Confining phase of superstrings and axionic strings

Soo-Jong Rey
Department of Physics, University of California, Santa Barbara, California 93106
(Received 5 October 1990)

In the four-dimensional effective field theory of the heterotic superstring compactified on a Calabi-Yau manifold, we find a new class of spacetime string instantons in the gravity sector that involves graviton, dilaton, and Kalb-Ramond (axion) fields. The instantons satisfy generalized (anti-)self-duality equations, and saturate a Bogomolnyi bound provided by the underlying supersymmetry. The instantons lead to several intriguing nonperturbative phenomena to low-energy string theory. A mass gap is generated to the Kalb-Ramond gauge fields, and results in the confinement of superstrings into Kalb-Ramond domain walls below the compactification scale. More interestingly, they provide a new example of the nonperturbative breakdown of the superpotential nonrenormalization theorem. Together with the world-sheet string instantons, this implies that all would-be Peccei-Quinn symmetries that arise in four-dimensional string theories are completely broken around the string compactification scale. Being self-dual, the instantons provide precisely two gravitino zero modes, thus might induce dynamical supersymmetry breaking. We also point out an underlying analogy with the invisible-axionic domain walls bounded by axionic strings.

I. INTRODUCTION: GRAVITY SECTOR OF $D=4$, $N=1$ SUPERSYMMETRIC STRINGS

The Kalb-Ramond gauge field[1] is naturally linked with various kinds of strings. First compactified four-dimensional fundamental (super)string theories give rise to a Kalb-Ramond gauge field that may be interpreted as a spacetime torsion.[2] Another type of string is the Abrikosov-Nielsen-Olesen vortex.[3] They are topological defects in an Abelian Higgs model in which the U(1) gauge symmetry is spontaneously broken. The magnetic field is confined inside the vortex tubes [the interaction of these flux tubes is described by their intersections; the intersection point is nothing but an instanton configuration of the spontaneously broken U(1) gauge field, and properly interpreted as joining and splitting processes of the flux tubes] whose diameters are of order of the vector-boson Compton wavelength. The massive U(1) gauge field in the Higgs phase can be described in terms of a massive Kalb-Ramond field that couples to the world sheet of vortex tubes.[4] They neatly summarize the short-range interactions between vortex flux tubes. On the other hand, if a global U(1) symmetry is broken, global strings[5] form at which point the order parameter vanishes. Unlike the above gauge vortex tubes, the massless Goldstone bosons mediate a long-range interaction among the global strings. In four-dimensional spacetime, the U(1) Goldstone boson can be equally represented by a compact U(1) Kalb-Ramond gauge field. In QCD, the confinement phase should show a physical spectrum consisting of glueballs (closed color-electric flux tubes) and mesonic or baryonic strings. The Kalb-Ramond gauge field arises out of their collective excitations and is presumably responsible for the color-confinement phenomena.[1]

All these are not coincidences at all. Much like photons couple to the world line of an electrically charged particle, the Kalb-Ramond field couples to the world sheet of an oriented string. Depending on the origin of strings as exemplified above, the Kalb-Ramond field becomes either massless or massive. The underlying physics is all intertwined with one another in an intriguing way.

In this paper, we investigate nonperturbative dynamics of the U(1) Kalb-Ramond field coupled to four-dimensional $N=1$ supersymmetric heterotic strings. In string theory, the Kalb-Ramond gauge field is naturally paired with gravity and the dilaton field as massless modes. Thus, any nontrivial dynamics of the Kalb-Ramond field should be considered together with nontrivial dilaton and graviton fields. This will lead to several unique modifications unexpected from pure Kalb-Ramond theory, as will be elaborated in this paper.

In the rest of this section, we summarize the gravitational sector of the low-energy effective Lagrangian arising from compactified four-dimensional heterotic superstrings. In Secs. II and III, to the lowest order of the string world sheet and the loop perturbation expansion, we explicitly construct self-dual string instantons and non-self-dual wormholes. In Sec. IV, we find fermionic zero modes in the instanton and wormhole backgrounds. It is shown that only instantons provide two chiral gravitino zero modes, the right number to induce dynamical supersymmetry breaking. Carrying nontrivial "magnetic" Kalb-Ramond charges, the instantons and wormholes immediately lead to a nonperturbative generation of the mass gap of the Kalb-Ramond gauge field.[6] In four-dimensional spacetime, the consequence is a confinement of oriented strings into domain walls.[7] (For a discussion of the phase structure of pure Kalb-Ramond gauge theory, refer to Ref. 4, and references therein.) Furthermore, the "spacetime" instantons invalidate the superpotential nonrenormalization theorem[8,9] and lead to a generation of superpotential F terms on a nonperturbative level. All these are shown in Sec. V. In fact, as explained in Sec. VI, these phenomena are very analogous to the confinement of invisible-axion[10] strings in the standard model extended to incorporate Peccei-Quinn symmetry. The QCD instantons break explicitly the nonlinearly realized Peccei-Quinn symmetry, and generate a nonperturbative axion potential. Because of the electric disorder provided by the QCD instantons, the "magnetic" flux out of the axionic string, originally radially symmetric, collapses down to a wall whose thickness is of the order of the QCD scale. Thus axionic strings get confined to axionic domain walls, forming the boundaries of the domain walls.

Let us consider the ten-dimensional heterotic string whose internal six dimensions are compactified on a six-dimensional Calabi-Yau manifold, denoted as K_6.[11] Because of the vanishing first Chern class of K_6, this compactification is known to lead to four-dimensional, $N=1$ spacetime supersymmetry. [In fact, our subsequent considerations are largely independent of the precise compactification scheme as long as the four-dimensional field theory preserves $N=1$ spacetime supersymmetry. Thus, we may abstract the relevant compactifications as any $c=9$ (2,2) conformal field theories.] Since the second cohomology group $H^{(1,1)}(K_6)$ is nonempty, the Kaluza-Klein modes of the ten-dimensional Kalb-Ramond field give rise to four-dimensional massless scalar fields (would-be axions) whose number equals the second Betti number b_2 of the Calabi-Yau manifold. These massless scalar fields are axions since, on the quantum level, they possess Peccei-Quinn symmetries and axionlike couplings to the gauge and the matter fields. However, it is known that world-sheet instantons[12,13] break Peccei-Quinn symmetries associated with most would-be axions. Fermionic zero modes of the world-sheet instanton suppress this breaking for (at least) one of the compactification generated axion field. Together with the internal dilaton field and their superpartners coming from the shape and size deformations of the Calabi-Yau internal manifold, the remaining truly axion field forms one chiral scalar superfield T of the gravity sector in the four-dimensional $N=1$ supergravity. There exists also a model-independent chiral scalar superfield S, directly coming from the ten-dimensional dilaton and the Kalb-Ramond gauge field and their superpartners.[14] Our main concern in this paper is the dynamics of these chiral scalar superfields in four-dimensional compactified heterotic superstring theory. The Kähler Lagrangian of these fields is compactly written as a nonlinear σ model over a coset space that will be identified shortly:

$$\frac{1}{\sqrt{-g}}\mathcal{L}=-\frac{1}{2\kappa^2}\mathcal{R}(g)+(K_{,Z\bar{Z}}g^{\mu\nu}\nabla_\mu\bar{Z}\nabla_\nu Z)\ . \quad (1.1)$$

We abbreviate S and T fields by Z, suppress all higher-order terms in gradient expansion, and denote the four-dimensional gravitational coupling constant $8\pi G_N=\kappa^2$. Scaling symmetry at the string tree level and four-dimensional $N=1$ spacetime supersymmetry uniquely determine the leading-order Kähler potential:

$$K(S,T)=-\frac{1}{\kappa^2}\ln(S+\bar{S})-\frac{3}{\kappa^2}\ln(T+\bar{T})\ . \quad (1.2)$$

Inserting the Kähler potential equation (1.2), the Lagrangian equation (1.1) is written as

$$\frac{1}{\sqrt{-g}}\mathcal{L}=-\frac{1}{2\kappa^2}\mathcal{R}(g)+\frac{1}{\kappa^2}g^{\mu\nu}\left[\frac{\nabla_\mu\bar{S}\nabla_\nu S}{(S+\bar{S})^2}+3\frac{\nabla_\mu\bar{T}\nabla_\nu T}{(T+\bar{T})^2}\right]\ . \quad (1.3)$$

The two scalar supermultiplets may be decomposed as

$$S=\exp(D)+iF_a\kappa A\ \cdots$$

and (1.4)

$$T=\exp\left[\frac{d}{\sqrt{3}}\right]-i\frac{f_a\kappa}{\sqrt{3}}a\ \cdots\ ,$$

where F_a and f_a denote the Peccei-Quinn scales associated with S and T fields, respectively. Ellipses denote the dilatinos and auxiliary fields. Using Eq. (1.4), bosonic part of the Lagrangian equation (1.3) can be put into

$$\frac{1}{\sqrt{-g}}\mathcal{L}=-\frac{1}{2\kappa^2}\mathcal{R}(g)+\frac{1}{4\kappa^2}(\nabla D)^2+\frac{F_a^2}{4}e^{-2D}(\nabla A)^2+\frac{1}{4\kappa^2}(\nabla d)^2+\frac{f_a^2}{4}e^{-(2/\sqrt{3})d}(\nabla a)^2\ . \quad (1.5)$$

If one considers the gauge sector at the same time, there are other scalar fields in the definition of the chiral superfield T coming from the dimensionally reduced gauge fields. Primarily interested in the gravity sector, we shall set them to zero in what follows.

