Str}_{M < M_E} \left[ \frac{\xi}{2\pi} \left( \bar{Q}_H^2 - \frac{1}{12} \right) f_4(M, a) \right] \\ & + \text{Str}_{M \geq M_E} \left[ \frac{\xi}{2\pi} \bar{Q}_H^2 f_4(M, a) \right] \\ & - \text{Str}_{M_L \geq M_E} \left[ \frac{\xi}{24\pi} f_4(M_L, a) \right] \\ & - \text{Str}_{0 < M < M_E} \left[ 2 \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2 f_5(M, a) \right] \\ & - \text{Str}_{M \geq M_E} [2\bar{Q}_H^2 Q_G^2 f_5(M, a)] \\ & + \text{Str}_E \left[ \frac{1}{6} Q_G^2 f_5(M_L, a) \right]. \end{aligned} \quad (3.122)$$

Although this expression has more individual terms than its two predecessors, it explicitly shows that the entwinement is wholly restricted to string states with  $M_L \geq M_E$ . Indeed, all terms that receive contributions from states with masses  $M < M_E$  depend not on  $\bar{Q}_H^2$  but rather on the explicit unentwined combination  $\bar{Q}_H^2 - 1/12$ . This statement includes the contribution on the second line of Eq. (3.122) once we realize that the  $1/12$  term that would otherwise appear there has vanished as a result of the identity  $\text{Str} \mathbf{1} = 0$ .

Equations (3.106), (3.114), (3.117), and (3.122) represent fully modular-invariant evaluations of the reduced string amplitude  $P(a)$  in Eq. (3.104), expressed purely in

terms of supertraces over our string states. Given that these supertraces are to be evaluated over the states within the spectrum of whatever the string model happens to be, these results are completely general and model agnostic, applicable to any four-dimensional string model—with or without spacetime supersymmetry—so long as the model lacks physical tachyons. Although the modular invariance of  $P(a)$  in each of these expressions is not manifest, it is hidden in supertrace identities that relate the various terms in these expressions to each other.

Using these expressions in conjunction with Eq. (3.105) we can then trivially evaluate our full desired amplitude  $\hat{\Delta}_G(\mu)$  for the running of the gauge couplings. Indeed, we see from Eq. (3.105) that we can turn  $P(a)$  into  $\hat{\Delta}_G(\mu)$  simply by replacing each term within  $P(a)$  according to the schematic substitution  $\text{Str}[\mathbb{X}f_i(a)] \rightarrow \text{Str}[\mathbb{X}\phi_i(\mu)]$  where the operators  $\mathbb{X}$  are unchanged and where the new functions  $\phi_i(\mu)$  for each  $f_i(a)$  are given by

$$\phi_i(\mu) \equiv \frac{1}{1 + \rho a^2} \frac{\rho}{\rho - 1} a^2 \frac{\partial}{\partial a} [f_i(\rho a) - f_i(a)] \quad (3.123)$$

where we first evaluate the right side of Eq. (3.123) as a function of  $\rho$  and  $a$ , and then identify  $\mu^2 \equiv \rho a^2 M_s^2$  with  $\rho = 2$  chosen as a benchmark value. Indeed, given the  $f_i$ -functions in Eq. (3.107), we find

$$\begin{aligned} \phi_1(\mu) &= \frac{\pi}{3} \frac{\mu^2/M_s^2}{1 + \mu^2/M_s^2}, \\ \phi_2(\mu) &= \frac{\pi}{3} \frac{1}{1 + \mu^2/M_s^2}, \\ \phi_3(\mu) &= \frac{2}{1 + \mu^2/M_s^2} \log\left(\frac{2\sqrt{2}eM_s}{\mu}\right), \\ \phi_4(M, \mu) &= \frac{1}{1 + \mu^2/M_s^2} \frac{1}{\pi} \left(\frac{M}{\mathcal{M}}\right)^2 [\mathcal{K}_0^{(0,1)}(z) + \mathcal{K}_2^{(0,1)}(z)], \\ \phi_5(M, \mu) &= \frac{2}{1 + \mu^2/M_s^2} [\mathcal{K}_1^{(1,2)}(z) - 2\mathcal{K}_0^{(0,1)}(z)], \end{aligned} \quad (3.124)$$

where  $z \equiv 2\sqrt{2}\pi M/\mu$  and where we have defined the Bessel-function combinations [2]

$$\mathcal{K}_\nu^{(n,p)}(z) \equiv \sum_{r=1}^{\infty} (rz)^n [K_\nu(rz/\rho) - \rho^p K_\nu(rz)]. \quad (3.125)$$

Note that  $\phi_1(\mu) + \phi_2(\mu) = \pi/3$ .

These  $\phi_i$ -functions are extremely important and have direct physical interpretations. While the specific charges that enter into the  $\mathbb{X}_i$  expressions tell us which specific quantity is under study (such as the Higgs mass versus the gauge coupling), and while the particular numerical values of these charges tell us about the particular string model under study, the  $\phi_i$  functions are essentially universal and tell us how these phenomenological quantities *run* as

functions of the scale  $\mu$  in the corresponding EFT. As we have seen, these running functions are universal for all quantities (such as the one-loop Higgs mass or gauge couplings) which have at most a logarithmic divergence in string theory prior to regularization. More specifically, substituting  $P(a) \rightarrow \hat{\Delta}(\mu)$  and  $f_i(a) \rightarrow \phi_i(\mu)$  within our previous expression for  $P(a)$  in Eq. (3.114), we find that

- (i)  $\phi_1(\mu)$  is the contribution to  $\hat{\Delta}_G(\mu)$  per unit  $(\mathbb{A}^{(1)} + \mathbb{B}^{(1)})$  charge from each massless physical state;
- (ii)  $\phi_2(\mu)$  is the additional contribution to  $\hat{\Delta}_G(\mu)$  per unit  $(\mathbb{A}^{(1)} + \mathbb{B}^{(1)})$  charge from each physical state, *regardless* of mass;
- (iii)  $\phi_3(\mu)$  is the additional contribution to  $\hat{\Delta}_G(\mu)$  per unit  $(\mathbb{A}^{(2)} + \mathbb{B}^{(2)})$  charge from each massless physical state;
- (iv)  $\phi_4(M, \mu)$  is the additional contribution to  $\hat{\Delta}_G(\mu)$  per unit  $\mathbb{A}^{(1)}$  charge for each physical state of nonzero mass  $M$ , while  $\phi_4(M_L, \mu)$  is  $\chi_r^{-1}$  times the additional contribution to  $\hat{\Delta}_G(\mu)$  per unit  $\mathbb{B}^{(1)}$  charge for each physical *or unphysical* string  $(m, n)$  state with left-moving mass  $M_L$  for which  $n - m \equiv r$  with  $r \geq 0$ ; and
- (v)  $\phi_5(M, \mu)$  is the additional contribution to  $\hat{\Delta}_G(\mu)$  per unit  $\mathbb{A}^{(2)}$  charge for each physical state of nonzero mass  $M$ , while  $\phi_5(M_L, \mu)$  is  $\chi_r^{-1}$  times the additional contribution to  $\hat{\Delta}_G(\mu)$  per unit  $\mathbb{B}^{(2)}$  charge for each physical *or unphysical* string  $(m, n)$  state with left-moving mass  $M_L$  for which  $n - m \equiv r$  with  $r \geq 0$ .

Here the  $\mathbb{A}^{(i)}$  and  $\mathbb{B}^{(i)}$  charges are given in Eq. (3.69), and the above results are quoted for bosonic states; fermionic states of course contribute with opposite signs. Once again, we stress that these results are completely general for all phenomenological quantities which diverge at most logarithmically when unregulated; it is only when we substitute the particular forms of  $\mathbb{A}^{(i)}$  and  $\mathbb{B}^{(i)}$  in Eq. (3.69) that we limit our attention to  $\hat{\Delta}_G(\mu)$  of the gauge couplings. Indeed, in the case of the Higgs mass in Ref. [2], no entanglement occurs and we have  $\mathbb{B}^{(1)} = \mathbb{B}^{(2)} = 0$ .