The T superfield in Eq. (1.4) possesses a hidden non-compact $\mathrm{SL}(2,\mathbb{R})$ symmetry, associated with the coset space $\mathrm{SU}(1,1)/\mathrm{U}(1)\approx\mathrm{SL}(2,\mathbb{R})/\mathrm{U}(1)$ on which the matter Lagrangian in Eq. (1.1) is defined

$$iT\to\frac{aiT+b}{ciT+d},\quad a,b,c,d\in\mathbb{R},\quad ad-bc=1\ . \quad (1.6)$$

In particular, its discrete subgroup $\mathrm{SL}(2,\mathbb{Z})$ represents modular invariance (generalized duality symmetry) relating between strong and weak couplings. (Even though the modular invariance has been established only for the T moduli field, we suspect that there is a stringy Montonen-Olive[15]-type duality associated with S moduli field. This possibility is further hinted by a stringy cosmic-string solution recently found by Dabholkar and Harvey as a dual soliton to the string instanton we will find in this paper.[16]) This symmetry will play an intriguing role when we discuss the physical implications of the string instantons and wormholes in Sec. V.

The two chiral superfields S and T span the submanifold of the moduli space of the $N=1$ supersymmetric four-dimensional heterotic string theory. The tree-level string theory possesses a flat direction associated with them. It is of prime importance to find any mechanism to lift the tree-level flat directions of the superpotential. However, there is a powerful nonrenormalization theorem of the superpotential that forbids any perturbative renormalization of the superpotential.[8,9] On the other hand, it has been observed that nonperturbative physics may evade the perturbative nonrenormalization theorem. Such examples known so far include worldsheet string instantons,[12,13] and gluino condensation[17] in the strongly interacting, hidden gauge sector in $E_8 \times E_8$ heterotic string theory. As will be shown, the spacetime string instantons we will discuss in this paper provide yet another mechanism of generating nonperturbative superpotentials for both S and T moduli fields.

II. SPACETIME STRING INSTANTONS

We first study the Kähler Lagrangian of chiral scalar superfields in Eq. (1.5) in the gravitational background $g_{\mu\nu}$. Since the S and T moduli fields are completely decoupled, we find it sufficient to study the model

$$\mathcal{S} = \int d^4x \sqrt{-g} \left[\tfrac{1}{2}(\nabla D)^2 + \frac{f_a^2}{2\times 3!} e^{2\beta D} H^2 \right] . \tag{2.1}$$

There are two coupling constants in the theory: the Kalb-Ramond coupling constant f_a and the dilaton coupling constant β. Inferring Eq. (1.4), we see that β takes values $\beta=1$ and $\beta=1/\sqrt{3}$ for S and T fields, respectively. Recall that the scaling symmetry of classical strings and the $N=1$ supersymmetry fix the numerical values of β's.

We have set the gravitational constant $\kappa^2 \equiv 8\pi G_N = 1$ so that the dimensional analysis gives

$$[D]=[H_{\mu\nu\lambda}]=[f_a]=1/[\beta]=(\text{mass unit}) .$$

The Kalb-Ramond field strength $H=dB$ is related to the imaginary part of S or T moduli fields (axions) through the pseudoscalar representation

$$\nabla_\mu A = (e^{2\beta D} {}^*H)_\mu \equiv \frac{1}{3!} \frac{e^{2\beta D}}{\sqrt{-g}} \epsilon_{\mu\nu\lambda\sigma} H^{\nu\lambda\sigma} . \tag{2.2}$$

Since we are considering only leading order of the low-energy effective action, appropriate modifications of the Kalb-Ramond gauge field by the Yang-Mills and gravitational Chern-Simons three-forms are ignored. From Eq. (2.2), the "global" U(1) axion symmetry (nonlinearly realized Peccei-Quinn symmetry)

$$A(x) \to A(x) + C, \quad C=\text{const} \tag{2.3}$$

is related in the dual formalism to a compact U(1) "gauge" symmetry of the Kalb-Ramond field

$$B(x) \to B(x) + d\Lambda(x) , \tag{2.4}$$

where Λ is a one-form potential. In both representations, the symmetry generators are the closed zero-form C and two-form $d\Lambda(x)$, respectively. These two symmetries protect the Kalb-Ramond and the axion field from getting mass terms.

The action (2.1) possesses Poincaré duality between the dilaton and the Kalb-Ramond gauge fields. This is because the dilaton "field strength" dD is a one-form while the Kalb-Ramond gauge field H is a three-form.

The above argument leads us to expect an instanton solution. In fact, it is more evident by rewriting the action equation (2.1) as

$$\begin{aligned} S &= \tfrac{1}{2}\int d^4x \sqrt{-g}\, (\nabla_\mu D \pm f_a e^{\beta D} {}^*H_\mu)^2 \\ &\quad \mp \int d^4x \frac{f_a}{\beta} d(e^{\beta D}) \wedge H \\ &\geq \frac{f_a}{\beta} \left| \int d^4x\, d(e^{\beta D}) \wedge H \right| . \end{aligned} \tag{2.5}$$

Since the first term is manifestly positive semidefinite as long the dilaton and the Kalb-Ramond gauge fields are real valued, the Euclidean action is minimized once the dilaton field and the Kalb-Ramond gauge field satisfy

$$dD = \mp f_a e^{\beta D} H^* \quad \text{and } dH=0 . \tag{2.6}$$

[We have used "duality" in two different contexts here. The Kalb-Ramond gauge field may be equally represented by a pseudoscalar field that we called the axion. On the other hand, the duality of Eq. (2.6) relates "field strengths" of two different fields: i.e., the dilaton and the Kalb-Ramond fields. In effect, Eq. (2.6) interchanges the scalar dilaton and the pseudoscalar axion fields.] One may be suspicious of the existence of any nontrivial solutions of Eq. (3.6), by invoking Derrick's theorem. Namely, the action equation (3.5) changes as $\mathcal{S} \to \lambda \mathcal{S}$ under the rescalings[14]

$$g_{\mu\nu} - \lambda g_{\mu\nu}, \quad S \to \lambda S, \quad T \to T . \tag{2.7}$$

Thus, it appears that any nontrivial solution may be rescaled to a trivial vacuum configuration. However, the theorem implicitly assumes that the action comes from a local field configuration. In our case, the action is entirely from the asymptotic boundary behavior of fields as in topologically nontrivial situations, thus invalidating Derrick's theorem.

Indeed, the field configuration that satisfies Eq. (2.6) admits a topological interpretation. The action equation (2.5) is bounded by the Bogomolnyi inequality,[18] realized through spacetime supersymmetry. Minimum action instantons satisfy the (anti-)self-duality given by Eq. (2.6). The action of the instanton (or anti-instanton) is a topological charge defined as an intersection number:

$$\begin{aligned} \mathcal{S} &= \mp \int d^4x \frac{f_a}{\beta} (de^{\beta D}) \wedge H \\ &= \beta^{-2} \int d^4x\, (de^{\beta D})(de^{-\beta D}) \\ &= \frac{f_a}{\beta} \oint d\Sigma_3 e^{\beta D} H . \end{aligned} \tag{2.8}$$

In the second equality, we used the (anti-)self-duality condition equation (2.6). The last equality exhibits that the action can be expressed entirely in terms of a surface integral, as we argued above.

One can easily check that solutions to the first-order (anti-)self-duality condition equation (2.4) are automatically solutions of the equations of motion for the dilaton

$$\nabla^2 D - \frac{\beta f_a^2}{3!} e^{2\beta D} H^2 = 0 \tag{2.9a}$$

and the Kalb-Ramond gauge field

$$\nabla_\mu (e^{2\beta D} H^{\mu\nu\lambda}) = 0 \ . \tag{2.9b}$$

The readers should not have missed the close analogy with the (anti-)self-dual configurations in the classical Yang-Mills gauge theories. One minor difference is that, in the present case, we have two "gauge" fields: the dilaton field one-form dD and the Kalb-Ramond field three-form $H = dB$. There is another difference between the Yang-Mills gauge theory and the present system. Not surprisingly, the model equation (2.1) is not scale invariant. This is because the coupling constant of the Kalb-Ramond gauge field carries dimension one. More explicitly, we have

$$T_{\mu\nu} = \tfrac{1}{2}\partial_\mu D \partial_\nu D + \frac{f_a^2}{2\times 2!} e^{2\beta D} H_{\mu\alpha\beta} H_\nu^{\ \alpha\beta} - \tfrac{1}{4}\delta_{\mu\nu} \left[(\partial_\mu D)^2 + \frac{f_a^2}{3!} e^{2\beta D} H^2 \right] . \tag{2.10}$$

This is not traceless in general. However, if the dilaton and the Kalb-Ramond gauge fields satisfy the (anti-)self-dual condition equation (2.6), the energy-momentum tensor vanishes identically. This is also what happens to (anti-)instanton solutions of Yang-Mills gauge theory. In fact, the energy-momentum tensor equation (2.10) may be rewritten as

$$T_{\mu\nu} = \tfrac{1}{4}(\partial_\mu D + f_a e^{-\beta D} \partial_\mu A)(\partial_\nu D - f_a e^{-\beta D} \partial_\nu A) + (\mu \leftrightarrow \nu) - \tfrac{1}{4}\delta_{\mu\nu}[(\partial D)^2 - f_a^2 e^{-2\beta D} (\partial A)^2] \ . \tag{2.11}$$

Thus, the energy-momentum tensor vanishes identically for the (anti-)self-dual configurations. At the same time, the energy-momentum tensor expressed in Eq. (2.11) shows quite naturally the origin of the minus sign in front of the axionic part which was a confusing point in recent literature. The reason why we get it without any effort is because the Bianchi identity $dH = 0$ is automatically retained at every step in deriving Eq. (2.11).[19] Precisely the same phenomenon arises for a complex scalar field theory in a fixed global charge sector.[20] The dilaton field plays an analogous role to the modulus field of the complex scalar field. [In fact, one can shortcut the argument of Ref. 20 using a dual formulation. This is most easily done by rewriting the complex scalar field $\phi \equiv Re^{i\theta}$ and using the dual formulation to change the angular field into a Kalb-Ramond field $\nabla_\mu = R^{-2} H_\mu^*$. One immediately finds that the "radial" field R plays a similar role to the dilaton field in Eq. (2.1).]