As we have seen, the results quoted above for  $P(a)$  in Eqs. (3.114), (3.117), and (3.122) come directly from the result in Eq. (3.106). This in turn is taken directly from Eq. (A15) of Ref. [2]. Although the derivation given in Ref. [2] is sufficient for the Higgs mass, it makes the implicit assumption that supertraces of the general form  $\text{Str}[\tau_2 \mathbb{X}_2 f(M)]$  are all vanishing as a result of the explicit factor of  $\tau_2$  within the supertrace. Otherwise, such terms would also have appeared in Eq. (3.106). At first glance, the absence of such terms from *all* calculations might appear to be justified, given that our supertraces are defined in Eq. (3.44) in terms of a limiting procedure that involves taking the  $\tau_2 \rightarrow 0$  limit. Indeed, in most circumstances (including those considered in Ref. [2], where the Higgs mass was calculated), the extra factor of  $\tau_2$  inside the

supertrace would drive the overall supertrace to vanish, as assumed. However, as discussed earlier, it is possible (especially near the borders of moduli space) that our spectrum of string states can become extremely dense. In such cases, the accumulated “pile-up” of states can cause quantities such as  $\text{Str} \mathbb{X}_2 f(M)$  to diverge, thereby allowing supertraces such as  $\text{Str}[\tau_2 \mathbb{X}_2 f(M)]$  to have nonzero values. This “pile-up” phenomenon will be discussed in more detail in Ref. [21].

In the present calculation of gauge couplings, we would like to maintain complete generality and allow our results to remain valid even as we approach the boundaries of moduli space. For this reason, we must amend our results for  $P(a)$  quoted above. However, it turns out that this is relatively straightforward and amounts to introducing only one additional contribution

$$\text{Str}_{M>0} \tau_2 \mathbb{X}_2 [f_2(a)] \quad (3.126)$$

within Eq. (3.106). This extra term will then propagate into Eqs. (3.114), (3.117), and (3.122). In fact, given that the “pile-up” phenomenon that gives rise to this term involves the infinite towers of massive states, we note that  $\text{Str}_{M=0} \tau_2 \mathbb{X}_2 [f_2(a)] = 0$ . The contribution in Eq. (3.126) can thus be equivalently written as

$$\text{Str} \tau_2 \mathbb{X}_2 [f_2(a)], \quad (3.127)$$

with no restriction on the masses of the states in the supertrace.

At this stage, we have now completed Step 8, as outlined in Sec. II. This enables us to extract a considerable amount of information about the running of  $\hat{\Delta}_G(\mu)$ . For example, let us consider the behavior of  $\hat{\Delta}_G(\mu)$  in the deep-IR limit, i.e., as  $\mu \rightarrow 0$ . As  $\mu \rightarrow 0$ , we find that  $\phi_1(\mu)$ ,  $\phi_4(\mu)$ , and  $\phi_5(\mu)$  all vanish; in the latter two cases this happens because the Bessel functions  $K_2(z)$  in Eq. (3.124) all vanish exponentially as  $z \rightarrow \infty$ . Thus, only  $\phi_2(\mu)$  and  $\phi_3(\mu)$  survive in the deep-IR limit. Of course,  $\phi_3(\mu)$  actually diverges in this limit. This divergence is not a surprise, however, since the deep-IR limit corresponds to the limit  $a \rightarrow 0$  in which our regulator is removed. *Thus this divergence corresponds to the logarithmic divergence of our original unregulated quantity.* As anticipated in Sec. III B, and as apparent from each of our above expressions for  $P(a)$ , this divergence is proportional to

$$\text{Str}_{M=0} (\mathbb{A}^{(2)} + \mathbb{B}^{(2)}) = -2 \text{Str}_{M=0} \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2. \quad (3.128)$$

However, in theories for which this quantity vanishes, we find that only  $\phi_2(\mu)$  survives, with  $\lim_{\mu \rightarrow 0} \phi_2(\mu) = \pi/3$ . In such cases we find from Eqs. (3.114) and (3.127) that

$$\begin{aligned} \lim_{\mu \rightarrow 0} \hat{\Delta}_G(\mu) &= \frac{\pi}{3} \text{Str}(\mathbb{X}_1 + \tau_2 \mathbb{X}_2) \\ &= \frac{\xi}{6} \text{Str} \bar{Q}_H^2 - \frac{2\pi}{3} \left( \text{Str} \tau_2 \bar{Q}_H^2 Q_G^2 - \frac{1}{12} \text{Str}_E \tau_2 Q_G^2 \right). \end{aligned} \quad (3.129)$$

The fact that  $\hat{\Delta}_G(\mu)$  asymptotes to a constant as  $\mu \rightarrow 0$  is not particularly surprising, given the assumed vanishing of the quantity in Eq. (3.128). However, what *is* surprising is that the asymptotic value in Eq. (3.129) receives contributions not only from the light or massless string states, but from the *entire tower of string states*, all the way up into the UV. Indeed, *all* of the string states contribute to the unrestricted supertraces in Eq. (3.129). This is a graphic example of the UV/IR mixing inherent in a modular-invariant theory such as string theory.

Let us now consider the behavior of  $\hat{\Delta}_G(\mu)$  as we proceed upwards in energy scale  $\mu$  away from the deep-IR limit. Indeed, much of the following discussion mirrors the discussion for the Higgs mass in Ref. [2], to which we refer the reader for details not provided here. Let us first focus on energy scales for which  $\mu \ll M_s$ . In this regime, we find that  $\phi_1$  and  $\phi_4$  continue to remain vanishingly small. However, whether  $\phi_5$  remains small as well for a particular state of mass  $M$  depends on the value of  $z \sim M/\mu$ —i.e., on whether the state whose contribution we are assessing is heavier or lighter than  $\mu$ . In this connection it is important to realize that our supertraces receive contributions from the entire string spectrum. This necessarily includes states with masses  $M \gtrsim M_s$ , but may also include potentially light states with nonzero masses far below  $M_s$ . The existence of such light states depends on our string construction and on the specific string model in question. Thus, even though we are considering situations in which  $\mu \ll M_s$ , there need not be any fixed hierarchical relationship between  $\mu$  and  $M$ .

In Fig. 4, we have plotted  $\phi_5(M, \mu)$  as a function of  $\mu/M$ . Recalling that this is the contribution to  $\hat{\Delta}_G(\mu)$  per unit  $\mathbb{A}^{(2)}$  charge from a given physical bosonic string state of mass  $M$ , certain aspects of this behavior are easy to understand. For example, when  $\mu \ll M$  the state is much heavier than the relevant energy scale  $\mu$  and is effectively “integrated out” of our theory. Thus all running stops, and  $\phi_5(M, \mu)$  becomes flat. Mathematically, this occurs because

$$\mathcal{K}_\nu^{(n,p)}(z) \sim \sqrt{\frac{\pi\rho}{2}} z^{n-1/2} e^{-z/\rho} \quad \text{as } z \rightarrow \infty. \quad (3.130)$$

Thus all running is exponentially suppressed as  $z \sim M/\mu \rightarrow \infty$ , leaving behind only an exponential tail. By contrast, for energy scales  $\mu \gg M$ , our state is still dynamical. We then see from Fig. 4 that the effective contribution to the running from this state is effectively *logarithmic*. Indeed, as  $z \rightarrow 0$ , we find that [22]

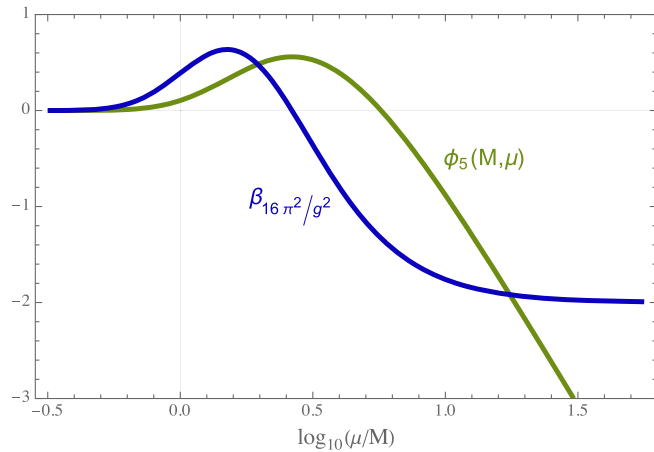


FIG. 4. The function  $\phi_5(M, \mu)$  in Eq. (3.124) and the corresponding contribution to the beta function  $\beta_\Delta$  in Eq. (3.137), plotted as functions of  $\log_{10}(\mu/M)$  (green and blue, respectively). Note that  $\phi_5(M, \mu)$  is the Bessel-function contribution to  $\hat{\Delta}_G(\mu)$  per unit  $\mathbb{A}^{(2)}$  charge from a given physical bosonic string state of nonzero mass  $M$ . When  $\mu \ll M$ , the state is effectively integrated out of the theory, whereupon the running contribution  $\phi_5(M, \mu)$  asymptotes to a constant. However, at larger energy scales  $\mu \gg M$ , the state is fully dynamical and produces a running which is effectively logarithmic. Finally, within the intermediate  $\mu \sim M$  region, the Bessel-function expression for  $\phi_5(M, \mu)$  in Eq. (3.124) provides a smooth connection between these two asymptotic behaviors and even gives rise to a transient “hump” in the value of  $\hat{\Delta}_G(\mu)$ , or equivalently a “dip” in the value of the running coupling  $g_G(\mu)$ .

$$\begin{aligned} \mathcal{K}_0^{(0,1)}(z) &\sim -\frac{1}{2} \log z + \frac{1}{2} [\log(2\pi) - \gamma], \\ \mathcal{K}_1^{(1,2)}(z) &\sim 1, \end{aligned} \quad (3.131)$$

where  $\gamma$  is the Euler-Mascheroni constant. For  $\mu \gg M$ , this leads to an asymptotic logarithmic running for  $\phi_5(M, \mu)$  of the form

$$\phi_5(M, \mu) \approx -2 \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma+1)} \frac{\mu}{M} \right]. \quad (3.132)$$

Finally, between these two extremes, we see that  $\phi_5(M, \mu)$  interpolates smoothly and even gives rise to a transient “hump” in  $\hat{\Delta}_G(\mu)$ , or equivalently a “dip” in  $g_G$ . This behavior results from the specific combination of Bessel functions within  $\phi_5(M, \mu)$ . Of course, the statistics factor  $(-1)^F$  within the supertrace flips the sign of this contribution for degrees of freedom which are fermionic.

Likewise, for any fixed scale  $\mu$ , adjusting the mass  $M$  upwards or downwards simply corresponds to shifting this curve rigidly to the right or left, respectively. In this way one can imagine summing over all such contributions to the running (while also weighting each contribution by the appropriate net numbers of states at each mass level) as one

takes the supertrace over the entire string spectrum. Of course, for any energy scale  $\mu$ , the contributions from states with  $M \gg \mu$  are exponentially suppressed, as discussed above. Thus, at any energy scale  $\mu$ , the only states which contribute meaningfully to  $\hat{\Delta}_G(\mu)$  are those with  $M \lesssim \mu$ .

Thus, combining these Bessel-function contributions with those from Eq. (3.129) and keeping only those (leading) terms which dominate when  $M \lesssim \mu \ll M_s$ , we see that we can approximate the exact result in Eq. (3.122) as

$$\begin{aligned} \hat{\Delta}_G(\mu) &\approx \frac{\pi}{3} \text{Str}(\mathbb{X}_1 + \tau_2 \mathbb{X}_2) \\ &\quad + 2 \text{Str}_{M=0}(\mathbb{A}^{(2)} + \mathbb{B}^{(2)}) \log \left( \frac{2\sqrt{2}eM_s}{\mu} \right) \\ &\quad - 2 \text{Str}_{0 < M \lesssim \mu}(\mathbb{A}^{(2)} + \mathbb{B}^{(2)}) \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma+1)} \frac{\mu}{M} \right] \end{aligned} \quad (3.133)$$

or equivalently

$$\begin{aligned} \hat{\Delta}_G(\mu) &\approx \frac{\xi}{6} \text{Str} \bar{Q}_H^2 \\ &\quad - \frac{2\pi}{3} \left( \text{Str} \tau_2 \bar{Q}_H^2 Q_G^2 - \frac{1}{12} \text{Str}_E \tau_2 Q_G^2 \right) \\ &\quad - 4 \text{Str}_{M=0} \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2 \log \left( \frac{2\sqrt{2}eM_s}{\mu} \right) \\ &\quad + 4 \text{Str}_{0 < M \lesssim \mu} \left( \bar{Q}_H^2 - \frac{1}{12} \right) Q_G^2 \log \left[ \frac{1}{\sqrt{2}} e^{-(\gamma+1)} \frac{\mu}{M} \right]. \end{aligned} \quad (3.134)$$

Given these results, our gauge couplings  $g_G(\mu)$  can exhibit a variety of running behaviors. These will ultimately depend on the spectrum of states associated with the string model under study. Of course, the final terms in Eqs. (3.133) and (3.134) do not exhibit any running until we reach  $\mu \sim M_{\text{lightest}}$ , where  $M_{\text{lightest}}$  is the mass of the lightest massive string state carrying a nonzero  $(\mathbb{A}^{(2)} + \mathbb{B}^{(2)})$  charge. Therefore, if we first restrict our attention to energy scales  $\mu \lesssim M_{\text{lightest}}$ , the only running that arises is due to the massless  $(\mathbb{A}^{(2)} + \mathbb{B}^{(2)})$ -charged states. These are the contributions that appear on the second and third lines of Eqs. (3.133) and (3.134), respectively.

This running can be expressed in a manner which is more traditional for describing the running of gauge couplings in string theory, namely in terms of an RGE that relates the couplings  $g_G(\mu)$  to their values at the string scale  $M_s$  (see, e.g., Refs. [13,14,23,24]). From Eq. (3.2), we obtain the general running equation for the total gauge couplings  $g_G(\mu)$ :

$$\frac{16\pi^2}{g^2(\mu)} - \frac{16\pi^2}{g_{\text{tree}}^2} = \hat{\Delta}_G(\mu) \quad (3.135)$$

where from Eq. (3.134) we can write

$$\begin{aligned} \hat{\Delta}_G(\mu) &= \hat{\Delta}_G(M_s) - 2 \text{Str}_{M=0}(\mathbb{A}^{(2)} + \mathbb{B}^{(2)}) \log\left(\frac{\mu}{M_s}\right) \\ &\equiv \hat{\Delta}_G(M_s) + \beta_\Delta(0) \log\left(\frac{\mu}{M_s}\right). \end{aligned} \quad (3.136)$$

Here the quantity  $\beta_\Delta(0)$  may be regarded as the  $\mu = 0$  value of the general beta function  $\beta_\Delta(\mu)$  for  $\hat{\Delta}_G(\mu)$ , which in turn is defined as

$$\beta_\Delta(\mu) \equiv \frac{\partial \hat{\Delta}_G(\mu)}{\partial \log \mu} = -\frac{32\pi^2}{g_G^3} \beta_g(\mu) \quad (3.137)$$

where  $\beta_g(\mu) \equiv \partial g_G / \partial \log \mu$  is the usual beta function for the gauge coupling  $g_G$ . Indeed, we see from Eq. (3.136) that  $\beta_\Delta(0)$  is precisely  $-2$  times the quantity given in Eq. (3.128), i.e.,

$$\begin{aligned} \beta_\Delta(0) &= -2 \text{Str}_{M=0}(\mathbb{A}^{(2)} + \mathbb{B}^{(2)}) \\ &= 4 \text{Str}_{M=0} \left( \bar{Q}_H^2 - \frac{1}{12} \right) \mathcal{Q}_G^2. \end{aligned} \quad (3.138)$$

Likewise, using an asterisk “\*” to indicate the couplings that would have arisen in our theory if  $\beta_\Delta(0)$  had vanished, we can write

$$\hat{\Delta}_G(M_s) = \hat{\Delta}_G^*(0) + \kappa \quad (3.139)$$

where  $\hat{\Delta}_G^*(0)$  is given in Eq. (3.129) and where [within our regulator scheme defined by our regulator function  $\hat{\mathcal{G}}_\rho(\mu)$  with  $\rho = 2$ ] we have

$$\kappa = -\beta_\Delta(0)[1 + \log(2\sqrt{2})]. \quad (3.140)$$

Thus, putting the pieces together, we have the RGE

$$\frac{16\pi^2}{g^2(\mu)} - \frac{16\pi^2}{g_{\text{tree}}^2} = \hat{\Delta}_G(M_s) + \beta_\Delta(0) \log\left(\frac{\mu}{M_s}\right) \quad (3.141)$$

where  $\hat{\Delta}_G(M_s)$  is given in Eq. (3.139).

Thus far, Eq. (3.141) captures the effects of the massless  $(\mathbb{A}^{(2)} + \mathbb{B}^{(2)})$ -charged string states. However, as  $\mu$  increases still further, additional  $(\mathbb{A}^{(2)} + \mathbb{B}^{(2)})$ -charged string states enter the EFT and contribute their own individual logarithmic contributions. Of course, if these additional states have masses  $M \gg M_{\text{lightest}}$ , the logarithmic nature of the running shown in Fig. 4 from the state with mass  $M_{\text{lightest}}$  will survive intact until  $\mu \sim M$ . However, if the spectrum of

states is relatively dense beyond  $M_{\text{lightest}}$ , the logarithmic contributions from each of these states must be added together, leading to a far richer behavior.