Let us now solve the coupled equations (2.6) with the SO(4)-symmetry ansatz

$$ds^2 = d\tau^2 + \tau^2 d\Omega_3^2 \ . \tag{2.12}$$

From Eq. (2.6), we find that

$$0 = dH = \pm \frac{1}{\beta f_a} \nabla^2 e^{\beta D} \ . \tag{2.13}$$

A class of solutions which are regular everywhere except at the center of instanton are

$$e^{-\beta D(\tau)} = A + \frac{\beta |Q|}{2\pi f_a^2 \tau^2} \quad \text{and} \quad H_{abc}(\tau) = \frac{Q}{\pi f_a^2} \frac{\epsilon_{abc}}{\tau^3} \ . \tag{2.14}$$

An integration constant is denoted by A, and three independent Killing indices on the three-sphere Ω_3 are denoted by a,b,c. The "Kalb-Ramond charge" Q takes both positive and negative values. The absolute value of charge is taken for the dilaton solution in Eq. (2.14). This means that the self-dual instanton is taken for positive Kalb-Ramond charges, while the anti-self-dual instanton is taken for their negative charges.

There exists a constraint to the Kalb-Ramond charge Q from the requirement of consistent coupling of strings to the Kalb-Ramond gauge field. A string wave function must be single valued when the string sweeps a nontrivial Kalb-Ramond gauge field source. Thus, the Kalb-Ramond gauge flux should fall in the third integral cohomology group $H_3(M,\mathbb{Z})$.[7] This gives rise to a quantization of the Kalb-Ramond charge Q

$$2\pi\mathbb{Z} = f_a^2 \int_M H \rightarrow Q \in \mathbb{Z} \ . \tag{2.15}$$

The same quantized charge Q also enters the dilaton field configuration. *A priori*, there is no reason that the dilaton charge is quantized. However, in our instanton configuration, the self-duality equation (2.6) also relates the dilaton charge to the axion charge, thus allowing only quantized values.

Incidentally one may express Eq. (2.14) in a manifestly covariant form

$$e^{-\beta D(x)} = A + \frac{\beta |Q| 2\pi f_a^2}{|x - x_0|^2}$$

and (2.16)

$$H_{\mu\nu\lambda} = \frac{Q}{\pi f_a^2} \frac{\epsilon_{\mu\nu\lambda\sigma}(x-x_0)^\sigma}{|x-x_0|^4} \ .$$

The integration constant A is determined by dilaton boundary conditions at $\tau \rightarrow \infty$, the Euclidean asymptotic region far away from the center of the instanton. If the dilaton field approaches zero, $A = 1$, while a divergent dilaton field sets $A = 0$. We will call these type-A and type-B instantons, respectively. We shall see that these two types give completely different features of dilaton-axion dynamics. For example, one may measure the long-range field strength of the axion and the dilaton fields to define their respective charges. In the case $A = 1$, the dilaton and the Kalb-Ramond fields fall off as R^{-3} (monopole potential) and their charges, denoted by q_D and q_H, respectively, are type A:

$$q_D = -|Q| \quad \text{and} \quad q_H = Q \ . \tag{2.17}$$

On the other hand, for $A = 0$, the dilaton field falls off like R^{-1} (confining potential) while the Kalb-Ramond field does like R^{-3}. Their charges are type B:

$$q_D = \frac{2\pi}{\beta} \quad \text{and} \quad q_H = Q \; . \tag{2.18}$$

Thus, it appears the two charges are not related to each other. This sounds odd since the (anti-)self-duality condition relates them to each other. However, after the nontrivial dilaton field configuration is taken into account, an "effective" Kalb-Ramond charge read off from $e^{\beta D}H_{\mu\nu\lambda}$ becomes the same as q_D and, thus is independent of the "topological" charge Q.

What if we turn on the gravity? Since the energy-momentum tensor vanishes identically, the above self-dual spacetime instantons do not interact gravitationally. One can understand this from cancellations between the dilaton and the Kalb-Ramond field Euclidean energy densities

$$\begin{aligned} T_{00} &= \tfrac{1}{4}(\dot{D}^2 - f_a^2 e^{-2\beta D}\dot{A}^2) \\ &= \frac{1}{4}\left[\dot{D}^2 - \frac{f_a^2}{3!}e^{2\beta D}H_{abc}^2\right] \; . \end{aligned} \tag{2.19}$$

The dilaton field furnishes a positive energy density while the Kalb-Ramond gauge field furnishes a negative one with precisely the same magnitude. This fact gives us another bonus. Since the instantons are gravitationally neutral, one can immediately get (anti-)self-dual multi-instanton solutions by superposing individual single instanton solutions:

$$e^{-\beta D(x)} = A + \sum_a \frac{|Q_a|/2\pi f_a^2}{|x-x_a|^2}$$

and (2.20)

$$H_{\mu\nu\lambda}(x) = \sum_a \frac{Q_a}{\pi f_a^2}\frac{\epsilon_{\mu\nu\lambda\sigma}(x-x_a)^\sigma}{|x-x_a|^4} \; .$$

This phenomenon is reminiscent of the Prasad-Sommerfield limit of non-Abelian magnetic monopoles. There, the vector field and the Higgs field exchange forces cancel each other. In our string instantons, the attractive dilaton exchange interaction is counterbalanced by the repulsive Kalb-Ramond exchange interaction. That is, the dilaton field plays a role similar to the Higgs field in non-Abelian magnetic monopoles. In fact, this identification is quite well justified from the analogy of the dilaton-axion system to the complex scalar field system.

Let us calculate a single instanton action. For type A, the action is calculated to be finite

$$S_A(\text{instanton}) = \frac{2\pi}{\beta f_a}|Q| \; . \tag{2.21}$$

On the other hand, for type-B instantons, the action is quadratically divergent. If we infrared regularize so that $\tau \le T$,

$$S_B(\text{instanton}) = \frac{4\pi^2}{\beta^2}T^2 \; . \tag{2.22}$$

However, using the dilaton solution equation (2.14), we can compactly express the instanton actions for both cases as

$$S(\text{instanton}) = \frac{2\pi}{\beta f_a}|Q|e^{\beta D(\infty)} \; . \tag{2.23}$$

This is obtained as follows. One may imagine a topologically trivial vacuum configuration with a background dilaton field at some constant value $D(\infty)$. One may cut out a four-dimensional ball B_4 out of the background, and patch the instanton configuration given in Eq. (2.14). The dilaton field should match the background value smoothly. The excess value of the action to paste such an instanton configuration is precisely Eq. (2.23). The infrared divergence of a type-B instanton action may be interpreted due to the divergent dilaton field far away from the instanton.

We emphasize that the spacetime string instantons exist for both S and T moduli fields. Namely, the solution is insensitive to the dilaton coupling constant β. This is in sharp contrast to other interesting instanton solutions in string theory. For example, the existence of world-sheet instantons[12,13] crucially depends upon the topological structure of the initial manifold K_6, and thus only to the T moduli field. Some of the spacetime wormholes we will discuss in the next section exist only with the T moduli field again.