One interesting possibility arises in cases of string theories with large compactification radii  $R \gg M_s^{-1}$ . In such cases, our theory will have Kaluza-Klein (KK) modes with masses  $m_n \sim n/R$  that appear well below  $M_s$ . Thus, as  $\mu$  increases towards  $M_s$ , increasingly many KK states enter the EFT. Although each KK state contributes the same logarithmic running, our natural field-theoretic expectation is that the full supertrace over the string spectrum will begin to experience an accumulated effective *power-law* growth, with  $\hat{\Delta}_G(\mu) \sim \mu^\delta$  where  $\delta$  is the number of spacetime dimensions whose inverse compactification radii  $R^{-1}$  lie below  $\mu$ . Indeed, this is precisely the field-theoretic behavior discussed in Refs. [25,26], which can algebraically be interpreted as resulting from a beta function  $\beta_\Delta(\mu)$  which itself is growing polynomially with  $\mu$ . However, as we shall shortly see, in a string context we also have a scale-duality symmetry under  $\mu \rightarrow M_s^2/\mu$ . This means that even at energy scales  $\mu \ll M_s$  the *winding* modes associated with such compactifications can also contribute. Remarkably, these have the generic effect of *cancelling* this power-law running (and even the original logarithmic running), thereby producing a situation in which there can be no running at all. We will refer to the region in which such running terminates as a “fixed-point region.” This nonrenormalization phenomenon, surprising as it is, is actually quite general and will be discussed in detail in Ref. [21].

Thus far our results are valid for energy scales below the string scale. However, as mentioned above, it turns out that we also have information about what happens in the opposite region, namely that with  $\mu > M_s$ : *we simply enter a “dual” infrared region in which this same behavior again emerges, but in reverse.* This is a direct consequence of the modular invariance which we have been careful to maintain throughout our calculations. Indeed, modular invariance ensures that the running is symmetric under the *scale-inversion* duality transformation

$$\mu \rightarrow \frac{M_s^2}{\mu}. \quad (3.142)$$

As a result, when plotted as a function of  $\log(\mu/M_s)$ , the behavior of  $\hat{\Delta}_G(\mu)$  for  $\mu \ll M_s$  is reflected symmetrically through the self-dual point  $\mu_* = M_s$  to yield the reverse behavior for  $\mu \gg M_s$ .

As discussed in Ref. [2], the origins of this scale-duality symmetry are easily understood. To see this, we note that in general the contribution of a string state of mass  $M$  to the one-loop partition function experiences a Boltzmann-like suppression factor  $\sim e^{-\pi\alpha' M^2 \tau_2}$ . Thus, for any particular benchmark value  $\tau_2 = t$ , we can separate our string spectrum into two groups: “heavy” states whose

Boltzmann suppressions at  $\tau_2 = t$  are significant according to some convention, and whose contributions therefore do not require regularization, and “light” states whose Boltzmann suppressions are not significant, and whose contributions therefore require regularization. On this basis, with an eye towards interpreting these results in terms of an EFT with a running scale  $\mu$ , we are directly led to identify  $\mu^2$  inversely with  $t$ . However, modular invariance tells us that any physical quantities which depend on  $\tau$  must be invariant under  $\tau \rightarrow -1/\tau$ . Along the  $\tau_1 = 0$  axis, this becomes an invariance under  $\tau_2 \rightarrow 1/\tau_2$ . This then immediately implies an invariance under  $t \rightarrow 1/t$ , or equivalently under  $\mu \rightarrow \mu_*^2/\mu$  where  $\mu_*$  is an arbitrary self-dual mass scale. Of course, the choice of normalization for  $\mu$  in relation to  $t$  is purely a matter of convention, since the former is a dimensionful spacetime quantity while the latter is a dimensionless worldsheet quantity. In keeping with the traditional string-theoretic conventions relating worldsheet and spacetime physics, we take this conversion factor to be given by  $\alpha'$ . This then tells us that  $\mu_* = M_s$ .

Although this scale-duality symmetry follows directly from modular invariance, its implications are profound. *Ultimately, the existence of such a symmetry implies the existence of a fundamental limit on the extent to which an EFT perspective can possibly remain valid in string theory.* We have already noted in the Introduction that string theory is rife with duality symmetries which defy EFT notions: an immediate example of this is T-duality, under which the physics associated with a closed string propagating on a spacetime with a compactified dimension of radius  $R$  is indistinguishable from the physics associated with a closed string propagating on a spacetime with a compactified dimension of radius  $R' \equiv \alpha'/R$ . This is true as an exact symmetry not only for the string spectrum but also for all interactions. Thus such strings cannot distinguish between large and small compactification geometries, thereby preventing us from establishing a Wilsonian EFT-like ordering of length scales from large to small, or equivalently from IR to UV. Phrased somewhat differently, the existence of a T-duality symmetry tells us that there is a fundamental limit to which we may consider a spacetime compactification radius to be “small.” However, what we are seeing now is that a somewhat different phenomenon—namely the scale-duality symmetry under  $\mu \rightarrow M_s^2/\mu$  which is guaranteed by modular invariance—implies a fundamental limit on the extent to which our EFT can exhibit UV behavior. Indeed, pushing  $\mu$  beyond  $M_s$  only serves to reintroduce the original IR behavior of our theory—a behavior which we may now associate with the *dual* energy scale  $\mu' \equiv M_s^2/\mu$  associated with a “dual” EFT. In this sense, the energy scales near  $M_s$  exhibit the “most possible UV” behavior that can be realized.

At first glance, it may be tempting to associate this scale-duality symmetry with T-duality, since both tend to place limits on UV behavior and have similar algebraic forms.

We stress, however, that T-duality is a *spacetime* symmetry, and as such comes with certain assumptions about the spacetime geometry. By contrast, modular invariance is a fundamental *worldsheet* symmetry which is required for the self-consistency of the theory itself. As such, modular invariance and T-duality are unrelated. Indeed, T-duality relates two *a priori* different string theories to each other, one with a large compactification volume and the other with a small compactification volume, and maps a given state with KK and winding numbers  $(m, n)$  in the first theory to the *equally massive*  $(n, m)$  state in the other. By contrast, modular invariance is a symmetry that operates *within* a single string theory and involves Poisson resummations across the *entire* string spectrum simultaneously. As such, no two string states in the theory are directly related to each other under modular transformations. Indeed, only through such nontrivial resummations involving the *entire* string spectrum—including the oscillator states as well—could we ever hope to obtain features such as misaligned supersymmetry and supertrace constraints (such as  $\text{Str } \mathbf{1} = 0$ ) that simultaneously balance all of our string states at all energy scales against each other within a single string theory, even without supersymmetry.

Taken together, all of these observations lead to a running for  $\hat{\Delta}_G(\mu)$  as sketched in Fig. 5. In the deep IR,  $\hat{\Delta}_G(\mu)$  approaches a constant unless the quantity in Eq. (3.128) is nonzero. As  $\mu$  increases, our theory then passes through a “dip” region and a subsequent EFT region characterized by logarithmic running. If our theory has large compactification radii  $R \gg M_s^{-1}$ , the contributions from the corresponding Kaluza-Klein and winding states can then conspire to eliminate this running, leading to the existence of a higher-dimensional string-scale fixed-point region. Beyond  $M_s$ , our theory enters a “dual” regime in which further increases in  $\mu$  only reproduce the IR behavior we have already seen, only in reverse.

We conclude this section with three comments. First, we observe that the running of the gauge coupling is essentially the same as the running of the Higgs mass in Ref. [2]—indeed for  $\mu \lesssim M_E$  the only differences are the *coefficients* of the different running terms. These coefficients change because they tally the appropriate charges of our states across the string spectrum, and the charges that are appropriate or relevant change when we switch our attention from the Higgs mass to the gauge couplings.

Our second comment concerns the running of the gauge couplings themselves. Within our calculations we have implicitly assumed that these couplings remain perturbative throughout the running shown in Fig. 5; otherwise our one-loop calculation is no longer applicable and higher-loop (and even nonperturbative) calculations would be needed. Depending on the relative signs and magnitudes of the various supertraces involved, these couplings could be in



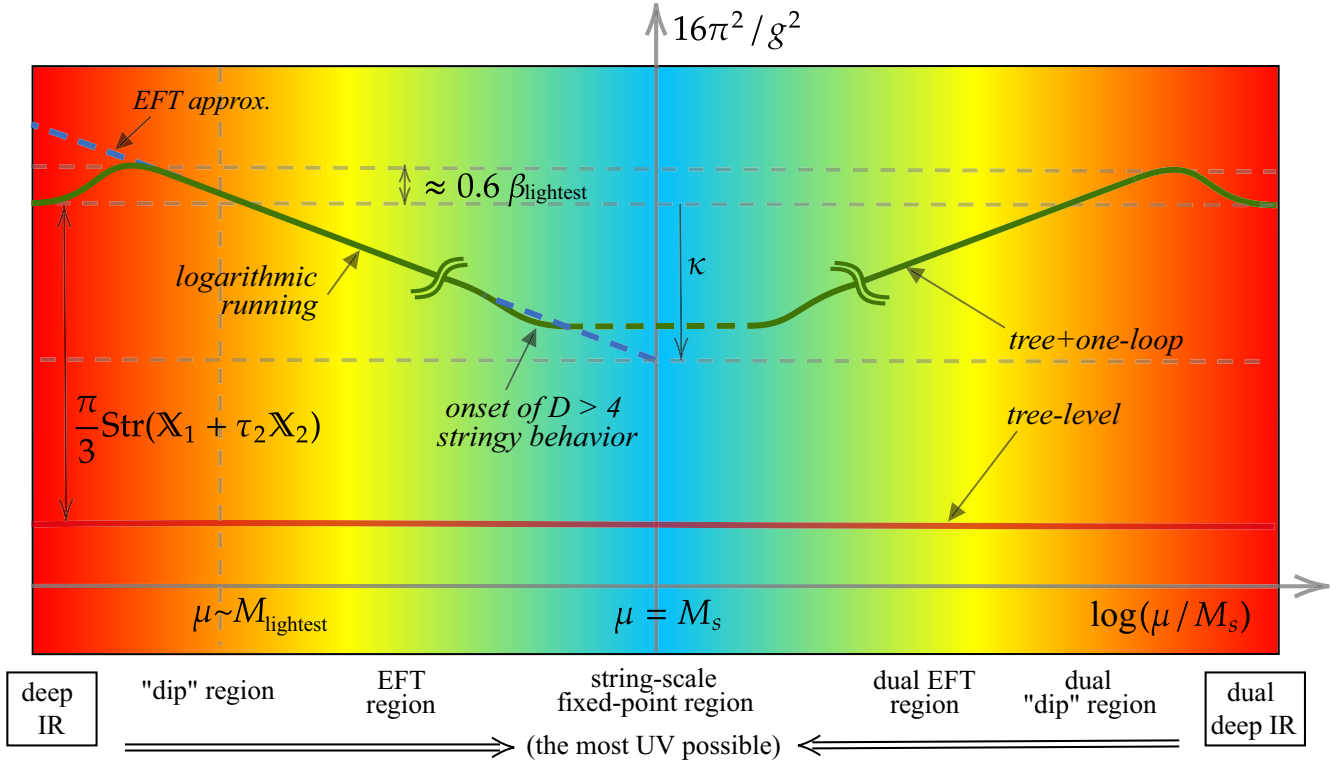


FIG. 5. The one-loop running of the inverse gauge coupling  $\hat{\Delta}_G = 16\pi^2/g_G^2$  for any gauge group  $G$ , as calculated from first principles in a fully modular-invariant string framework. The tree-level contribution is sketched in red, and the total one-loop coupling is sketched in green. In the deep IR, the coupling approaches an asymptotic value which receives contributions from *all* of the states in the string spectrum which carry nontrivial helicity  $\mathbb{A}^{(1)} \sim \bar{Q}_H^2$  charges. This assumes that our theory contains no net massless states charged under  $\mathbb{A}^{(2)} + \mathbb{B}^{(2)} \sim (\bar{Q}_H^2 - 1/12)Q_G^2$  where  $Q_G^2$  is the sum of the squares of the charges in the Cartan subalgebra of  $G$ ; otherwise  $\hat{\Delta}_G(\mu)$  diverges in the IR limit. Moving towards higher values of  $\mu$ , we see that a nontrivial scale dependence does not emerge until  $\mu \sim M_{\text{lightest}}$ , where  $M_{\text{lightest}}$  collectively represents the masses of the lightest massive states which are charged under  $\mathbb{A}^{(2)} + \mathbb{B}^{(2)}$ . The nonmonotonic “dip” in  $g_G$  (or “hump” in  $\hat{\Delta}_G$ ) that is observed in this region is a transient effect which smoothly connects the asymptotic deep-IR region  $\mu \ll M_{\text{lightest}}$  to an EFT-like region  $M_{\text{lightest}} \lesssim \mu \ll M_s$ . Beyond the dip region, the theory then enters an EFT-like region in which the gauge coupling experiences a logarithmic running. As  $\mu \rightarrow M_s$ , it is possible that we might cross the energy threshold  $R^{-1}$  associated with large compactification radii. In such cases, this logarithmic running can be modified by the appearance of Kaluza-Klein and winding states which might appear at mass scales significantly below  $M_s$  and which might tend to *cancel* this logarithmic running, leading to the existence of a higher-dimensional fixed-point regime, as shown. The subtleties involved in this behavior will be discussed further in Ref. [21]. However, as a general principle, modular invariance requires that the running of  $\hat{\Delta}_G$  exhibit an invariance under  $\mu \rightarrow M_s^2/\mu$ . Thus, as  $\mu$  increases beyond  $M_s$ , the theory inevitably begins to reenter an IR-like regime which we may associate with a “dual” EFT, followed by a dual dip region and then a dual deep-IR region. The background colors of this sketch indicate the transition from the deep IR (red) to the UV (blue) and then back to IR (red). As such, there is a maximum degree to which our theory can approach the UV: once the energy scale  $\mu$  passes the self-dual point  $\mu \sim M_s$ , further increases in  $\mu$  only push us towards increasingly IR behavior. The quantity  $\kappa$  is defined in Eq. (3.139).

danger of becoming nonperturbative either as  $\mu \rightarrow M_s$  (which represents one extremum of the gauge-function plotted in Fig. 5) or within the “dip” region.

Most importantly, however, there is a deep and fundamental difference between the running of the gauge couplings  $\hat{\Delta}_G$  and the running of the Higgs mass in Ref. [2]. As we see directly from Eqs. (3.114), (3.117), and (3.122), the gauge-coupling calculation now includes contributions from off-shell string states for which  $M_L \neq M_R$ . This is a strange but not entirely unexpected feature: states which are not *physical* in the underlying

string theory, and which therefore can only contribute in string loop diagrams, also contribute to the running of the gauge couplings in the corresponding low-energy EFT. This feature did not appear in the running of the Higgs mass in Ref. [2]. However, as we have seen, this feature ultimately stems from the fact that the contributions to the Higgs mass are proportional not to the square of the helicity charge  $\bar{Q}_H^2$ , but rather to this quantity *minus* 1/12. In field theory, this extra  $-1/12$  is not problematic. However, in string theory it has deep repercussions because a pure number such as  $-1/12$  cannot be subtracted from a

squared-charge operator such as  $\bar{Q}_H^2$  because a pure number has modular weight  $k = 0$  while the squared-charge operator has modular weight  $k = 2$ . Modular invariance thus requires that the  $-1/12$  term be “completed” to the weight-two modular function  $\bar{E}_2/12$ , and this in turn has reverberations throughout the string spectrum, shifting left-moving string masses  $M_L$  relative to right-moving string masses  $M_R$ . This is why the nonlevel-matched string states now survive the Rankin-Selberg procedure and appear in our running calculation.

That said, these states do not contribute to the low-energy running in a standard way. Normally, we would expect a string state to contribute in the low-energy theory according to its mass  $M^2 = (M_L^2 + M_R^2)/2$ . Indeed, this quantity in some sense tells us how much worldsheet energy (as measured by eigenvalues of  $\bar{L}_0$  and  $L_0$  respectively) has been “invested” in creating that state as an excitation in the underlying worldsheet theory. However, what we are now learning from Eqs. (3.114), (3.117), and (3.122) is that although a given entwined string state may have a string-theoretic mass given by  $M$ , it contributes to the low-energy EFT precisely as if it had a mass simply given by  $M_L$ . In other words, the combined *string-theoretic mass*  $M$  is irrelevant; what matters—and what we may therefore consider to be the effective *EFT mass* in such theories, at least as far as the gauge couplings are concerned—is determined by  $M_L$  alone. This, of course, is the effect of the shift in left-moving masses relative to right-moving masses induced by  $\bar{E}_2$ .