The cautious reader must have worried about the singular nature of the string instanton solutions we obtain in Eq. (2.14). The dilaton and the Kalb-Ramond fields diverge at the center of the instanton. After all, the field configuration is not much different from a singular Dirac magnetic monopole solution in Maxwell's electrodynamics. However, there are significant differences between the latter and our string instanton. As noted in Eqs. (2.21) and (2.22), the instanton action gets no contribution from the $\tau \approx 0$ region. The divergence we encountered in the type-B instanton were from the infrared region, not from the center of the instanton. This is in contrast with the Dirac magnetic monopoles, in which the self-energy diverges unless we regularize the monopole to a finite size. Furthermore, since we have used an effective low-energy four-dimensional effective field theory of superstrings, the fine structures of solutions Eq. (2.14) on a scale shorter than the compactification scale $\tau \le M_c^{-1}$ should be treated through a full-fledged ten-dimensional string theory. In particular, the internal six-dimensional manifold should start to show complicated massive mode excitations. From the point of view of four-dimensional effective field theory, these massive mode excitations are expected to act as a source to the massless gravity sector. In effect, the Bianchi identity of the Kalb-Ramond field gets modified to

$$dH = Q\delta^{(4)}(x-x_0) \; . \tag{2.24}$$

The vanishing contribution of the action from the core region of the instanton suggests that we may replace the short-distance field configurations of Eq. (2.14) by some regular string-theory solutions. Thus moduli space of our spacetime string instantons would be much larger than naively expected from the field-theory limit, once deformed through a full string configuration space. The situation is very analogous to the 't Hooft–Polyakov magnetic monopoles in non-Abelian gauge theories. Far outside the monopole core, only the long-range Abelian gauge field component is present. This is well approximated by a singular Dirac magnetic monopole. However, inside the monopole core region, the short-ranged non-Abelian gauge field and Higgs field excitations are present. These excitations render the monopole self-energy finite. In our spacetime string instantons, the dilaton field plays a role very similar to the Higgs field in 't Hooft–Polyakov magnetic monopoles. The nontrivial dilaton field configuration made the singularities of the Kalb-Ramond gauge field near the center of instanton wash out completely. Again, we encountered a close analogy of the dilaton field with the Higgs field. Another related example in which the apparent singularity is illusory is Kaluza-Klein magnetic monopoles.[21] Their field configurations look singular from the lower-dimensional point of view. However, once expressed in terms of the original higher-dimensional fields, the magnetic monopoles are perfectly regular.

III. STRING MERONS AND SPACETIME WORMHOLES

In gauge theories, it is known that there exist other singular classical solutions with negative energy-momentum tensor, so-called merons, in addition to the self-dual instantons. Merons and nested merons in non-Abelian gauge theories become wormholes[22,23] once the gravitational interaction is turned on. (The merons and meronic wormholes may well be unstable against polarized instanton configurations. This instability is partly because they do not carry conserved topological charges. The string wormholes we will describe below have nontrivial topological charges, thus stable.)

In this section, we construct similar solutions, string merons, that possess negative values of the energy-momentum tensor. Not surprisingly, once gravity is turned on again, the string merons become spacetime string wormholes the same way as non-Abelian gauge theory merons.[22]

To this end, we again use the spherical coordinates of Eq. (2.12). The equations of motion are the same as in Eq. (2.9). The Kalb-Ramond equation of motion is solved by the same configuration as in Eq. (2.14):

$$H_{abc}=\frac{Q}{\pi f_a^2}\frac{\epsilon_{abc}}{\tau^3} . \tag{3.1}$$

The first integral of the dilaton equation of motion reads

$$\tau^6 T_{00}=\frac{1}{4}\left[(\tau^3\dot{D})^2-e^{2\beta D}\frac{Q}{\pi f_a^2}\right]^2=-\tfrac{1}{4}E\le 0 . \tag{3.2}$$

Here, E is an integration constant. We find solutions exist only if E is positive semidefinite.

Introducing a new variable $\theta(\tau)$ such that

$$\sqrt{E}\,\cosh\theta(\tau)=\frac{|Q|}{\pi f_a^2}e^{\beta D(\tau)}$$

and (3.3)

$$\sqrt{E}\,\sinh\theta(\tau)=\tau^3\dot{D}(\tau) ,$$

one can explicitly solve the dilaton field equation of motion to get

$$e^{-\beta D(\tau)}=A+\frac{|Q|}{\pi f_a^2\sqrt{E}}\sin\left[\frac{\beta\sqrt{E}}{2\tau^2}\right] , \tag{3.4}$$

As in the instanton case, we have two possible types of wormholes: type A in which the asymptotic dilaton field becomes infinite; type B in which the asymptotic dilaton field approaches finite, weak coupling. As a consistency check, we find the solution (3.4) interpolates to the instanton solution equation (2.14) continuously as $E\to 0$.

The solution, however, develops infinitely many singularities and domains of the complex-valued dilaton field configuration. This is evident from the solution of Eq. (3.4). As "time" τ is decreased from infinity to a critical time

$$\tau_c^{-2}=\frac{2}{\beta\sqrt{E}}\arcsin\left[A\frac{\pi f_a^2\sqrt{E}}{|Q|}\right] , \tag{3.5}$$

the dilaton field diverges indefinitely. Below the critical time, the dilaton field becomes complex valued until the time reaches the next critical value. The dilaton field repeats its real and complex values as time τ is decreased toward zero. We note that the critical time τ_c depends upon the strength of parameter β. If β becomes larger, the physically allowed domains get narrower and the complex-valued dilaton field oscillates faster.

More transparently, one may examine phase trajectory of the dilaton field by rewriting Eq. (3.2) to

$$\frac{1}{2}\left[\frac{\tau^3}{\beta\sqrt{E}}\frac{\partial}{\partial\tau}e^{-\beta D}\right]^2+\tfrac{1}{2}(e^{-\beta D})^2=\frac{|Q|^2}{\pi^2 f_a^4 E} . \tag{3.6}$$

One can view Eq. (3.6) as describing a mechanical analog of one-dimensional simple harmonic oscillator with energy $\mathcal{E}=|Q|^2/\pi^2 f_a^4 E$ whose one-dimensional coordinate is $X=e^{-\beta D}$ measured in time $T=\beta\sqrt{E}/2\tau^2$. The analog particle trajectories are thus circles in phase space. However, unless the dilaton field becomes complex valued, it cannot complete a whole cycle of the trajectory through the region of the negative X axis. At best, one may hope the left-half of phase space is inaccessible by imposing a suitable boundary condition on the dilaton field. We do not find such a boundary condition. Thus, we regard the meron solution unphysical.

The situation improves significantly once we turn on the gravity. We first note that, due to the negative-energy-momentum tensor in Eq. (3.2), the merons turn into string wormholes. The reason why turning on the gravity improves the situation is because a finite wormhole neck size makes almost all of the singularities inaccessible. It is almost because, depending on the strength of β, one may still have the outermost singularity and a core region in which the dilaton field is complex valued. This gives a criterion of critical value β_c for the parameter β. The same phenomena occurs also in the charged black-hole solution with nontrivial dilaton fields.[24]

Let us see this in detail. We consider Euclidean spacetime metric to be

$$ds^2 = N(\tau)^2 d\tau^2 + a(\tau)^2 d\Omega_3^2 . \tag{3.7}$$

A convenient gauge is a Robertson-Walker metric with $N(\tau)=1$. Again, the nontrivial Kalb-Ramond field solution reads

$$H_{abc} = \frac{Q}{\pi f_a^2}\frac{\epsilon_{abc}}{a^3(\tau)} . \tag{3.8}$$

The first integral of the dilaton equation of motion is properly modified from Eq. (3.2) to

$$\left[(a^3\dot{D})^2 - e^{2\beta D}\left[\frac{Q}{\pi f_a^2}\right]^2 \right] = -E . \tag{3.9}$$

Using this, the zeroth component of the Einstein equation of motion reads

$$\dot{a}^2(\tau) - 1 = -\frac{E}{12}a^{-4}(\tau) . \tag{3.10}$$

One can solve these equations straightforwardly. The conformal factor geometry is

$$a(\tau) \sim (\tau-\tau_0)\left[1+\frac{E/72}{(\tau-\tau_0)^4}\right] . \tag{3.11}$$

From this, one finds that there exists a minimum neck size of $\bar{a}^2 = -\sqrt{E/12}$ at which the extrinsic curvature vanishes. Using Eq. (3.9) and a similar parametrization as in Eq. (3.3), the dilaton solution reads

$$\arcsin\left[\sqrt{E}\frac{\pi f_a^2}{|Q|}(e^{-\beta D(\tau)} - C)\right] = \beta\sqrt{E}\int_\infty^\tau dt\, a^{-3}(t) . \tag{3.12}$$

Again, the integration constant C signifies an asymptotic behavior of the dilaton field: for type A, $C=1$, while, for type B, $C=0$.

Using the Einstein equation (3.11), we have

$$\sqrt{E}\int_\infty^\tau dt\, a^{-3}(t) = \frac{1}{\beta_c}\arcsin\left[\sqrt{E}\frac{\beta_c/2}{a^2(\tau)}\right] . \tag{3.13}$$

Here, $\beta_c = 1/\sqrt{3}$. Combining Eqs. (3.12) and (3.13),

$$\arcsin\left[\sqrt{E}\frac{\pi f_a^2}{|Q|}(e^{-\beta D(\tau)} - C)\right] = \frac{\beta}{\beta_c}\arcsin\left[\sqrt{E}\frac{\beta_c/2}{a^2(\tau)}\right] . \tag{3.14}$$

As a consistency check, we take $E\to 0$, the string instanton limit. It is straightforward to verify that the solution equations (3.7) and (3.13) reduce to the string instanton solution equation (2.14). Contrary to what it appears, this limit is not so trivial. This is because the curvature becomes singular as the wormhole neck pinches down in the limit $E\to 0$.

For type A wormholes, the dilaton field in Eq. (3.14) is regular everywhere outside the wormhole neck. For type B wormholes, the dilaton field has no naked singularity at finite proper time $\tau<\infty$ if and only if

$$\beta < \beta_c \equiv \frac{1}{\sqrt{3}} . \tag{3.15}$$

This determines the critical coupling strength of the dilaton field to have a regular solution: wormholes out of the S moduli field possess naked singularities, while those out of the T moduli field are regular except the spacetime infinity.