We also observe that the entwined resonances all have left-moving masses that exceed the string scale:  $M_L \geq M_E = \sqrt{2}M_s$ . Thus, one might be tempted to argue that these states have no effects at energy scales below  $M_s$ . However, this would not be correct. Thanks to scale-inversion duality, any state that affects the running of quantities *above* the string scale will also affect the running of these quantities *below* the string scale. This is not a new phenomenon unique to the entwined states. After all, we have already seen that the behavior of our amplitudes in the deep IR is in part determined by the extremely heavy string states in the deep UV. In a similar way, the entwined states also have effects below the string scale and thereby also have an indirect role in affecting the low-energy EFT below the string scale.

Ultimately, this can be understood from the perspective of modular invariance and misaligned supersymmetry. In the two cases studied in Ref. [2] (namely the cosmological constant and the Higgs mass), the relevant supertraces involved only physical states and closed under modular transformations into themselves. Thus the corresponding amplitudes in each case were fully modular invariant and yet could be written purely in terms of supertraces over only physical string states. However, for the present gauge-coupling calculation, slightly different supertraces are involved. Amongst these supertraces, those involving only

physical string states do not close into each other under modular transformations. Rather, closure for these supertraces also involves certain entwined supertraces as well. This is ultimately why the entwinement occurs in these theories. Through the entwinement, the off-shell string states continue to make explicit contributions to the relevant amplitudes.

This last observation is in fact part of a more general lesson. In ordinary quantum field theory, one can meaningfully seek to identify the physical effects that arise due to the existence of specific states in the spectrum. For example, we might attempt to determine the energy scales at which a given state of mass  $M$  contributes to the running of a given quantity. However, in a modular-invariant theory, this question has no meaning because of UV/IR mixing. Every state within the spectrum is deeply connected to the states at all other energy scales. It is therefore impossible to uniquely isolate the contributions of a single state within the spectrum because there is no modular-invariant way to perform such an analysis.

As a dramatic example of this phenomenon, let us consider the unphysical  $(m, n) = (0, -1)$  proto-graviton states [19] that arise within all string models, and ask whether these states contribute to the corresponding one-loop cosmological constant  $\Lambda$ . On the one hand, we might claim that the proto-graviton states do not contribute to  $\Lambda$  because we know that we can write  $\Lambda$  as the supertrace of  $M^2$  over the purely physical states in the spectrum, as in Eq. (2.3). However, a direct calculation of the one-loop torus integral associated with  $\Lambda$  demonstrates that these states not only contribute, but actually provide contributions that dominate over those of all other states (see, e.g., Table 2 of Ref. [27]). The underlying reason for this apparent contradiction is that our question about which states contribute does not have a modular-invariant answer. Modular transformations allow us to reshuffle our contributions so that the effects of one state can be reinterpreted as the resummed effects of other states instead. Indeed, as we have already asserted, the unentwined supertraces that contribute to the gauge-coupling running do not close into themselves under modular transformations; they also involve the entwined supertraces. All supertraces—both entwined and unentwined—therefore contribute together in a modular-invariant way.

In this connection, we observe that the lightest entwined states actually have vanishing string-theoretic masses. These are the states that populate the  $(m, n) = (-1/2, +1/2)$  square in Fig. 3. Such states require only a minimal amount of energy to produce on the worldsheet—indeed, exactly the same amount of energy as required to produce the physical massless states that populate the  $(0, 0)$  square and presumably include the Standard-Model states. It will be interesting to explore the ramifications of this observation [20].

#### IV. CONCLUSIONS, DISCUSSION, AND FUTURE DIRECTIONS

In this paper we developed a general framework for analyzing the running of gauge couplings within closed string theories. Unlike previous discussions in the literature, our calculation fully incorporates the underlying modular invariance of the string and includes the contributions from the infinite towers of string states which are ultimately responsible for many of the properties for which string theory is famous, including its enhanced degree of finiteness and UV/IR mixing. In order to perform our calculations, we adopted a formalism [2] that was recently developed for calculations of the Higgs mass within such theories.

In general, this formalism—which builds upon the Rankin-Selberg technique [4,5] but which also includes additional critical features such as an identification between worldsheet parameters and an effective spacetime energy scale  $\mu$ —gives rise to an “on-shell” EFT description in which the final results are expressed in terms of supertraces over the physical string states, and in which these quantities exhibit an EFT-like “running” as a function of the scale  $\mu$ . We found, however, that the calculation of the gauge couplings differs in one deep way from that of the Higgs mass: while the latter results depend on purely on-shell supertraces, the former results have a different modular structure which causes them to depend on supertraces over *off-shell* string states as well. Indeed, as explained at the end of Sec. III D, the entwinement induced by the modular completion of the helicity operator needed for calculating the gauge couplings has “deformed” the notion of physicality for the string states, allowing states which are not level matched to nevertheless act as physical states which contribute to the physical supertraces describing the values of physical string amplitudes. We have also seen that although our results yield the expected logarithmic running of the gauge couplings within certain energy scales, they also yield a number of intrinsically stringy behaviors that transcend what might be expected within an effective field theory approach.

A central feature of our treatment is our use of a modular-invariant regulator  $\hat{G}_\rho$  to define a physical energy scale  $\mu$  in the system and simultaneously eliminate any logarithmic divergences that might arise from the contributions of certain massless states in the theory. Since this regulator is modular invariant, it suppresses the contributions from the lighter string states in a smooth manner which is consistent with UV/IR mixing and which therefore naturally incorporates the contributions from the infinite towers of string states in a modular-invariant way. Use of this regulator not only eliminates logarithmic IR divergences but more importantly but also allows us to study how the gauge couplings “run” as a function of the spacetime energy scale  $\mu$ . Indeed this procedure can be thought of as a “functional renormalization group” (FRG) approach to

scaling [28–30] for UV-complete modular-invariant theories. Notably, the use of such a regulator allows us to sidestep the need to introduce a sharp cutoff which would be very difficult if not impossible to reconcile with our fundamental UV/IR-mixed string symmetries. This also allows us to avoid the need to designate which states are “light” with respect to  $\mu$  (and which therefore contribute to the running), and which are “heavy” (and therefore do not). Through such modular-invariant regulator functions, we can develop a notion of “running” gauge couplings and beta functions, with our modular-invariant regulator allowing us to extract this apparent EFT-like behavior as a function of the spacetime energy scale.

The final global picture that emerges is shown in Fig. 5. Perhaps the most prominent feature is the existence of a scale-duality symmetry, i.e., an invariance under  $\mu \rightarrow M_s^2/\mu$ . As we have discussed, this is an inevitable consequence of the modular invariance that underlies our calculations. However, the impact of this scale-duality symmetry is felt even at energy scales below the string scale. For example, we have seen that the IR value of the one-loop contribution to the gauge coupling is given by  $\frac{\pi}{3} \text{Str}(\mathbb{X}_1 + \tau_2 \mathbb{X}_2)$  where the supertrace is over *all* of the states in the theory, regardless of their masses. From a naïve field-theory perspective, such a supertrace would appear to control a quadratic UV divergence. However, the deep IR in such theories is also equivalent to the deep UV, where one would expect all of the states to play a role. Indeed, these IR predictions are RG invariants, in the sense that they define fixed-point values. These predictions are also independent of our choice of regulator function.

As we move away from the  $\mu \rightarrow 0$  limit and proceed towards higher energy scales, the system evolves away from this asymptotic IR behavior. Once  $\mu$  exceeds the masses of the lowest-lying states, the one-loop contribution to the gauge couplings passes through a localized “dip” and then begins to experience a nontrivial logarithmic running that can be associated with an EFT-like description. This running then continues towards higher energy scales, possibly passing through a sequence of EFT-like descriptions. If our string compactification geometry has effective radii  $R_i^{-1} \ll M_s$ , then this running continues until  $\mu \sim R_i^{-1}$ . Above this scale we find a surprising new behavior: the running stops, with the gauge coupling entering a string-scale “fixed-point” region. This surprising behavior will be discussed in more detail in Ref. [21] and ultimately results from the combined effects of both KK states and winding states. It may seem strange that both kinds of states should be playing a role at scales  $\mu \ll M_s$ , but this is a direct consequence of the scale-duality symmetry under  $\mu \rightarrow M_s^2/\mu$ . The fact that both KK and winding states are simultaneously playing important roles further implies that the behavior in this region is not only fully higher-dimensional but also intrinsically stringy. This transition to entirely stringy behavior is an inevitable and profound

consequence of an RG procedure which is consistent with modular invariance: a modular-invariant regulator cannot distinguish between KK and winding modes and therefore can only act to suppress the contributions of both or neither.

Beyond  $\mu \sim M_s$ , we enter the “dual” phase in which the running of the gauge coupling is inverted. This inversion of the running is quite remarkable. Indeed, from a naïve field theoretic perspective, this kind of complete reversal would be disallowed (by e.g., the  $a$ -theorem). Of course, we do not expect such a theorem to hold in a UV/IR-mixed theory such as string theory.

Indeed this inversion of the running of the gauge couplings can best be understood by recognizing that the fundamental degrees of freedom within the dual phase of the theory are not those of the original low-energy theory. The original theory and its dual are nothing but modular transformations of each other—indeed, the relationship between these two “phases” of the theory is outlined in Fig. 4 of Ref. [2], where they are shown to lie along different but equivalent “spokes” of the same fundamental diagram. Thus each phase carries the same information and can be viewed as representing the same fundamental theory, consistent with the idea that modular symmetries (like gauge symmetries) do not relate physically different theories to each other, but rather represent redundancies of description. However, under modular transformations, the states within our string theory at all mass levels are nontrivially mixed with each other. Thus, a given state in the original theory is mapped to a highly nontrivial combination of states in the dual theory, while each state in the dual theory is likewise mapped to a highly nontrivial combination of states in the original theory. Demanding that this mapping nevertheless produce the same theory is the essence of what it means for a theory to be modular invariant. More explicitly, as discussed in Ref. [2], the scale-duality map  $\mu \rightarrow M_s^2/\mu$  is intimately related to the  $\tau \rightarrow -1/\tau$  modular transformation evaluated along the  $\tau_1 = 0$  line. This modular transformation induces a Poisson resummation amongst the states of the original theory, so that the degrees of freedom in the dual phase of the theory with  $\mu \geq M_s$  are Poisson-resummed versions of the degrees of freedom in the original phase of the theory with  $\mu < M_s$ . It is ultimately this Poisson resummation which is responsible for the inversion in the running of the gauge couplings once we cross between the  $\mu < M_s$  and  $\mu > M_s$  regions in any modular-invariant theory.

Although this physical picture is relatively simple, it actually encapsulates a considerable amount of nontrivial physics. As we would expect in any UV/IR-mixed theory, string modes that are light are being mixed with those which are heavy. However, such heavy string states are super-Planckian, and may (depending on the string coupling) include black holes. They are also likely to include so-called “long” strings, i.e., strings with large numbers of

oscillator excitations. However, as long as we maintain a constant definition of the physical spacetime energy scale  $\mu$ , modular invariance requires that these states all conspire (through Poisson resummations) to achieve this apparent reversal in the directionality of the gauge-coupling running at  $\mu = M_s$ .

It would be an interesting exercise to develop an understanding of the dual running directly in terms of these dual degrees of freedom. Moreover, although we have concentrated in this paper on the one-loop running of the gauge couplings, we expect similar results to apply to other one-loop amplitudes, such as might be involved in string scattering. Here too one must identify a physical string scale  $\mu$  in terms of certain renormalization conditions and then study how such amplitudes depend on  $\mu$ .

Above the string scale any alternative RG prescription derived from such amplitudes—e.g., a prescription based on suppressing the contributions of certain momentum modes—should therefore be defined in terms of *dual momenta* for those asymptotic eigenstates that can be prepared in this regime. These would comprise the long-string modes discussed above. In this way the  $\mu \rightarrow M_s^2/\mu$  symmetry would be faithfully respected. It is for this reason that our worldsheet regulator prescription gives a correct physical picture of renormalization, one which is applicable for all values of  $\mu$ . Indeed, such a regulator prescription is also correctly aligned with both cosmological [31] and thermal [32,33] dualities. We leave such investigations for future work.

Given these observations, we now discuss possible new approaches to hierarchy problems. We begin by recalling from Sec. II that within the string context there is no notion of a spacetime energy scale  $\mu$  before we insert a regulator and identify  $\mu$  in terms of the parameters of this regulator. Thus, the unregulated modular integrals that govern the couplings of the theory—integrals such as  $\Delta_G$  or equivalently  $\hat{\Delta}_G(\mu = 0)$ —are simply one-loop contributions to the effective action. Such integrals are either logarithmically divergent (if the string spectrum contains a nonzero net number of exactly massless  $\mathbb{X}_2$ -charged states), or finite otherwise. Thus, upon introducing the modular-invariant regulator  $\hat{\mathcal{G}}_\rho(a, \tau)$ , it is inevitable that the values of the couplings in the deep IR (i.e., as  $\mu \rightarrow 0$ ) are also the “bare” couplings that we would expect to obtain as  $\mu \rightarrow \infty$ . To see this, we note that within the FRG approach the “average effective action” is normally taken to interpolate between the effective action at  $\mu = 0$  and the “bare” action at  $\mu \rightarrow \infty$ . However, within our UV/IR-mixed context, the effective actions that one obtains in the  $\mu \rightarrow 0$  and  $\mu \rightarrow \infty$  limits are one and the same. This is because the modular invariance of our regulator implies that when we are regulating the IR, we are also equivalently regulating the UV. Indeed, any divergence that would arise within the  $\tau \rightarrow i\infty$  region of the modular integral (and which therefore would normally be interpreted as an IR divergence) is the same as the divergence that would arise within the

$\tau \rightarrow 0$  region of the modular integral (and which would therefore be interpreted as a UV divergence).

This relation between UV and IR divergences has an important implication. By locking these two types of divergences together, our theory cannot exhibit any UV divergence that is not also present as an IR divergence. However, as we have seen, the  $\tau \rightarrow i\infty$  limit of our theory can support at most a logarithmic IR divergence arising from the contributions of certain massless states. This is an expected divergence that indicates that in the low-energy regime the theory will behave like a quantum field theory. However, modular invariance then implies that we cannot have quadratic divergences as  $\tau \rightarrow 0$  (i.e., in the UV). Indeed, in any modular-invariant theory, stringy “miracles” (such as the cancellation of certain supertraces) have no choice but to eliminate the quadratic divergences because there is simply no place left for them. In this connection, we note that this argument relies directly on modular invariance alone, and is not specific to the specific form chosen for our regulator so long as it is modular invariant.

This suggests that there is a fundamental difference between hierarchy problems in field theory and hierarchy problems in string theory (and in UV/IR mixed theories more generally). In order to analyze hierarchy problems within a Wilsonian field theory, one starts by separating operators into “relevant” ones that grow in the IR and “irrelevant” ones that grow in the UV. Relevant operators typically begin at (possibly Gaussian) fixed points in the UV. Their RG trajectories are then determined by a set of “unpredicted” free parameters that are chosen by hand in order to make the associated couplings agree with the desired (presumably measured) IR values. Meanwhile the irrelevant operators flow to attractive fixed points in the IR, thereby becoming “predictions” of the theory. In this language, hierarchy problems arise whenever there is an extreme sensitivity of the relevant operators (which control the RG trajectory) to the intermediate physics. Such sensitivity confronts us with what is essentially a “shooting problem” because it requires us to keep fine-tuning our RG trajectory in order to hit the desired IR values. However the UV/IR mixing in string theory (and its attendant  $\mu \rightarrow M_s^2/\mu$  symmetry) removes the underpinnings of this entire picture, as operators cannot even be designated as “relevant” or “irrelevant” until we decide which direction corresponds to “UV” and which corresponds to “IR” within our definition of the energy scale. In this context we refer the reader to Fig. 4 in Ref. [2], which graphically demonstrates the different possibilities. Indeed, the only reliable quantities before we make this choice are the values of supertraces over the infinite towers of states. As we have mentioned, these are by definition invariant under our choice of regulator and also invariant under changes in the relevant energy scale. As such, they remain invariant within the emergent EFT, simultaneously determining from the outset both the bare action and

the effective action to which the theory must flow in the deep IR regardless of what intermediate physics may exist.