In principle, we can imagine a continuously varying dilaton coupling constant β. As the dilaton coupling constant is increased beyond the critical value β_c, the dilaton exchange attraction becomes so strong that a naked singularity starts to show up at a finite proper distance away from the center of instanton.

We now calculate the wormhole action in the limit $E\approx 0$. We find

$$S(\text{wormhole}) = S(\text{instanton}) + \sqrt{E}\frac{\sqrt{3}\pi^2}{18}(2\beta_c^{-2}-\beta^{-2}) + O(E) . \tag{3.16}$$

Again, the type B wormholes show infrared divergences, while type A wormholes are finite. The second term in Eq. (3.16) is a leading contribution from the wormhole neck in a perturbative expansion in powers of $\sqrt{E}$. Also, for nonzero E, the wormhole action is larger than that of the instantons. This is in agreement with generalized Bogomolnyi inequality

$$S(\text{wormhole}) \geq S(\text{instanton}) . \tag{3.17}$$

The equality is saturated in the limit that the wormhole energy $E\to 0$ at which the (anti-)self-duality of the dilaton and the Kalb-Ramond field holds.

As in the instanton case, the wormhole action can be written compactly in terms of the asymptotic dilaton field configuration and the wormhole neck correction that we denote as $\Delta(\sqrt{E})$:

$$S(\text{wormhole}) = \frac{2\pi}{\beta f_a}|Q|e^{\beta D(\infty)} + \Delta(\sqrt{E}) . \tag{3.18}$$

From the above wormhole action calculation, we expect the correction $\Delta(\sqrt{E})$ to be an analytic, positive function of $\sqrt{E}$.

Finally, let us compare our string wormholes with those found by Giddings and Strominger.[25] They imposed a different boundary condition such that the time derivative of dilaton field vanishes at the wormhole neck. This changes the dilaton field configuration but the Kalb-Ramond and metric field configurations remain the same as ours. Furthermore, their solution also shows a singularity to the dilaton field at a finite proper distance away from the center unless $\beta<\beta_c=1/\sqrt{3}$. This is precisely the same condition for our type B wormholes as given in Eq. (3.15). Furthermore, it is intriguing to observe that, in the case $\beta=\beta_c$ of the T moduli field, our type B wormhole solution agrees with the one of Giddings and Strominger. The infrared-divergent wormhole action agrees in both cases, modulo a correction from the finite wormhole neck we have explicitly calculated.

IV. FERMIONIC ZERO MODES

In this section, we investigate fermionic zero modes in the instanton or wormhole backgrounds. We restrict ourselves to fermionic superpartners of the graviton and the two dilaton-axion moduli fields: spin-$\frac{3}{2}$ gravitino $\psi_\mu(x)$ and two spin-$\frac{1}{2}$ dilatinos $\lambda(x)$. We will explicitly count the zero modes of fermionic partners of the S moduli field, but the same counting goes through to the T moduli field too.

The supersymmetry transformations of fermionic fields are

$$\begin{aligned}\delta\psi_\mu(x)&=(\nabla_\mu-\tfrac{1}{2}e^{D}H_\mu)\epsilon(x)\ ,\\ \delta\lambda(x)&=(\nabla D+e^{D}H)\epsilon(x)\ .\end{aligned}\tag{4.1}$$

We adopted the notation

$$\begin{aligned}&\nabla D\equiv\gamma^\mu\nabla_\mu D,\quad H_\mu\equiv\frac{1}{2!}H_{\mu\nu\lambda}\gamma^{[\nu}\gamma^{\lambda]}\ ,\\ &H\equiv\frac{1}{3!}H_{\mu\nu\lambda}\gamma^{[\mu}\gamma^{\nu}\gamma^{\lambda]}\ .\end{aligned}\tag{4.2}$$

Also, our notation differs from that of Chapline and Manton by a redefinition of the gravitino (a similar trick has been noted by the authors in Ref. 26; we thank Strominger for informing us of these references):

$$\psi^\mu=\psi^\mu_{\rm CM}-\tfrac{1}{2}\gamma^\mu\lambda\ .\tag{4.3}$$

The supersymmetry transformations of bosonic fields read

$$\begin{aligned}\delta g_{\mu\nu}&=\bar\epsilon(x)\gamma_{(\mu}\psi_{\nu)}\ ,\\ \delta B_{\mu\nu}&=\bar\epsilon(x)\gamma_{[\mu}\psi_{\nu]}\ .\end{aligned}\tag{4.4}$$

The bosonic transformations show that the Kalb-Ramond gauge field can be interpreted as a spacetime torsion. More precisely, the covariant derivative in a nontrivial dilaton and Kalb-Ramond field gets modified into

$$\nabla_\mu\rightarrow\hat\nabla_\mu\equiv\nabla_\mu-\tfrac{1}{2}e^{\beta D}H_\mu\ .\tag{4.5}$$

a new covariant derivative with nonzero torsion. This is precisely the structure that enters into the gravitino supersymmetry transformation given in Eq. (4.1).

Let us first count the gravitino zero modes. The gravitino zero modes are provided by those components of supersymmetry transformation spinor $\epsilon(x)$ that approach constant at Euclidean spacetime infinity, and, thus, are not normalizable, while their covariant derivatives fall off sufficiently fast so that the gravitino wave functions are normalizable. Unless there exist covariantly constant spinors (with nonvanishing torsion provided by the dilaton and the Kalb-Ramond fields), we therefore expect four gravitino zero modes.

Suppose there exist a covariantly constant spinor, $\epsilon_0(x)$. It gives a vanishing gravitino wave function

$$\delta\psi_\mu(x)=(\nabla_\mu-\tfrac{1}{2}e^{D}H_\mu)\epsilon_0(x)\ .\tag{4.6}$$

In this case, the spinor must satisfy an integrability condition

$$\nabla_{[\mu}\delta\psi_{\nu]}(x)=[\nabla_\mu-\tfrac{1}{2}e^{D}H_\mu,\nabla_\nu-\tfrac{1}{2}e^{D}H_\nu]\epsilon_0(x)=0\ .\tag{4.7}$$

Expanding out the left-hand side, we have

$$(R_{\mu\nu}-\nabla_{[\mu}e^{D}H_{\nu]}+\tfrac{1}{4}e^{2D}H_{[\mu}H_{\nu]})^{ab}\sigma_{ab}\epsilon_0(x)=0\ .\tag{4.8}$$

In the background of spacetime instantons, the dilaton and the Kalb-Ramond gauge fields satisfy a generalized self-duality condition

$$\nabla_\mu e^{-D}=\pm H^*_\mu\ ,\tag{4.9}$$

so that Eq. (4.8) becomes

$$\nabla_{[\mu}D\nabla^\alpha D\sigma_{\alpha\nu]}(1\mp\gamma_5)\epsilon_0(x)=0\ .\tag{4.10}$$

For a nontrivial covariantly constant spinor, Eq. (4.10) is satisfied if and only if it has a positive chirality $\gamma_5\epsilon_0(x)=+\epsilon_0(x)$. Namely, there exist two covariantly constant spinors of the same chirality in the spacetime string instanton background. This in turn implies that there exist two gravitino zero modes of negative chirality, since the other two states are annihilated by the supersymmetry transformation. This is precisely the right number one needs to induce dynamical supersymmetry breaking since it gives the nonvanishing gravitino two-point function

$$\langle|\psi_{\mu a}\psi_{\nu b}|\rangle_{\rm instanton}\neq0\ .\tag{4.11}$$

Indeed, our spacetime instanton may be regarded as a concrete realization of Witten's idea[27] that gravitational instantons with nontrivial torsion might break supersymmetry. One can regard our spacetime string instantons as gravitational instantons. This is more transparent once the field variables are transformed into those of σ model variables to which the strings couple directly. The dilaton and the Kalb-Ramond gauge fields serve as a warp factor and a spacetime torsion, respectively. A string consistency of having nontrivial torsion provided by the Kalb-Ramond field requires us to include a nontrivial dilaton field at the same time. (Supersymmetry breaking using Eguchi-Hanson gravitational instantons has been proposed recently.[28] However, these instantons have a global topology that is not asymptotically Euclidean. In this regard, it is not clear whether one should include the Eguchi-Hanson instanton to the quantum level of supergravity theory. Thus, we find the proposed mechanism less appealing than other possibilities with asymptotically Euclidean instantons such as ours.)

Next, we consider the dilatino zero modes. From Eq. (4.1), the dilatino wave function in the spacetime instanton background reads

$$\delta\lambda(x) = \nabla e^{D}(1 \pm \gamma_5)\epsilon(x) . \tag{4.12}$$

Therefore, we find two dilatino zero modes of positive helicity in the spacetime string instanton background. Half of the supersymmetry transformation annihilates the instanton states.

Overall, we thus have two chiral gravitino zero modes and two chiral dilatino zero modes, half of the number one naively expects. (These numbers could have been derived by explicitly solving the Dirac and the Rarita-Schwinger equations. I checked them and found that the number of fermionic zero modes agrees with the above counting.) As shown above, the reduction of zero modes by half is due to the self-duality that the spacetime string instanton backgrounds provide for. Also, we observe a definite relationship between the two chiralities of gravitino and dilatino zero modes. Had we chosen a gauge of $\epsilon(x)$ orthogonal to the covariantly constant spinors, there would be two gravitino zero modes but no dilatino zero modes. On the other hand, if we had chosen a guage such that $\epsilon(x)$ is proportional to the covariantly constant spinor, there would be no gravitino zero modes but two dilatino zero modes.

Finally, we consider the spacetime wormhole background. Since the wormhole backgrounds are not self-dual, we do not find any nontrivial solution of covariantly constant spinor $\epsilon_0(x)$ to Eq. (4.7). Thus, we have four gravitino zero modes. Similarly, we find four dilatino zero modes, since there is no chiral projection. Since wormholes generate four gravitino zero modes, we find them not directly relevant to the possibility of dynamical supersymmetry breaking.

V. LOW-ENERGY IMPLICATIONS OF STRING INSTANTONS

What are the low-energy implications of these spacetime string instantons and wormholes? We observed that the essential features of instantons and wormholes are axionic charges carried by them. This fact leads to several intriguing consequences for low-energy physics: for instance, a string confinement and a nonperturbative renormalization of superpotential.

A. String confinement

Recall that the instantons carry a nontrivial Kalb-Ramond magnetic field. Once the instantons and/or wormholes are turned on, the Kalb-Ramond vacuum is "magnetically" disordered. Thus, the "electric" field flux of the Kalb-Ramond gauge field gets confined. In particular, the four-dimensional Kalb-Ramond gauge theory admits only a confinement phase for all values of coupling constant, once the nonperturbative instanton effects are taken into account. No free Kalb-Ramond gauge field is a physical spectrum. The gauge theory sector develops a nonperturbative mass gap through a Debye screening. For pure Kalb-Ramond gauge theory in four dimensions, this can be seen easily from the following argument. Let us consider an order parameter $\mathcal{L}_A[\Sigma^*] = \exp(i\int_{\Sigma^*} dA)$ of the dual axion field $dA = H^*$. Recall that the dual axion field A possesses a global U(1) symmetry, that is dual to the local U(1) symmetry of the original Kalb-Ramond field. The order parameter of the latter is a Wilson "surface" $\mathcal{W}[\Sigma] = \exp(i\int_{\Sigma} H)$. The "world histories" Σ and Σ^* are dual to each other. Let us consider a topological linking correlation function defined by

$$\exp(i\mathcal{G}_2) = \langle 0|\mathcal{W}[\Sigma]\mathcal{L}_A[\Sigma^*]|0\rangle . \tag{5.1}$$

Since the Wilson "surface" measures the "magnetic" flux of the Kalb-Ramond gauge field as the A field traverses around, the two-point function $\mathcal{G}_2$ in Eq. (5.1) is topologically determined by a mutual Gauss' linking number and the gauge group. Once the "magnetic" instanton fluctuations are important, the global U(1) symmetry associated with the A field is explicitly broken. Thus, the phase of Eq. (5.1) changes appreciably only for a finite distance along a dual link direction Σ^*. This implies that the Wilson "surface" shows a "volume" law behavior, the right behavior for the confinement phase of pure Kalb-Ramond gauge theory. In terms of the pseudoscalar axion field A, the "electric" instantons induce a domain wall whose boundary is nothing but the closed string $\mathcal{L}_A[\Sigma^*]$. This argument is applied to our case almost the same, except for a small but significant modification due to a nontrivial dilaton field as we shall see below.

Suppose a supersymmetric string sitting without the instanton fluctuations was turned on. Since the superstring also acts as a source of the Kalb-Ramond gauge field, the Kalb-Ramond field around the string is radially symmetric and long ranged. Once the instanton fluctuations are important, however, the radially symmetric Kalb-Ramond field is no longer energetically favorable. The Kalb-Ramond electric field collapses down to a domain wall to form an electric flux surface. The energy of the domain wall is the surface tension times the area. The domain walls may form out of a single string or multiple strings. Any oriented strings (they are the ones carrying the Kalb-Ramond gauge charge) should be confined into domain walls.

The confinement of superstrings to axionic domain walls has been already suggested by Witten[7] for the S moduli field. We expect that the same phenomena arises also from those instantons and wormholes composed out of the T moduli fields. This is because both the S and T moduli fields have Peccei-Quinn symmetries and couplings in low-energy string theory. The Peccei-Quinn scale of both moduli fields is naturally the compactification scale. The spacetime string instantons and/or wormholes give rise to "magnetic" disorder and a nonperturbative confinement of strings into domain walls. Thus, the string deconfinement phase transition is directly related to the spontaneous compactification of the extra six dimensions. (Recall that the pure Kalb-Ramond gauge theory has a Coulomb phase if the spacetime dimension is larger than four. Furthermore, once the strings are coupled to the Kalb-Ramond field, there exists also a Higgs phase continuously connected to a confinement phase.)

We can estimate the domain-wall surface tension. If a string defines one of the boundaries of the domain wall, the surface tension and the string tension compete with each other. The total energy is roughly

$$T2\pi R-\mu\pi R^2\sim E(R)\ . \qquad (5.2)$$

The minimum of energy is at $R\sim T/\mu$. Since the string tension $T\approx M_{\rm Pl}^2$, while the surface tension $\mu\approx M_c^2 M_{\rm Pl}$, an average radius of the confined string to domain wall is about $M_{\rm Pl}/M_c^2 > M_c^{-1}$ Since the size is larger than the compactification scale, the four-dimesnional approximation might be well justified. This string confinement is largely independent of the type of underlying string theories, details of compactification schemes, thus fundamental strings are intrinisically confined within a microscopic scale.

However, in the above argument, we relied upon nonperturbative physics of pure Kalb-Ramond gauge theory. Once we take into account the nontrivial dilaton field configuration, the story is quite modified. In Ref. 29, we studied the modification of axion domain-wall structure once the dilaton field is included. Unless there exists a stable vacua to the dilaton field, we found that the dilaton field induces an instability to the axionic domain wall. The wall thickness is increased by a large factor. One could think of the dilaton field as a Higgs field to Kalb-Ramond gauge theory. The confinement phase of Kalb-Ramond gauge theory is contaminated by the dilaton field. Because of this, the wall tension is considerably reduced. The string confinement on a microscopic scale around the compactification scale would arise only if some unknown mechanism furnishes a vacuum expectation value to the dilaton field well before the Kalb-Ramond dynamics is operative. One such possibility will be discussed in the next subsection.

B. Nonperturbative breakdown of nonrenormalized theorem

The spacetime string instantons have another consequence. Dine and Seiberg[8] proved a nonrenormalization theorem of the moduli field superpotential to all finite orders of string loop expansion. (A proof based upon microscopic analysis was completed by Martinec.[9]) However, we will see that our spacetime string instantons give rise to a nonperturbative breakdown of superpotential nonrenormalization theorem. (In fact, Dine and Seiberg[30] already anticipated that some unknown spacetime string instantons might generate a nonperturbative renormalization of the superpotential along the S moduli direction.) The nonrenormalization theorem of the superpotential in the four-dimensional effective Lagrangian of string theory is based upon the analytic structure of the possible superpotentiial $\mathcal{W}$ and a classical symmetry associated with the dilaton and axion fields. If the four-dimensional supersymmetry is not anomalous, the moduli fields S and T should enter the effective Lagrangian in a manifestly supersymmetric and analytic way. Scattering amplitudes, with an insertion of the Kalb-Ramond vertex operator at zero momentum, are easily shown to vanish[8] (the polarization tensor is denoted by $E_{\mu\nu}$):

$$\left\langle \cdots \int d^2\sigma\, E_{\mu\nu}\nabla X^\mu(\overline{\nabla}X^\nu+iK\cdot\psi\psi^\nu)\exp(iK\cdot X)\cdots\right\rangle_{K\to 0}=0 \qquad (5.3)$$

independent of the world-sheet topology and, thus, to all finite orders of string loop expansions. This is because the Kalb-Ramond vertex operator at zero momentum is a total derivative on any compact world sheet. (The subtlety associated with the zero-momentum limit is overcome either by an analytic continuation or by the manifest supersymmetry form of the vertex operators.)

In Sec. II, we learned that a gauge symmetry of the Kalb-Ramond field is equally well represented by a global symmetry of the axion field in a pseudoscalar field representation. Thus, the Peccei-Quinn symmetry associated with the spacetime axion field ImS is valid to all finite orders of string loop expansions. Since we assumed that the spacetime supersymmetry is nonanomalous, the superpotential should be an analytic function of the moduli fields S or T. This implies that the superpotential is independent of ReS as well and, thus, not renormalized to all finite orders of string loop expansions.

Once the spacetime string instantons are included, the story is different. Let us define a charge operator associated with the Peccei-Quinn symmetry $\hat{Q}_5=\int d^3\mathbf{X}\, i\dot{A}(t,\mathbf{X})$. If the spacetime string instanton (wormhole) induces a modification of $\delta L_{\rm eff}$ to the low-energy effective Lagrangian, the microscopic Peccei-Quinn symmetry dictates that

$$[\hat{Q}_5,\delta L_{\rm eff}]=Q\delta L_{\rm eff}\ . \qquad (5.4)$$

Thus

$$\delta L_{\rm eff}\propto \exp[iQA]\ . \qquad (5.5)$$

The effective operator induced by the anti-instanton (wormhole) is a complex conjugate of Eq. (5.5). Furthermore, in Secs. III and IV, we calculated the action of the spacetime instantons and wormholes. We refer to Eqs. (2.23) and (3.16). The asymptotic dilaton field configuration denotes an expectation value in much the same way as a θ angle dependence originates from an expectation value of the axion field A. Combining these two contributions, we finally get

$$\delta L_{\rm eff}=e^{-|Q|e^D}e^{iQA}(\text{fermion part})=\exp(-Q\cdot S_b)\mathcal{F}_1\cdots\mathcal{F}_{2N}\ . \qquad (5.6)$$

Here, we denoted the bosonic part of the chiral superfield S by S_b and yet-to-be determined fermion operators by $\mathcal{F}_a$, $a=1,\ldots,2N$. We emphasize that the fermions carry *no* Kalb-Ramond charge. There is no more quantum correction to this expression, thanks to the nonanomalous supersymmetry.