We find these observations to be compelling foundations for future, more general, phenomenological studies. In particular, the transformation of apparent hierarchy problems into statements about the properties of supertraces over infinite towers of physical string states suggests that seemingly miraculous cancellations and “magic zeros” are unavoidable features of the effective field theory stemming from a natural UV/IR-mixed modular-invariant theory. In such a framework, the solutions to hierarchy problems such as the gauge-hierarchy problem and the cosmological-constant problem rely on conspiracies between physics at all energy scales simultaneously, and would thus be essentially invisible to low-energy observers. In this sense, they might be considered to exhibit what has recently been dubbed “neutral naturalness,” except in a form that does not involve pairwise cancellation mechanisms that operate scale-by-scale but rather through seemingly miraculous cancellations that operate at all scales simultaneously. Within such frameworks, retaining the full spectrum of states within our calculations is therefore critical for obtaining a full understanding of naturalness.<sup>1</sup>

Given these comments, there exist many promising avenues for future research targeting the development of a fuller understanding of naturalness in UV/IR-mixed theories. Indeed, one of these is to study the manner in which the running of the gauge couplings is deformed in the presence of large extra dimensions. For example, following Ref. [35], we might study the running of the gauge couplings when our theory has a compactification geometry of the form  $\mathbb{K} \times \mathbb{T}^2$  where  $\mathbb{K}$  has a characteristic volume near the string scale and where  $\mathbb{T}^2$  indicates a two-torus with radii  $R_i \gg M_s^{-1}$ . The results of this analysis will be presented in Ref. [21], where we shall find that the promotion of logarithmic running to power-law running—as expected from a field-theoretic analysis [25,26]—does not occur in string theory. Indeed, as far as power-law running is concerned, this is a kind of “nonrenormalization” theorem for string theory. As we shall see in Ref. [21], this cancellation of power-law running is the result of a delicate conspiracy between modular invariance and the manner in which extra spacetime dimensions emerge in string theory as their radii become large. Indeed, this cancellation ultimately reflects a subtle entanglement between the properties of renormalization in higher dimensions and the requirements of modular invariance. Moreover, as we shall demonstrate in Refs. [17,21], there are also many additional supertrace relations which govern the spectra of modular-invariant string theories with large compactification radii—supertrace relations whose role is to enforce these

<sup>1</sup>This point of view regarding modular symmetries addressing hierarchy problems was first advocated in Ref. [1]. More recent discussions along these lines can be found in Refs. [2,34].

remarkable cancellations. Such supertrace relations are thus also responsible for the finiteness properties of these theories.

The results shown in Fig. 5 also call for a reappraisal of the possibilities for gauge-coupling unification in modular-invariant theories. Although the gauge couplings continue to exhibit logarithmic running—thereby suggesting that a traditional logarithmic unification may continue to be viable—the existence of a possible fixed-point regime near the string scale has the potential to alter this situation. This is especially true given that the scale of unification typically assumed for heterotic strings is only one or two orders of magnitude below the string scale [3,14]. It is even possible that the existence of this fixed-point regime may serve to reconcile the long-standing discrepancy [14] between the string-predicted perturbative heterotic unification scale and the traditional GUT scale extrapolated from experimental measurements of the Standard-Model gauge couplings at low energies.

Overall, the message seems clear. Hierarchy problems (and even issues related to gauge-coupling unification) assume traditional field-theory relationships between UV

and IR physics. By contrast, string theory tells us that we have UV/IR mixing, misaligned supersymmetry, softened divergences (even finiteness), scale duality, and so forth. Thus, within the context of string theory, hierarchy problems may not be fundamental or survive in the manner we normally assume.

## ACKNOWLEDGMENTS

The research activities of S. A. A. were supported by the STFC grant ST/P001246/1, and the Institut Pascal at Université Paris-Saclay with the support of the P2IO Laboratory of Excellence, the P2I axis of the Graduate School Physics of Université Paris-Saclay, and the IN2P3 master project UCMN. The research activities of K. R. D. were supported in part by the U.S. Department of Energy under Grant DE-FG02-13ER41976/DE-SC0009913, and also by the U.S. National Science Foundation through its employee IR/D program. The opinions and conclusions expressed herein are those of the authors, and do not represent any funding agencies.

- 
- [1] K. R. Dienes, *Nucl. Phys.* **B611**, 146 (2001).  
 [2] S. Abel and K. R. Dienes, *Phys. Rev. D* **104**, 126032 (2021).  
 [3] V. S. Kaplunovsky, *Nucl. Phys.* **B307**, 145 (1988); **B382**, 436(E) (1992).  
 [4] R. Rankin, *Proc. Cambridge Philos. Soc.* **35**, 351 (1939).  
 [5] R. Rankin, *Proc. Cambridge Philos. Soc.* **35**, 357 (1939).  
 [6] A. Selberg, *Arch. Math. Naturvid.* **43**, 47 (1940).  
 [7] J. Polchinski, *Commun. Math. Phys.* **104**, 37 (1986).  
 [8] K. R. Dienes, M. Moshe, and R. C. Myers, *Phys. Rev. Lett.* **74**, 4767 (1995).  
 [9] K. R. Dienes, *Nucl. Phys.* **B429**, 533 (1994).  
 [10] E. Kiritsis and C. Kounnas, *Nucl. Phys.* **B442**, 472 (1995).  
 [11] E. Kiritsis, C. Kounnas, P. M. Petropoulos, and J. Rizos, *Nucl. Phys.* **B483**, 141 (1997).  
 [12] E. Kiritsis, C. Kounnas, P. M. Petropoulos, and J. Rizos, *Nucl. Phys.* **B540**, 87 (1999).  
 [13] K. R. Dienes and A. E. Faraggi, *Nucl. Phys.* **B457**, 409 (1995).  
 [14] K. R. Dienes, *Phys. Rep.* **287**, 447 (1997).  
 [15] D. Zagier, *J. Fac. Sci. Tokyo* **28**, 415 (1982), <https://people.mpim-bonn.mpg.de/zagier/files/rankin-selberg/fulltext.pdf>.  
 [16] D. Kutasov and N. Seiberg, *Nucl. Phys.* **B358**, 600 (1991).  
 [17] S. Abel, K. R. Dienes, and L. A. Nutricati, “New constraints on string spectra: Vanishing supertraces without supersymmetry” (to be published).  
 [18] C. Angelantonj, M. Cardella, S. Elitzur, and E. Rabinovici, *J. High Energy Phys.* **02** (2011) 024.  
 [19] K. R. Dienes, *Phys. Rev. Lett.* **65**, 1979 (1990).  
 [20] S. Abel, K. R. Dienes, and L. A. Nutricati, “Probing the off-shell structure of string theory: Off-shell misaligned supersymmetry and off-shell supertraces” (to be published).  
 [21] S. Abel, K. R. Dienes, and L. A. Nutricati, “Nonrenormalization theorems without supersymmetry” (to be published).  
 [22] R. B. Paris, *Math. Eterna* **8**, 71 (2018), <https://www.longdom.org/abstract/the-evaluation-of-single-bessel-function-sums-4807.html>.  
 [23] K. R. Dienes and A. E. Faraggi, *Phys. Rev. Lett.* **75**, 2646 (1995).  
 [24] K. R. Dienes, A. E. Faraggi, and J. March-Russell, *Nucl. Phys.* **B467**, 44 (1996).  
 [25] K. R. Dienes, E. Dudas, and T. Gherghetta, *Phys. Lett. B* **436**, 55 (1998).  
 [26] K. R. Dienes, E. Dudas, and T. Gherghetta, *Nucl. Phys.* **B537**, 47 (1999).  
 [27] K. R. Dienes, *Phys. Rev. D* **73**, 106010 (2006).  
 [28] J. Polchinski, *Nucl. Phys.* **B231**, 269 (1984).  
 [29] C. Wetterich, *Phys. Lett. B* **301**, 90 (1993).  
 [30] T. R. Morris, *Int. J. Mod. Phys. A* **09**, 2411 (1994).  
 [31] R. H. Brandenberger and C. Vafa, *Nucl. Phys.* **B316**, 391 (1989).  
 [32] K. R. Dienes and M. Lennek, *Phys. Rev. D* **70**, 126005 (2004).  
 [33] K. R. Dienes and M. Lennek, *Phys. Rev. D* **70**, 126006 (2004).  
 [34] K. R. Dienes, talk given at CERN Theory Workshop, “Exotic Approaches to Naturalness,” (2023); slides available at [https://indico.cern.ch/event/1204192/contributions/5217557/attachments/2584637/4458574/dienes\\_cern.pdf](https://indico.cern.ch/event/1204192/contributions/5217557/attachments/2584637/4458574/dienes_cern.pdf).  
 [35] L. J. Dixon, V. Kaplunovsky, and J. Louis, *Nucl. Phys.* **B355**, 649 (1991).