To complete the derivation, we now count the fermionic operators in Eq. (5.6). In a functional-integral formalism, we calculate scattering amplitudes in the string instanton or wormhole backgrounds to derive the low-energy effective Lagrangian. The functional integral with the instanton or wormhole background gives various bosonic (metric, dilaton, and axion) and fermionic (gravitino and dilatino) determinants. The contributions coming from nonzero modes cancel out, due to the nonanomalous spacetime supersymmetry. On the other hand, there are uncanceled bosonic and fermonic zero modes. Integration over four translational bosonic zero modes restores energy-momentum conservation. We have derived the fermionic zero modes in Sec. IV; there exist two gravitino zero modes and two dilatino zero modes for instanton backgrounds, and four gravitino zero modes and four dilatino zero modes for wormhole backgrounds. Thus, we find they contribute to the fermionic operator insertions in Eq. (5.6). They constitute new corrections to the low-energy effective Lagrangian coming from the spacetime instantons or wormholes. [In the case of spacetime wormholes, there may be additional wormhole creation and annihilation operators multiplied to Eq. (5.6).]

The above derivation shows that the spacetime instantons induce a nonperturbative correction to the superpotential for both S and T moduli fields. Of particular interest is the correction of the S moduli field superpotential

$$\mathcal{W}(S) \approx M_{\rm Pl}^3 \exp(-2\pi S) \ . \tag{5.7}$$

However, by itself, the scalar potential calculated from a standard supergravity

$$V_0 = \exp(K)[K^{,A\bar{A}}(\nabla_A \mathcal{W} + \mathcal{W} K_A)(\nabla_{\bar{A}} \overline{\mathcal{W}} + \overline{\mathcal{W}} K_{\bar{A}}) - 3|\mathcal{W}|^2] \tag{5.8}$$

does not show a stable minimum at a finite value of the dilaton ReS. As argued by Dine and Seiberg,[31] the vacuum energy approaches zero as the dilaton field ReS "runs away" to infinity. It is not clear if one can find nontrivial minima at the strong-coupling regime or at the weak-coupling regime at which the potential turns around positive along the dilaton field direction (a mechanism that realizes analogous to the latter has been proposed by Coleman[32] in the context of spacetime topology change; the zero cosmological constant (coupling constant associated with an identity operator) feels a deep, exponentially attractive potential).

However, it is known that there exists another mechanism to induce a nonperturbative correction to the superpotential $\mathcal{W}(S)$: gluino condensation of the hidden E_8' gauge group of heterotic superstrings.[17] This mechanism is known to have its own problems too.

On the other hand, if we consider a combined effect of both due to our string instanton and gluino condensation in the gauge subgroup $G \in \mathrm{E}_8'$ in the heterotic superstrings, the total superpotential reads

$$\mathcal{W}(S) \approx M_{\rm Pl}^3 \left[C_d \exp(-2\pi S) + C_g S \exp\left[-\frac{3}{2\beta_0} S \right] \right] \ . \tag{5.9}$$

The first and the second terms come from the dilatino and gluino condensations, respectively. The one-loop β function coefficient of the subgroup $G \in \mathrm{E}_8'$ is denoted by β_0. The prefactors C_d and C_g in Eq. (5.9) are complex-valued numerical coefficients. They may be obtained from a careful evaluation of functional integrations for dilatino and gluino bilinear operators in the spacetime string instanton and Yang-Mills instanton backgrounds. The relative phase between C_d and C_g generically takes an arbitrary value. To see if there can be a nontrivial minima of the dilaton potential, let us concentrate on physics below the dilatino and gluino condensation scales. One can calculate the dilaton potential using Eqs. (5.8) and (5.9). For the gauge subgroup $G \in \mathrm{E}_8'$ not too small, one finds that the two exponential terms in Eq. (5.9) compete with each other, if

$$|C_d / C_g| \gg 1 \quad \text{and} \quad \arg(C_d/C_g) \approx \pi \ . \tag{5.10}$$

In this case, a minimum of the potential sets in:

$$\langle S \rangle \approx \frac{4\pi\beta_0}{3} |C_d / C_g| \gg 1 \ . \tag{5.11}$$

Thus, the dilaton field gets a vacuum-expectation value in a weak-coupling regime, and the instanton approximation is well justified. Since this idea depends upon the conditions in Eq. (5.10), it remains to derive the prefactors C_d and C_g.

Next, the induced superpotential for the T moduli field takes a form

$$\mathcal{W}(T) \approx F_a^3 \exp(-2\sqrt{3}\pi T) \ . \tag{5.12}$$

The superpotential seems to break explicitly the noncompact SU(1,1) symmetry of the original Lagrangian in Eq. (1.1). In fact, the Peccei-Quinn symmetry is one of SU(1,1) symmetry. The spacetime string instantons and wormholes break not only the Peccei-Quinn symmetry but also the whole noncompact SU(1,1) invariance. As it stands, its subgroup SL(2,$\mathbb{Z}$) appears also broken. Recall that this is the modular group relating between the small and large size of the internal compactified space. Similar types of superpotential are also generated by world-sheet string instantons.[12,13] However, their effects do not break the modular symmetry since it is an exact symmetry to all orders of string loop expansions. It is because the unitarity of string perturbation theory is valid to all finite orders. Thus, the nonperturbative superpotential generated by world-sheet instantons is subject to modular symmetry. On the other hand, it is not clear whether this is also the case for the spacetime instantons. In fact, it is conceivable that modular invariance is spontaneously broken by the nonperturbative string effects. If so, this suggests that the duality of the internal Calabi-Yau manifold is spontaneously broken by spacetime string instantons. On the other hand, it is also very possible that the apparent violation of the SL(2,$\mathbb{Z}$) duality symmetry by the superpotential (5.12) is due to our approximation to use the low-energy effective Lagrangian in the weak coupling regime.

VI. DISCUSSIONS

In this paper, we studied two types of Euclidean solutions of low-energy heterotic superstring theory: instantons and wormholes. Both solutions carry nonzero Kalb-Ramond "magnetic" charges and nontrivial dilaton field configurations. The instantons may arise from both S and T moduli fields. From many physical considerations, we find the self-dual string instantons more interesting than the wormholes. This could be compared again with Yang-Mills gauge theory. Yang-Mills merons are singular classical solutions that are not stable against instanton polarizations. Furthermore, the wormholes do not provide chirality projection which is crucial to dynamical supersymmetry breaking.

Since the spacetime string instanton solutions were derived from the field-theory limit, it is of prime interest to find conformal field theories that can be identified with the instantons in the low-energy field theory limit. Not only providing a full-fledged string theoretic investigation of nonperturbative physics, such a conformal field theory of instanton will certainly shed hints to several technical issues. For example, little is known about the nature of zero modes and the collective-coordinate quantizations for topologically nontrivial configurations in string theory.

We now discuss a few further intriguing points.

A. Analogy with the invisible-axionic strings

The underlying physics of the string spacetime instantons apply equally well to the standard model with invisible axions.[10] Suppose a theory indeed has a built-in or dynamically generated Peccei-Quinn symmetry. After the Peccei-Quinn symmetry is spontaneously broken, the Goldstone boson and invisible axions show up in the spectrum of the low-energy physics down to the QCD scale. There, prominent QCD instanton fluctuation will generate an axion potential $V \approx f_\pi^2 m_\pi^2 (1-\cos A)$, and break explicitly the nonlinearly realized Peccei-Quinn symmetry $A(x) \to A(x) + \text{const}$. The axion field thus becomes massive. Now suppose we had a single axionic (thus global) string. Before the QCD instanton effects are taken into account, the axion field emanated from the string is cylindrically symmetric and is long ranged. The axion field increases by 2π at a uniform rate as one traverses a closed path around the string. Once the QCD instantons make axions massive, this symmetric axion field configuration is energetically less favorable. The axion field flux out of the string collapses down to a thin sheet with a characteristic thickness of the order of the inverse of QCD scale.

The analogy goes further. Let us imagine the QCD instantons fluctuate in the axion vacuum. Since the axion couples to the θ term $\mathrm{Tr}\int F\cdot\tilde{F}$, the anomalous equation of motion of axion field reads

$$\nabla^2 a(x) = \sum_a Q_a \delta^{(4)}(x - x_a) \ . \tag{6.1}$$

Here, the instantons are approximated by a pointlike structure and their integrally quantized topological charges are denoted by Q_a. This is precisely the structure of the instanton or wormhole solutions in Eqs. (2.14) and (3.14), for instance, Eq. (2.24). As we stressed in Sec. V, the underlying physics is the "electrically" disordered axion vacuum. Thus, the axionic strings are boundaries of axionic domain walls, much like the fundamental superstrings are boundaries of the Kalb-Ramond domain walls.

B. Self-duality of the instantons and Bogomolnyi bound

In Sec. II, we showed that the minimum action configuration of dilaton and Kalb-Ramond fields saturates the Bogomolnyi bound.[18] This is indeed due to the underlying supersymmetry, and the dilatino supersymmetry transformation in Eq. (4.1) manifests saturation of the bound. On the other hand, the S and T chiral superfield sectors of the low-energy effective field theory of superstrings could have been derived from the compactification of type II A superstrings on a six-dimensional Calabi-Yau manifold. This compactification gives $N=2$ spacetime supersymmetry in four dimensions. This implies that the S and T scalar field sector is universal and admits an extension to $N=2$ chiral supersymmetry. On the other hand, $N=2$ supersymmetry algebra is known to admit central charge extensions. Witten and Olive[33] showed that the central charge is nonzero in the topologically nontrivial sectors, i.e., soliton sectors. Furthermore, they showed that the mass bound of the solitons are saturated even on a quantum level. From this, we suspect that the saturation of Bogomolnyi bound, the quantized Kalb-Ramond "magnetic" charge as a topological charge, and the chiral fermionic zero modes could be understood from the Witten-Olive argument. The result should be nothing but a positive action theorem applied to a topologically nontrivial sector.

We argued in Sec. II that the underlying properties of the spacetime string instantons are reminiscent of the Prasad-Sommerfield limit in the non-Abelian magnetic monopoles. This can be understood in yet another way from the following argument. Suppose the dilatons get massive by turning on some dilaton potential. This corresponds to going away from the Prasad-Sommerfield limit in the non-Abelian magnetic monopole through nonzero Higgs potential. Now being massive, the dilaton exchange attraction becomes short ranged. The effective range is an inverse of dilaton Compton wavelength. Since the Kalb-Ramond gauge field interaction still remains long ranged, the neutrality of multi-instanton or anti-instantons does not hold any more. The instanton gas gets balanced like a classical Coulomb gas, interacting with long-range force. Self-dual configuration is critical in the sense that the instanton gas configuration is topological, insensitive to positions of individual instantons. The dilaton and Kalb-Ramond gauge field forces cancel each other out. In the Prasad-Sommerfield limit of non-Abelian magnetic monopoles, the exchange force between the monopoles are balanced out between the Higgs field and the gauge field. The transition from Coulomb-like instanton gas to a "topological" instanton gas corresponds to the Prasad-Sommerfield limit in our instanton configurations.

This observation naturally leads to a speculation that the self-duality of the dilaton and the Kalb-Ramond gauge fields extends to other solitonic configurations in string theories. Dabholkar and Harvey[16] wrote a very relevant paper along this direction. They imagined a closed cosmic-string state winding around one of the spatial directions compactified on a circle. They showed that the superstring tension is not renormalized in perturbation theory, due to delicate cancellation among the graviton, the dilaton, and the Kalb-Ramond field interactions. This is a particularly natural soliton configuration that the compactified string theory gives rise to. In fact, the string tension nonrenormalization follows from an $N=2$ extended supersymmetry algebra in three dimensions in which the Witten-Olive central charge is the winding number of the string around the compactified spatial direction. The winding string of Dabholkar and Harvey is not the same kind of configuration as the string instanton self-duality equation of Eq. (2.6). In fact, their winding string possesses a nontrivial Kalb-Ramond "electric" field as opposed to the "magnetic" field of our string instantons. Thus, it would be interesting to derive a Bogomolnyi bound associated with the "electric" Kalb-Ramond field, from which the winding string configuration would follow. Work along these directions is in progress.

ACKNOWLEDGMENTS

The main body of this work was completed while the author was at the Korea Advanced Institute of Science and the Aspen Center for Physics. The kind hospitality of both institutions is acknowledged. The author is grateful to M. Dine, J. A. Harvey, and G. T. Horowitz for many helpful discussions. He also thanks J. H. Schwarz, N. Seiberg, A. Strominger, and M. B. Wise for their useful comments. This research was supported in part by the National Science Foundation under Grant No. PHY 85-06686.

[1]M. Kalb and P. Ramond, Phys. Rev. D **9**, 2273 (1974); Y. Nambu, Phys. Rep. **23**, 250 (1976). See also E. Cremmer and J. Scherk, Nucl. Phys. **B72**, 117 (1974).

[2]J. Scherk and J. H. Schwarz, Phys. Lett. **81B**, 118 (1974); T. Yoneya, Prog. Theor. Phys. **51**, 1907 (1974).

[3]X. Abrikosov, Zh. Eksp. Teor. Fiz. **32**, 1442 (1957) [Sov. Phys. JETP **5**, 1174 (1957)]; H. B. Nielsen and P. Olesen, Nucl. Phys. **B61**, 45 (1973).

[4]S.-J. Rey, Phys. Rev. D **40**, 3396 (1989).

[5]T. W. B. Kibble, J. Phys. A **9**, 1387 (1976); Phys. Rep. **67**, 183 (1980); A. S. Schwarz and Ya. S. Tyupkin, Phys. Lett. **90B**, 135 (1980); T. W. B. Kibble, G. Lazarides, and Q. Shafi, *ibid.* **113B**, 237 (1982); D. Olive and N. Turok, *ibid.* **99B**, 113 (1981).

[6]A. M. Polyakov, Phys. Scr. **T15**, 191 (1987), Sec. 4; R. B. Pearson, Phys. Rev. D **26**, 2013 (1982); P. Orland, Nucl. Phys. **B205** [FS5], 107 (1982); see also S.-J. Rey, *ibid.* **B319**, 765 (1989).

[7]E. Witten, Phys. Lett. **149B**, 351 (1984); R. Rohm and E. Witten, Ann. Phys. (N.Y.) **170**, 454 (1986).

[8]M. Dine and N. Seiberg, Phys. Rev. Lett. **57**, 2625 (1986).

[9]E. Martinec, Phys. Lett. B **171**, 441 (1986).

[10]J.-E. Kim, Phys. Rev. Lett. **43**, 103 (1979); H. Georgi, L. Hall, and M. B. Wise, Nucl. Phys. **B192**, 409 (1981); M. Dine, W. Fischler, and M. Srednicki, Phys. Lett. **104B**, 199 (1981).

[11]P. Candelas, G. T. Horowitz, A. Strominger, and E. Witten, Nucl. Phys. **B258**, 46 (1984).

[12]M. Dine, N. Seiberg, X.-G. Wen, and E. Witten, Nucl. Phys. **B278**, 769 (1987); **B289**, 319 (1987).

[13]J. Ellis, C. Gomez, D. V. Nanopoulos, and M. Quirós, Phys. Lett. B **173**, 59 (1986).

[14]E. Witten, Phys. Lett. **153B**, 243 (1984).

[15]D. Montonen and D. Olive, Phys. Lett. **72B**, 117 (1977).

[16]A. Dabholkar and J. A. Harvey, Phys. Rev. Lett. **63**, 478 (1989).

[17]J. P. Derendinger, L. E. Ibáñez, and H. P. Nilles, Phys. Lett. **155B**, 65 (1985); M. Dine, R. Rohm, N. Seiberg, and E. Witten, *ibid.* **156B**, 461 (1985).

[18]E. B. Bogomolnyi, Yad. Fiz. **24**, 861 (1976) [Sov. J. Nucl. Phys. **24**, 449 (1976)]; S. Coleman *et al.*, Phys. Rev. D **15**, 544 (1977).

[19]S.-J. Rey, Phys. Rev. D **39**, 3185 (1989).

[20]L. F. Abbott and M. B. Wise, Nucl. Phys. **B325**, 687 (1989); S. Coleman and K. Lee, *ibid.* **B329**, 387 (1990).

[21]R. Sorkin, Phys. Rev. Lett. **51**, 87 (1983); D. Gross and M. J. Perry, Nucl. Phys. **B226**, 29 (1983).

[22]S.-J. Rey, Nucl. Phys. **B336**, 146 (1990).

[23]A. Hosoya and W. Ogura, Phys. Lett. B **225**, 117 (1989); A. Gupta, J. Hughes, J. Preskill, and M. B. Wise, Nucl. Phys. **B333**, 195 (1990).

[24]T. Koikawa and M. Yoshimura, Phys. Lett. B **189**, 29 (1987); M. Yoshimura, Prog. Theor. Phys. **81**, 576 (1989).

[25]S. Giddings and A. Strominger, Phys. Lett. B **230**, 46 (1989).

[26]L. Romans and N. Warner, Nucl. Phys. **B273**, 320 (1986); A. Strominger, *ibid.* **B274**, 253 (1986).

[27]E. Witten, Nucl. Phys. **B185**, 513 (1981), especially Sec. 8.

[28]K. Konishi, N. Magnoli, and H. Panagopoulis, Nucl. Phys. **B300**, 201 (1988); **B323**, 441 (1989).

[29]S.-J. Rey (in preparation).

[30]M. Dine and N. Seiberg, Nucl. Phys. **B273**, 109 (1986).

[31]M. Dine and N. Seiberg, Phys. Lett. **162B**, 473 (1985).

[32]S. Coleman, Nucl. Phys. **B310**, 643 (1989).

[33]E. Witten and D. Olive, Phys. Lett. **78B**, 97 (1978).